ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ

Πρόγραμμα Μεταπτυχιακών Σπουδών
Μαθηματικά και Σύγχρονες Εφαρμογές

ΕΠΕΞΕΡΓΑΣΙΑ ΕΙΚΟΝΑΣ ΚΑΙ ΕΞΑΓΩΓΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΓΝΩΡΙΣΜΑΤΩΝ - ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Διπλάρη Σπύρου

Επιβλέπων καθηγητής: Όμηρος Ράγγος

Ιούνιος 2010

Περιεχόμενα

Πρόλογος 3

1 **Βασικές διαδικασίες επεξεργασίας εικόνας** 4

1.1 Γενική επισκόπηση 4

1.2 Ιστογράμματα 4

1.3 Επεξεργασία μεμονωμένου σημείου 5

1.4 Επεξεργασία σε ομάδες σημείων 11

2 **Εντοπισμός ακμών** 16

2.1 Γενική επισκόπηση 16

2.2 Βασικοί ανιχνευτές ακμών 16

3 **Επεξεργασία περιοχών συχνότητας** 27

3.1 Γενική επισκόπηση 27

3.2 Μετασχηματισμοί Fourier 27

3.3 Ο διακριτός μετασχηματισμός Fourier 33

3.4 Ιδιότητες του μετασχηματισμού Fourier 39

4 **Περιγραφή αντικειμένων** 42

4.1 Γενική επισκόπηση 42

4.2 Περιγραφή συνόρου 42

5 **Τεχνητά Νευρωνικά Δίκτυα** 60

5.1 Γενική επισκόπηση 60

5.2 Αρχιτεκτονική Νευρωνικών Δικτύων 60

5.3 Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων 65

5.4 Αλγόριθμος Backpropagation 65

6 **Αναγνώριση αριθμών** 71

Πρόλογος

Για εναν βιολογικό οργανισμό, το οπτικό μέσο με το οποίο δέχεται εικόνες από το περιβάλλον του είναι το μάτι. Αυτές οι εικόνες εισέρχονται στον εγκέφαλο ο οποίος τις επεξεργάζεται και παράγει διάφορα συμπεράσματα για το περιεχόμενο αυτών. Για έναν ηλεκτρονικό υπολογιστή το οπτικό μέσο με το οποίο δέχεται εικόνες από το περιβάλλον είναι συνήθως μια φωτογραφική μηχανή ή μια μηχανή λήψης βίντεο. Ο εγκέφαλος που θα επεξεργαστεί αυτές τις εικόνες είναι κάποιο λογισμικό που έχει κατασκευαστεί για να παράγει συγκεκριμένα συμπεράσματα πάνω στις εικόνες που δέχεται κάθε φορά. Για έναν άνθρωπο είναι πολύ απλό παρατηρώντας μια εικόνα να διακρίνει πολλά χαρακτηριστικά της και να περιγράψει το περιεχομενό της. Για έναν ηλεκτρονικό υπολογιστή δεν είναι τόσο απλή διαδικασία. Έχουν αναπτυχθεί και συνεχίζουν να αναπτύσσονται τεχνικές με τις οποίες ένα σύστημα μπορεί να εντοπίζει διάφορα χαρακτηριστικά μέσα σε μια εικόνα. Κάποιες βασικές τεχνικές θα αναφερθούν στις επόμενες παραγράφους. Το περιβάλλον όμως μεταβάλλεται και συνεπώς τα διάφορα αντικείμενα μέσα σε αυτό δεν είναι πάντα ίδια. Έτσι πολλές φορές η κατασκευή κάποιου αλγορίθμου, ο οποίος θα δίνει καλή απάντηση σε κάποιο πρόβλημα που έχει σχέση με την εξαγωγή χαρακτηριστικών μιας εικόνας, είναι δύσκολη. Σε αυτό το σημείο σημαντική είναι η βοήθεια που προσφέρουν τα Τεχνητά Νευρωνικά Δίκτυα με την ικανότητα μίμησης, μάθησης και μνήμης που διαθέτουν. Η επεξεργασία εικόνας σε συνδυασμό με τα Τεχνητά Νευρωνικά Δίκτυα παράγουν αποτελέσματα με πολύ ικανοποιητικό βαθμό επιτυχίας.

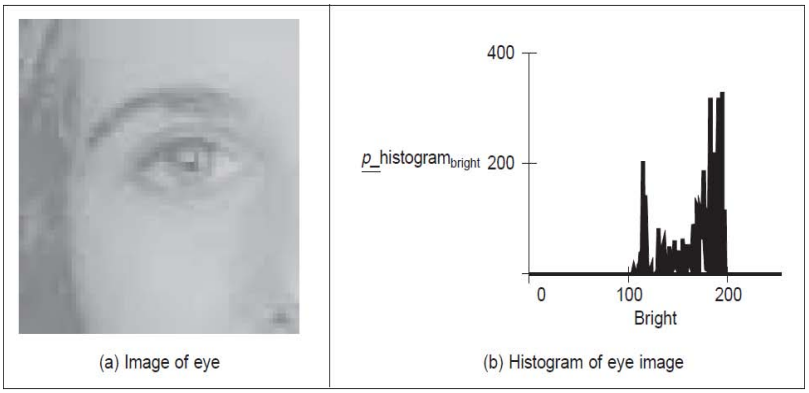
1 Βασικές διαδικασίες επεξεργασίας εικόνας

1.1 Γενική επισκόπηση

Σε αυτή την ενότητα αρχικά θα περιγράψουμε την επεξεργασία αλλαγής της φωτεινότητας μιας εικόνας χρησιμοποιώντας το ιστογραμμά της. Στη συνέχεια θα εξετάσουμε διάδικασίες με τις οποίες θα μπορούμε να χειριστούμε τις εικόνες ετσι ώστε να μπορέσουμε να αλλάξουμε το ιστογραμμά τους. Αυτές οι διαδικασίες μπορούν να μετατοπίσουν και να αλλάξουν την αρχική μορφή του ιστογράμματος, κάνοντας την αρχική εικόνα πιο φωτεινή ή πιο σκοτεινή με διάφορους τρόπους. Επίσης θα παρουσιάσουμε τεχνικές κάτω ορίου με τις οποίες η αρχική εικόνα μετατρέπεται σε εικόνα δύο χρωμάτων, του λευκού και του μαύρου. Αυτές οι τεχνικές αφορούν επεξεργασία σε ένα μεμονομένο σημείο. Ύστερα θα περάσουμε σε τεχνικές που αφορούν ομάδες σημείων με τη χρήση ενός προκαθορισμένου προτύπου. Πιο συγκεκριμένα θα ασχοληθούμε με τα πρότυπα του Gauss. Θα δείξουμε πως οι αυτές οι τεχνικές μπορούν να μειώσουν το θόρυβο σε μία εικόνα, όπου αυτό πολλές φορές είναι ένα πλεονέκτημα για τις τεχνικές εξαγωγής χαρακτηριστικών της εικόνας. Συνήθως αυτές οι τεχνικές που αναφέρθηκαν πιο πάνω, γίνονται στην αρχή της επεξεργασίας για να διαμορφώσουν ένα περιβάλλον που θα μπορούν να εφαρμοστούν πιο εύκολα οι τεχνικές εξαγωγής χαρακτηριστικών της εικόνας.

1.2 Ιστογράμματα

Το ιστόγραμμα μας δείχνει πως τα μεμονωμένα επίπεδα φωτεινότητας εμφανίζονται πάνω σε μία εικονα. Η αντίθεση σε μια εικόνα μετριέται από ένα σύνολο επιπέδων φωτεινότητας. Το ιστόγραμμα περιγράφει τον αριθμό των εικονοκυττάρων, που έχουν το ίδιο επίπεδο φωτεινότητας, κατά μήκος του συνόλου των επιπέδων φωτεινότητας. Για εικονοκύτταρα που αντιστοιχούν σε 8-bit, τα επίπεδα του φωτός παίρνουν ακέραιες τιμές απο το διάστημα [0,255] όπου το 0 αντιστοιχεί στο μαύρο ενώ το 255 στο λευκό. Το Σχήμα 1 δείχνει μια εικόνα με ένα μάτι μαζί με το ιστογραμμά της. Το ιστόγραμμα στο Σχήμα 1b δείχνει ότι δεν χρησιμοποιούνται όλες οι γκρίζες περιοχές. Το ιστόγραμμο έχει διαμορφωθεί μεταξύ των φωτεινών περιοχών 100 και 200. Οι περιοχές που είναι κοντά στο 100 έχουν σχέση με τις σκοτεινές περιοχές της εικόνας, όπως είναι τα μαλλιά και η ίριδα του ματιού ενώ τα φωτεινοτερα σημεία που αφορούν κυρίως το δέρμα είναι κοντά στο 200. Αν η εικόνα ήταν πιο σκοτεινή, τότε το ιστόγραμμα θα ήταν συγκεντρωμένο κοντά στο μηδέν. Αντίθετα αν η εικόνα ήταν πιο φωτεινή, τότε το ιστόγραμμα θα ήταν συγκεντρωμένο κατα μήκος των φωτεινών περιοχών.



Σχήμα 1: Ενα πρόσωπο και το ιστόγραμμα του.

Αυτό το ιστόγραμμα μας δείχνει ότι δεν έχουμε χρησιμοποιήσει όλα τα επίπεδα του γκρίζου. Συνεπώς μπορούμε να 'τεντώσουμε' το ιστόγραμμα προκειμένου να τα χρησιμοποιήσουμε όλα. Με αυτόν τον τρόπο η εικόνα θα φαίνεται πιο καθαρή. Το να κάνουμε την εικόνα πιο καθαρή δεν εξυπηρετεί μόνο λόγους αισθητικής, αλλά μας βοηθά να παρατηρήσουμε διάφορα χαρακτηριστικά της εικόνας με μεγαλύτερη λεπτομέρεια και αναγνωρίσουμε στοιχεία της που πρίν δεν ηταν τόσο ξεκάθαρα. Το ιστόγραμμα μπορεί επίσης να φανερώσει αν υπάρχει θόρυβος στην εικόνα, εάν το ιδανικό ιστόγραμμα είναι ήδη γνωστό. Είναι πολύ πιθανό να θελήσουμε να απομακρύνουμε τον θόρυβο από την εικόνα, όχι μόνο για να δώσουμε μια καλύτερη εμφάνιση σε αυτή, αλλά για να διαμορφώσουμε ένα περιβάλλον στο οποίο οι διάφορες τεχνικές επεξεργασίας εικόνας να μπορούν να χρησιμοποιηθούν πιο εύκολα και να δώσουν αποτελέσματα με μεγαλύτερη ακρίβεια. Πιο κάτω παρουσιάζεται τμήμα κώδικα που μπορεί να χρησιμοποιηθεί για την κατασκεύη του ιστογράμματος [1].

```
⋮
for bright = 0 : 255
    pixels_at_level(bright)=0
    for x = 0 : colums − 1
        for y = 0 : rows − 1
            if Image(x, y) == bright
               pixels_at_level(bright) = pixels_at_level(bright) + 1
            end
        end
    end
end
⋮
```

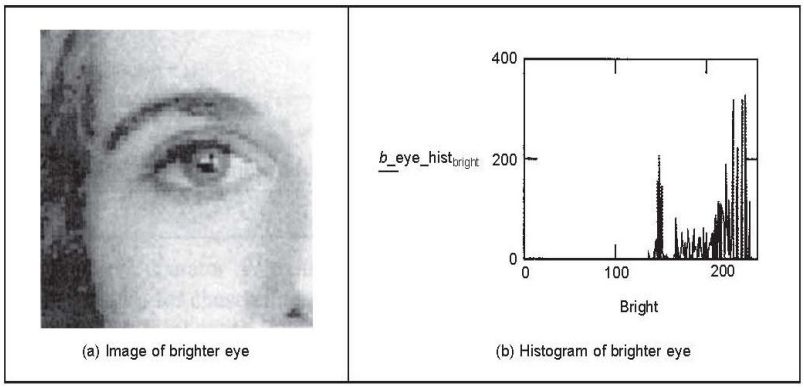
1.3 Επεξεργασία μεμονωμένου σημείου

1.3.1 Βασική επεξεργασία μεμονωμένου σημείου

Βασικές διαδικασίες στην επεξεργασία εικόνας είναι εκείνες που αφορούν την επεξεργασία σε ενα σημείο. Πιο συγκεκριμένα η τιμή του κάθε εικονοκύτταρου αντικαθίσταται από μια νέα τιμή η οποία παράγεται από την παλιά τιμή του. Εαν θέλουμε να αυξήσουμε την φωτεινότητα της εικόνας τότε αυτό γίνεται εύκολα πολλαπλασιάζοντας τις τιμές όλων των εικονοκύτταρων με τον ίδιο αριθμό. Αντίθετα εαν θέλουμε να μειώσουμε την φωτεινότητα μπορούμε να διαιρέσουμε τις τιμές όλων των εικονοκύτταρων με τον ίδιο αριθμό. Εάν θέλουμε γενικά να μεταβάλουμε την φωτεινότητα μιας εικόνας μπορούμε να χρησιμοποιήσουμε τον πιο κάτω γραμμικό μετασχηματισμό όπου N είναι η νέα εικόνα, O είναι η αρχική εικόνα και l ,k είναι πραγματικοί αριθμοί.

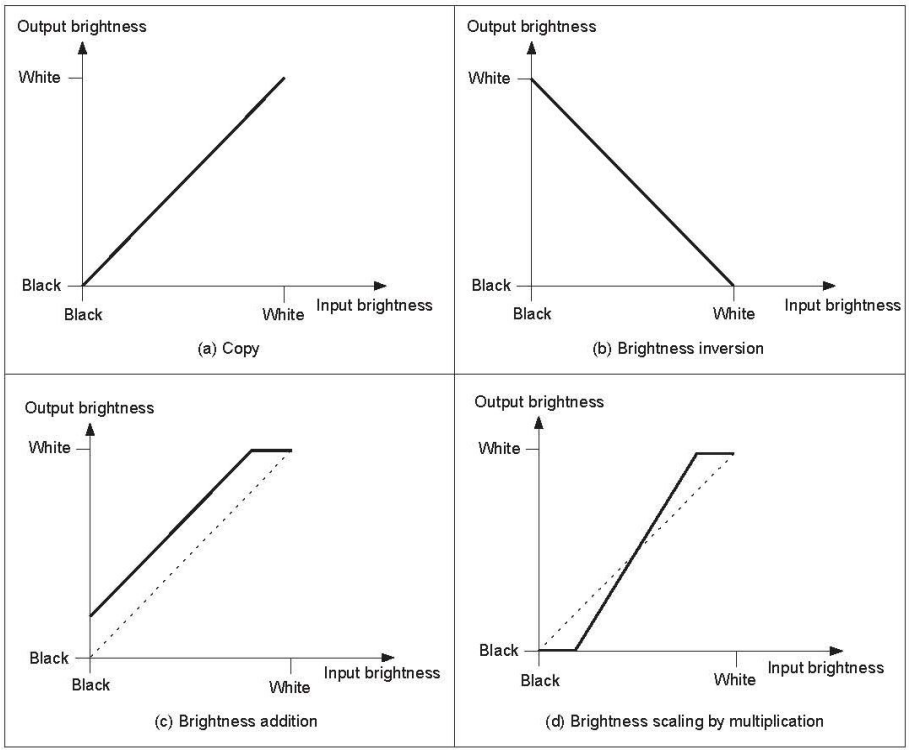
$$N_{x,y} = K \times O_{x,y} + l \qquad \forall\, x, y \,\epsilon\, 1, N \qquad (1.1)$$

Η τιμή του l ελέγχει τη γενική φωτεινότητα της εικόνας και είναι η μικρότερη τιμή που μπορεί να έχει ένα εικονοκύτταρο στη νέα εικόνα. Η τιμή k ελέγχει την αντίθεση. Όσο μεγαλύτερη είναι η τιμή του k τόσο αυξάνεται και η αντίθεση στην παραγόμενη εικόνα. Αυτή η διαδικασία παρουσιάζεται στο Σχήμα 2. Οπότε εάν στην αρχική εικόνα με το πρόσωπο (Σχήμα 1) εφαρμόσουμε τον πιο πάνω γραμμικό μετασχηματισμό με k=1.2 και l=10 τότε η εικόνα που θα προκύψει θα είναι πιο φωτεινή (Σχήμα 2a) και με καλύτερη αντίθεση, αφού τα φωτεινά σημεία της νέας εικόνας είναι πιο κοντά στην περιοχή του λευκού. Αυτές οι μετατροπές μπορούν να παρατηρηθούν στο ιστόγραμμα της νέας εικόνας (Σχήμα 2b). Η μεταβολή της αντίθεσης της εικόνας μπορεί να



Σχήμα 2: Πιο φωτεινή εικόνα

παρουσιαστεί και γραφικά, δείχνοντας τη μεταβολή του λευκού σε μαύρο και αντίστροφα (Σχήμα 3). Στο Σχήμα 3a η παραγόμενη εικόνα είναι ένα ακριβές αντίγραφο της αρχίκής εικόνας. Στο Σχήμα 3b γίνεται αντιστροφή του μαύρου και του λευκού. Δηλαδή οι φωτεινές περιοχές στην αρχική εικόνα θα είναι σκοτεινές και αντίστροφα. Το Σχήμα 3c δείχνει ότι μετατοπίζουμε την απεικόνιση της αρχικής εικόνας προσθέτωντας ένα θετικό αριθμό. Παρατηρούμε ότι η μέγιστη τιμή της νέας εικόνας δεν μπορεί να ξεπεράσει το 255 και, έτσι, σε αυτό το σημείο η γραφική παράσταση γίνεται παραλληλη με τον οριζόντιο άξονα. Το Σχήμα 3d είναι αποτέλεσμα μετατόπισης της αρχικής εικόνας μέσω διαίρεσης η αφαίρεσης [1].



Σχήμα 3: Μεταβολή Λευκού-Μαύρου

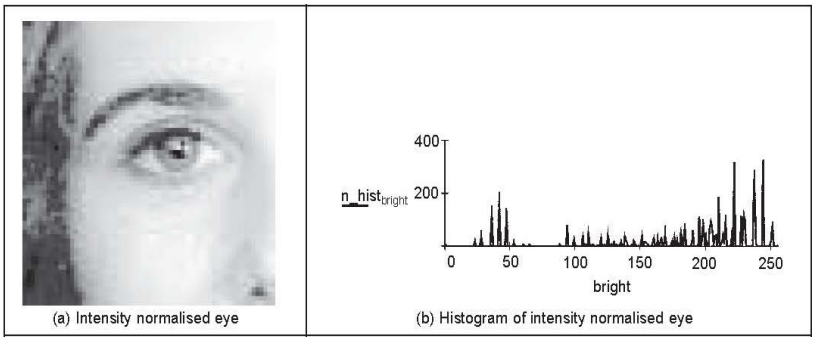
1.3.2 Κανονικοποίηση ιστογράμμων

Βασικός στόχος είναι το αρχικό ιστόγραμμα να 'τεντωθεί' και να μετατοπιστεί προκειμένου να καλύψει και τα

256 διαθέσιμα επίπεδα. Εαν το ιστόγραμμα της αρχικής εικόνας O ξεκινά από το O_{min} και καταλήγει στο O_{max} φωτεινό επίπεδο, τότε μπορούμε να μετατρέψουμε κλιμακωτά την εικόνα ετσι ώστε τα εικονοκύτταρα στην νέα εικόνα N να εκτείνονται μεταξύ ενός κάτω ορίου N_{min} και ενός άνω ορίου N_{max} . Αυτή η διαδικασία επιτυχγάνεται με τη χρήση της πιο κάτω σχέσης:

$$N_{x,y} = \frac{N_{max} - N_{min}}{O_{max} - O_{min}}(O_{x,y} - O_{min}) + N_{min} \qquad (1.2)$$

Το αποτέλεσμα αυτού του μετασχηματισμού παρουσιάζεται στο Σχήμα 4 και μπορεί να συγκριθεί με την αρχική εικόνα στο Σχήμα 1. Στο Σχήμα 4 φαίνεται καθαρά το τέντωμα που έγινε στο ιστόγραμμα (Σχήμα 4b) της αρχικής εικόνας. Το αποτέλεσμα είναι το Σχήμα 4α το οποίο έχει καλύτερη αντίθεση και είναι πιο ευδιάκριτο. Το ιστόγραμμα της κανονικοποιημένης εικόνας περιλαμβάνει όλα τα διαθέσιμα επίπεδα. Θα μπορούσαμε να ορίσουμε ένα διάστημα μεταξύ του 0 και του 255 που θα μας εξυπηρετούσε [1,2].



Σχήμα 4: Κανονικοποιημένη εικόνα

1.3.3 Εξισορρόπηση ιστογράμμων

Η εξισορρόπηση των ιστογράμμων είναι μια μη γραμμική διαδικασία που στοχεύει στο να δώσει έμφαση στη φωτεινότητα της εικόνας με έναν τρόπο που ταιριάζει ιδιαίτερα στην ανθρώπινη οπτική ανάλυση. Η εξισορρόπηση ιστογράμμων στοχεύει να αλλάξει μια εικόνα με τέτοιο τρόπο έτσι ώστε να παραχθεί μια νέα εικόνα με πιό επίπεδο ιστόγραμμα όπου όλα τα επίπεδα να είναι ισοπίθανα. Προκειμένου να αναπτυχθεί ο αλγόριθμος που θα πετυχαίνει την εξισορρόπηση του ιστογράμματος, θα πρέπει πρώτα να εξετάσουμε το ιστόγραμμα. Έστω ότι το ιστόγραμμα της αρχικής εικόνας αποτελείται απο M επίπεδα. Θεωρούμε ότι ο αριθμός των σημείων για κάθε επίπεδο είναι $O(l)$ για την αρχική εικόνα και $N(l)$ για την εικόνα που θα προκύψει από την εξισορρόπηση του ιστογράμματος, όπου $l \epsilon (0, M)$. Η αρχική και η τελική εικόνα έχουν ίδιο μέγεθος, έστω $n \times m$, οπότε πρέπει να ισχύει η πιο κάτω σχέση.

$$\sum_{l=0}^{M} O(l) = \sum_{l=0}^{M} N(l) \qquad (1.3)$$

Επίσης αυτό θα πρέπει να είναι το ίδιο και για ένα αυθαίρετα επιλεγμένο επίπεδο p δεδομένου ότι στοχεύουμε στην παραγωγή μιας εικόνας με ένα ομοιόμορφα επίπεδο ιστόγραμμα. Έτσι το αθροιστικό ιστόγραμμα της αρχικής εικόνας μέχρι το επίπεδο p θα πρέπει να μετασχηματιστεί έτσι ώστε να καλύψει το νέο ιστόγραμμα μέχρι το επίπεδο q :

$$\sum_{l=0}^{p} O(l) = \sum_{l=0}^{q} N(l) \qquad (1.4)$$

Επειδή το νέο ιστόγραμμα είναι ομοιόμορφα επίπεδο, τότε το αθροιστικό ιστόγραμμα πάνω από το επίπεδο p θα πρέπει να είναι ένα κλάσμα του ολικού αθροίσματος. Οπότε ο αριθμός των σημείων ανά επίπεδο στην παραγόμενη εικόνα είναι ο λόγος του συνολικού αριθμού των σημείων αυτής προς το μήκος του διαστήματος των επιπέδων της:

$$N(l) = \frac{n \times m}{N_{max} - N_{min}} \qquad (1.5)$$

Οπότε το αθροιστικό ιστόγραμμα της νέας εικόνας είναι:

$$\sum_{l=0}^{q} N(l) = q \times \frac{n \times m}{N_{max} - N_{min}} \qquad (1.6)$$

Απο την ισότητα της σχέσης 1.4 αυτό είναι ίσο με το αθροιστικό ιστόγραμμα της αρχικής εικόνας:

$$\sum_{l=0}^{q} O(l) = q \times \frac{n \times m}{N_{max} - N_{min}} \qquad (1.7)$$

Έτσι έχουμε μια χαρτογράφηση για τα εικονοκύτταρα της νέας εικόνας στο επίπεδο q από τα εικονοκύτταρα της αρχικής εικόνας στο επίπεδο p :

$$q = \frac{N_{max} - N_{min}}{n \times m} \times \sum_{l=0}^{q} O(l) \qquad (1.8)$$

Αυτό δίνει μια συνάρτηση χαρτογράφησης που παράγει μια εικόνα που έχει ένα περίπου επίπεδο ιστόγραμμα. Η συνάρτηση χαρτογράφησης δίνεται αν εκφράσουμε την Σχέση 1.8 ως μια συνάρτηση εξισορρόπησης E η οποία είναι συνάρτηση του επιπέδου q και της αρχικής εικόνας O :

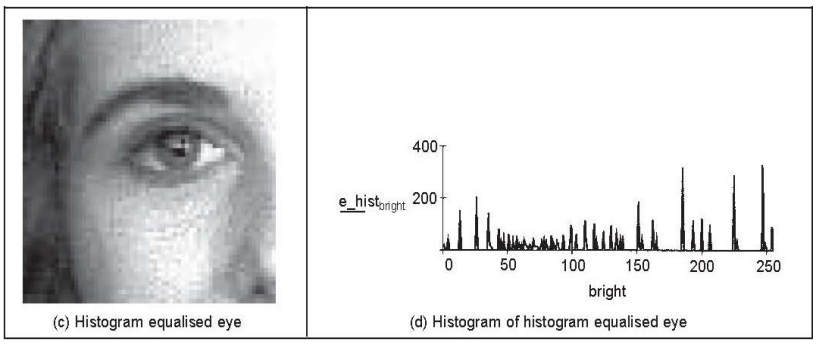
$$E(q, O) = \frac{N_{max} - N_{min}}{n \times m} \times \sum_{l=0}^{q} O(l) \qquad (1.9)$$

Η παραγόμενη εικόνα είναι :

$$N_{x,y} = E(O_{x,y}, O) \qquad (1.10)$$

Το αποτέλεσμα της πιο πάνω διαδικασίας φαίνεται στο Σχήμα 5. Στην εξισορροπημένη εικόνα (Σχήμα 5c) διακρίνονται πολύ καλύτερα τα διάφορα χαρακτηριστικά γνωρίσματα, ειδικά γύρο από τα μάτια σε σύγκριση με την αρχική εικόνα (Σχήμα 1α) . Το ιστόγραμμα (Σχήμα 5d) αποκαλύπτει τη μη γραμμική διαδικασία χαρτογράφησης σύμφωνα με την οποία στο λευκό και στο μαύρο δεν αντιστοιχεί η ίδια βαρύτητα. Η απόδοσή της είναι πολύ πειστική αφού η χαρτογράφηση που γίνεται είναι πολύ φιλική στην ανθρώπινη όραση. Εάν ένας γραμμικός μετασχηματισμός εφαρμοστεί στην αρχική εικόνα τοτε το εξισορροπημένο ιστόγραμμα θα είναι το ίδιο. Εάν αντικαταστήσουμε τις τιμές των εικονοκυττάρων με αυτές που υπολογίζονται σύμφωνα με την Σχέση 1.1 τοτε το εξισορροπημένο ιστόγραμμα θα είναι πάλι το ίδιο. Μια εναλλακτική ερμηνεία είναι ότι εάν εξισορροπήσουμε κάποια εικόνα τότε δεν χρειάζεται να ανησυχούμε για τους μετασχηματισμούς φωτεινότητας στην αρχική εικόνα. Ωστόσο ο θόρυβος στην εικόνα, ο οποίος δημιουργείται κατά τη λήψη της, μπορεί να επηρεάσει την μορφή του

ιστογράμματος της αρχικής εικόνας και στη συνέχεια θα επηρεάσει και αυτό της εξισοροπημένης μορφής. Οπότε το εξισοροπημένο ιστόγραμμα μιας εικόνας χωρίς θόρυθο δεν είναι το ίδιο με αυτό της ίδιας εικόνας στην οποία έχει προστεθεί θόρυβος. Δεν μπορούμε να αποφύγουμε τον θόρυβο στα ηλεκτρικά συστήματα, μπορούμε όμως με διάφορες τεχνικές να τον μειώσουμε. Η εξισοροπηση των ιστογράμμων βρίσκει λίγη χρήση σε γενικά συστήματα επεξεργασίας εικόνας, αν και μπορεί να είναι δραστική σε πολύ εξειδικευμένες εφαρμογές. Ένα μειονέκτημα της εξισορόπησης των ιστογράμμων είναι ότι δεν είναι γραμμική διαδικασία οπότε δεν μπορούμε να επιστρέψουμε στο αρχικό ιστόγραμμα ύστερα απο αυτή. Αντίθετα η κανονικοποίηση των ιστογράμμων είναι μια γραμμική διαδικασία και μπορούμε εύκολα να επιστρέψουμε στο πρωτότυπο ιστόγραμμα. Για αυτούς τους λόγους η κανονικοποίηση των ιστογράμμων χρησιμοποιείται πιο συχνά όταν χρειάζονται να γίνουν μορφοποιήσεις σε αυτά [1,2].



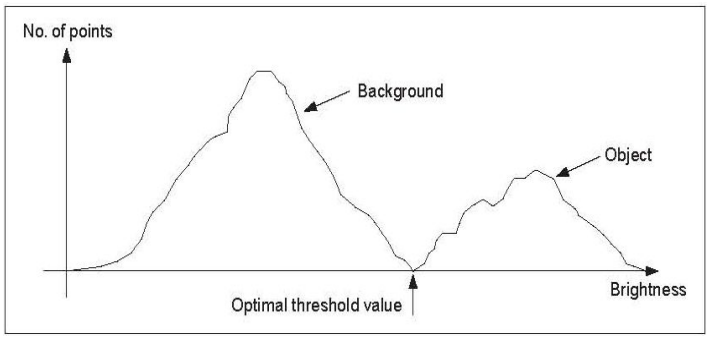
Σχήμα 5: Εξισορροπημένη εικόνα

1.3.4 Κατώτατο όριο

Ο μηχανισμός του κατώτατου ορίου επιλέγει εικονοκύτταρα τα οποία έχουν μια συγκεκριμένη τιμή ή βρίσκονται μέσα σε ένα καθορισμένο διάστημα. Αυτός ο μηχανισμός μπορεί να χρησιμοποιηθεί για να βρεθούν αντικείμενα μέσα σε μία εικόνα, αν για αυτά είναι γνωστά τα επίπεδα φωτεινότητας. Υπάρχουν δύο μορφές αυτού του μηχανισμού: ο ομοιόμορφος και προσαρμοστικός μηχανισμός του κατώτερου ορίου. Στον ομοιόμορφο μηχανισμό τα εικονοκύτταρα που έχουν τιμή μεγαλύτερη από έναν οριακό αριθμό(κατώφλι) γίνονται λευκά και αυτά που είναι έχουν τιμή κάτω απο αυτόν τον αριθμό γίνονται μαύρα. Στο Σχήμα 6 φαίνεται το αποτέλεσμα αυτού του μηχανισμού ο οποίος εφαρμόστηκε στο Σχήμα 1 αφού έχουμε ορίσει ώς κατώτερο όριο το 160. Όσα εικονοκύτταρα έχουν τιμή πάνω από 160 γίνονται λευκά και αυτά που έχουν τιμή κάτω από το 160 γίνονται μαύρα. Με αυτή τη διαδικασία τα μέρη που έχουν σχεση με το δέρμα του προσώπου ξεχωρίζουν από το υπόλοιπο περιβάλλον. Αυτό μπορεί επομένως να παρέχει έναν τρόπο απομόνωσης κάποιων σημείων με ιδιαίτερο ενδιαφέρον. Ο ομοιόμορφος μηχανισμός του κατώτατου ορίου απαιτεί τη γνώση των επιπέδων φωτεινότητας

Σχήμα 6: Ομοιόμορφος μηχανισμός

του αντικειμένου που ψάχνουμε, αλλιώς τα χαρακτηριστικά γνωρίσματα που έχουν επιλεχθεί σαν στόχοι δεν θα μπορέσουν να εντοπισθούν μέσα στην εικόνα. Γενικά στον ομοιόμορφο μηχανισμό προσπαθούμε με πειράματα να βρούμε ένα κατάλληλο όριο που θα μπορεί να χρησιμοποιηθεί μόνο για συγκεκριμένα αντικείμενα. Υπάρχουν πιο προηγμένες τεχνικές που είναι γνωστές σαν προσαρμοστικοί μηχανισμοί κατώτερου ορίου. Αυτές επιδιώκουν συνήθως να επιλέξουν μια τιμή για το κάτω όριο έτσι να μπορέσει να ξεχωρίσει ενα αντικείμενο από το υπόλοιπο περιβάλλον του. Αυτό προϋποθέτει ότι το αντικείμενο πρέπει να έχει διαφορετικά επίπεδα φωτεινότητας από το περιβάλλον του (Σχήμα 7).



Σχήμα 7: Προσαρμοστικός μηχανισμός

Η μέθοδος του Otsu είναι μία από τις διαδεδομένες τεχνικές που πετυχαίνουν τον προσαρμοστικό μηχανισμό του κάτω ορίου. Ουσιαστικά η μέθοδος του Otsu μεγιστοποιεί την πιθανότητα ώστε το κάτω όριο που επιλέχθηκε να είναι ικανό να χωρίσει την εικόνα από το περιβάλλον της. Αυτό επιτυγχάνεται με την επιλογή ενός κάτω ορίου που χωρίζει καλύτερα όλα τα εικονοκύτταρα της εικόνας σε δύο κατηγορίες. Η βάση της στηρίζεται στο κανονικοποιημένο ιστόγραμμα όπου ο αριθμός των σημείων του κάθε επιπέδου διαιρείται με τον συνολικό αριθμό των σημείων της εικόνας. Αυτό αντιπροσωπεύει την πιθανότητα ένα εικονοκύτταρο να ανοίκει σε ένα επίπεδο l :

$$p(l) = \frac{N(l)}{n \times m} \qquad (1.11)$$

Χρησιμοποιώντας αυτή την σχέση υπολογίζουμε την αθροιστική πιθανότητα και τον σταθμισμένο μέσο όρο μέχρι την επίπεδο k :

$$w(k) = \sum_{l=1}^{k} p(l) \qquad (1.12)$$

και

$$\mu(k) = \sum_{l=1}^{k} l \cdot p(l) \qquad (1.13)$$

Ο συνολικός σταθμισμένος μέσος όρος της εικόνας δίνεται από τον τύπο:

$$\mu T = \sum_{l=1}^{N_{max}} l \cdot p(l) \qquad (1.14)$$

Η διασπορά των κατηγοριών των εικονοκυττάρων ορίζεται από τον τύπο:

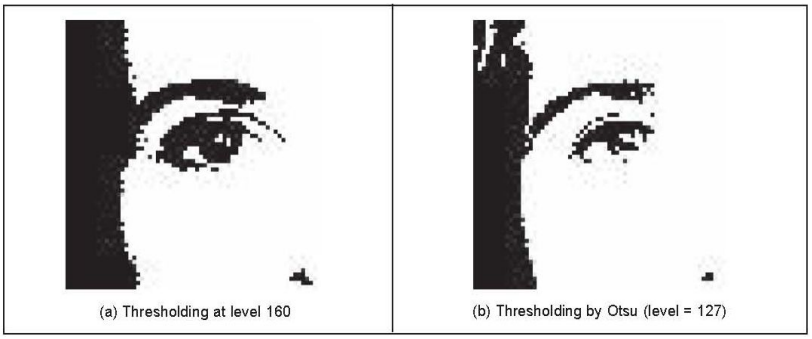
$$\sigma_\beta^2(k) = \frac{(\mu T \cdot w(k) - \mu(k))^2}{w(k)(1 - w(k))} \qquad \forall\, k \,\epsilon\, 1, N_{max} \qquad (1.15)$$

Το βέλτιστο κάτω όριο είναι το επίπεδο για το οποίο η διασπορά είναι μέγιστη και το συμβολίζουμε με T_{opt} :

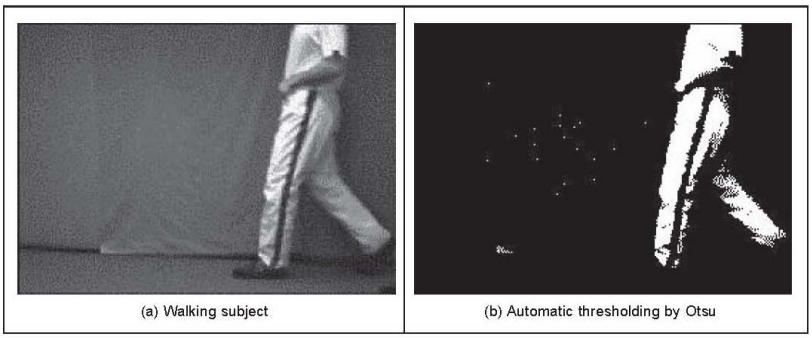
$$T_{opt} = \sigma_\beta^2(T_{opt}) = max(\sigma_\beta^2(k)) \qquad 1 \leq k < N_{max} \qquad (1.16)$$

Μια σύγκριση του ομοιόμορφου και του προσαρμοστικού μηχανισμού δίνεται στο Σχήμα 8. Το κάτω όριο που επιλέχθηκε από την μέθοδο του Otsu είναι ελαφρώς χαμηλότερο από την τιμή που επιλέχθηκε χειροκίνητα, οπότε η εικόνα ύστερα από την επεξεργασία της έχασε κάποια χαρακτηριστικά γύρω από το μάτι. Ωστόσο η επιλογή

του κάτω ορίου με την μέθοδο του Otsu είναι αυτόματη και αυτό είναι ένα πλεονέκτημα για τις διαδικασίες επεξεργασίας εικόνας. Βέβαια το όριο του Otsu μπορούμε να το τροποποιήσουμε όταν εργαζόμαστε πάνω σε συγκεκριμένο υλικό. Για παράδειγμα, μπορούμε να προσθέσουμε η να αφαιρέσουμε σε αυτό ένα σταθερό αριθμό c ή έναν αριθμό που θα μεταβάλλεται όπως $(10\% \cdot T_{opt})$. Στο Σχήμα 9α προκύπτει η ανάγκη να απομονώσουμε την ανθρώπινη φιγούρα. Αυτό μπορεί να γίνει αυτόματα με τη μέθοδο του Otsu όπως φαίνεται στο Σχήμα 9b . Παρατηρούμε ότι υπάρχουν μερικά επιπλέον σημεία λόγω του φωτισμού τα οποία έχουν εμφανιστεί στην προκύπτουσα εικόνα μαζί με την ανθρώπινη φιγούρα. Είναι εύκολο να αφαιρεθούν τα απομονωμένα σημεία, με τεχνικές που θα παρουσιάσουμε στη συνέχεια, όμως πιο δύσκολα θα μπορέσουμε να απομακρύνουμε αυτά που είναι συνδεδεμένα. Γενικά το μέγεθος της ανθρώπινης μορφής θα μας βοηθήσει δίνοντας μας πληροφορίες για να απομακρύνουμε τα περιττά σημεία.



Σχήμα 8: Χειροκίνητη και αυτόματη επιλογή κάτω ορίου



Σχήμα 9: Αυτόματη επιλογή κάτω ορίου με τη μέθοδο του Otsu

Μέχρι τώρα έχουμε εξετάσει τις γενικές τεχνικές και μεθόδους που λειτουργούν σε ολόκληρη την εικόνα. Υπάρχουν τοπικά προσαρμοστικές τεχνικές που χρησιμοποιούνται συχνά για την ανάλυση σε εικόνες εγγράφων. Μια έρευνα που έγινε από τον Trier το 1995 συγκρίνει τις γενικές και τις τοπικές τεχνικές σχετικά με την αποδοσή τους στην ανάλυση εικόνων εγγράφων. Οι τοπικές τεχνικές είχαν καλύτερη απόδοση στην αναγνώριση σχεδίων. Γενικά αυτές οι τεχνικές έχουν μικρή χρησιμότητα στην επεξεργασία εικόνων με ποικιλία αντικείμενα. Σε τέτοιες περιπτώσεις είναι καλύτερο να χρησιμοποιούνται τεχνικές όπως η μέθοδος του Otsu οι οποίες λειτουργούν ταυτόχρονα σε όλη την εικόνα και μπορούν να εντοπίσουν πολλά αντικείμενα ίδιου ενδιαφέροντος μέσα σε αυτή. [1,3]

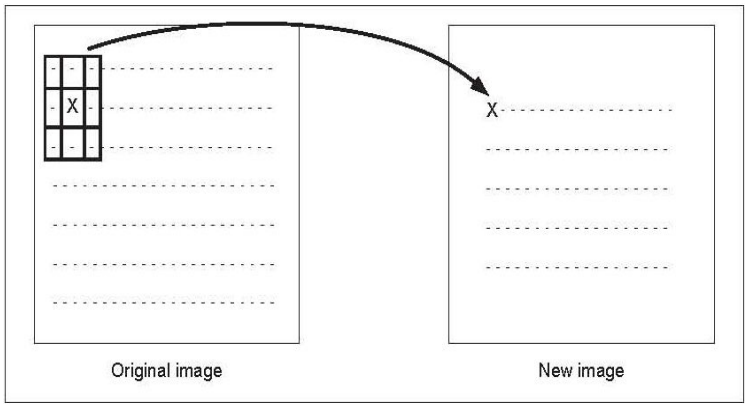
1.4 Επεξεργασία σε ομάδες σημείων

1.4.1 Συνέλιξη προτύπων

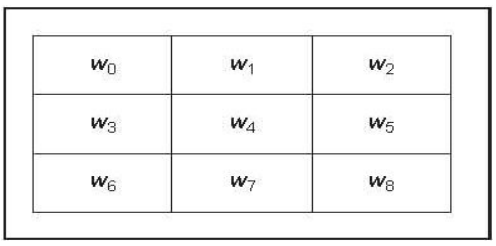
Η επεξεργασία σε ομάδες σημείων υπολογίζει νέες τιμές στα εικονοκύτταρα μιας εικόνας με τη χρήση των γειτονικών τους εικονοκυττάρων. Αυτή η ομαδική επεξεργασία χρησιμοποιεί ένα πρότυπο για να μπορέσει να

συνδυάσει γειτονικά εικονοκύτταρα σε ένα. Τα πρότυπα είναι πίνακες με βάρη τα οποία λειτουργούν σαν πολλαπλασιαστές των εικονοκυττάρων. Οι πίνακες αυτοί είναι συνήθως τετραγωνικοί $n \times n$ όπου το n είναι περιττός αριθμός για να εξασφαλιστεί ότι οι πίνακες μπορούν να τοποθετηθούν κατάλληλα στην εικόνα. Για παράδειγμα ένα πρότυπο $A_{3\times 3}$ είναι ένα πρότυπο με τρία εικονοκύτταρα μήκος και τρία εικονοκύτταρα ύψος. Νέες τιμές υπολογίζονται για τα εικονοκύτταρα με την τοποθέτηση του προτύπου στα ενδιαφερόμενα σημεία. Οι τιμές των εικονοκυττάρων πολλαπλασιάζονται με τα αντίστοιχα βάρη του προτύπου και προστίθενται στο συνολικό άθροισμα. Το άθροισμα εκτιμά μια νέα τιμή για το κεντρικό εικονοκύτταρο, πάνω στο οποίο έχει τοποθετηθεί το κέντρο του προτύπου, και αυτή η τιμή γίνεται η νέα τιμή του εικονοκυττάρου στην νέα εικόνα. Το πρότυπο μετακινείται οριζόντια και κάθετα πάνω στην εικόνα μέχρι να υπολογιστούν νέες τιμές για όλα τα εικοκύτταρα στα οποία μπορεί να τοποθετηθεί το κέντρο του προτύπου. Αυτή η διαδικασία παρουσιάζεται στο Σχήμα 10 όπου μια νέα εικόνα παράγεται από μια αρχική με συνέλιξη προτύπων. Το πρότυπο δεν μπορεί να επεκταθεί σε ολόκληρη την εικόνα, οπότε η νέα εικόνα θα είναι μικρότερη από την αρχική επειδή δεν μπορούν να υπολογιστούν νέες τιμές για το σύνορο της νέας εικόνας. Οταν το πρότυπο φθάσει στο τέλος μιας γραμμής τότε τοποθετείται στην αρχή της επόμενης γραμμής. Για μια γειτονιά διάστασης 3×3 χρειάζονται 9 πολλαπλασιαστικά βάρη w_t για να υπολογιστεί ένα σημείο στην νέα εικόνα. Για να υπολόγίσουμε ένα σημείο στη νέα εικόνα I με συντεταγμένες x, y το πρότυπο στο Σχήμα 11 λειτουργεί σε μια αρχική εικόνα O ως εξής:

$$I_{x,y} = \begin{array}{l} w_0 \cdot O_{x-1,y-1} + w_1 \cdot O_{x,y-1} + w_2 \cdot O_{x+1,y-1} + \\ w_3 \cdot O_{x-1,y} + w_4 \cdot O_{x,y} + w_5 \cdot O_{x+1,y} + \\ w_6 \cdot O_{x-1,y+1} + w_7 \cdot O_{x,y+1} + w_8 \cdot O_{x+1,y+1} \end{array} \qquad \forall\, x, y \,\epsilon\, 2, N-1 \qquad (1.17)$$



Σχήμα 10: Διαδικασία συνέλιξης προτύπων



Σχήμα 11: Πρότυπο διάστασης 3×3 με τα βάρη του

Όπως είπαμε, δεν μπορούν να αποδοθούν αριθμητικές τιμές στο σύνορο της εικόνας. Αυτό συμβαίνει διότι όταν

τοποθετούμε το κέντρο του προτύπου στο σύνορο της εικόνας, τμήμα του προτύπου πέφτει έξω από την εικόνα και δεν έχουμε καμία πληροφορία για να ακολουθήσουν οι σχετικοί υπολογισμοί. Το πλάτος του συνόρου είναι ίσο με το μισό μέγεθος του προτύπου. Για να δώσουμε τιμές στα εικονοκύταρα του συνόρου έχουμε τρείς τεχνικές:

1. Κάνε το σύνορο μαύρο.
2. Υποθέτουμε ότι η αρχή και το τέλος της εικόνας ως προς τον αξονα των x ταυτίζονται και εφαρμόζουμε την διαδικασία υπολογισμού που περιγράφτηκε πιο πάνω. Ομοίος ως προς τον άξονα των y.
3. Υπολόγισε τις τιμές των εικονοκυττάρων από μια μικρότερη περιοχή.

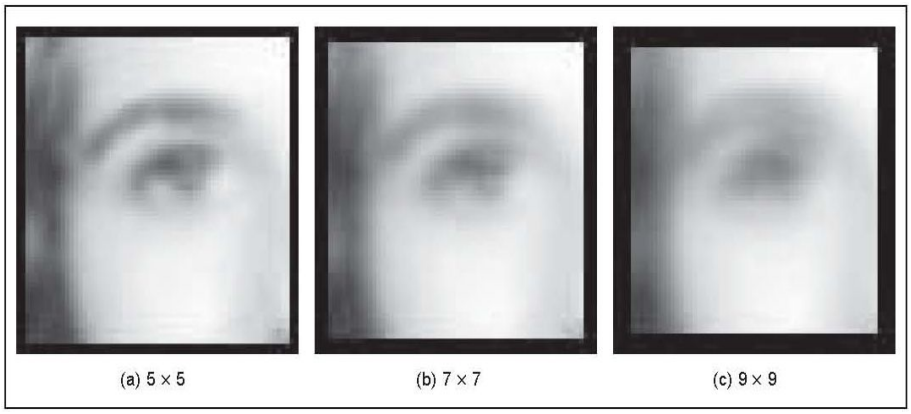
Καμία από αυτές τις τεχνικές δεν είναι βέλτιστη. Στα παραδείγματα που θα ακολουθήσουν έχει χρησιμοποιηθεί η πρώτη τεχνική. Δεν υπάρχει κάποιο κριτήριο που να επιλέγει την τεχνική που θα χρησιμοποιείται κάθε φορά. Σε πολλές εφαρμογές το αντικείμενο που ψάχνουμε στην εικόνα βρίσκεται συνήθως στο κέντρο αυτής. Σε τέτοιες περιπτώσεις οι πληροφοριές που θα χαθούν από το σύνορο της εικόνας είναι μικρού ενδιαφέροντος και επηρεάζουν ελάχιστα το συνολικό αποτέλεσμα. Ιδιαίτερα σημαντικό είναι η κατασκευή του προτύπου. Υπάρχουν διάφοροι τρόποι που το πετυχαίνουν αυτό. Ένας απλός αλλά συνηθισμένος τρόπος είναι αυτός του μέσου όρου του προτύπου. Για παράδειγμα, εάν το πρότυπο έχει διάσταση 3×3 τότε και τα 9 βάρη του προτύπου θα έχουν τιμή $1/9$ (Σχήμα 12). Το αποτέλεσμα της εφαρμογής αυτού του προτύπου φαίνεται στο Σχήμα 13. Παρατηρούμε ότι μαζί με τον θόρυβο στην εικόνα έχουν χαθεί και πολλές λεπτομέρειες αυτής όπως στα μάτια και στα φρύδια.

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

Σχήμα 12: Πρότυπο διάστασης 3×3

Σχήμα 13: Εικόνα ύστερα από φιλτράρισμα

Όπως είπαμε, το πλεονέκτημα του φιλτραρίσματος είναι η μείωση του θορύβου στην εικόνα, όμως ταυτόχρονα έχουμε και αλλοίωση των χαρακτηριστικών της. Με τη χρήση του μέσου όρου του προτύπου διατηρούμε τις χαμηλές συχνότητες και μειώνουμε τις πολύ υψηλές. Το μέγεθος του προτύπου παίζει σημαντικό ρόλο στη μείωση του θορύβου. Ένα πρότυπο μεγέθους 5×5 θα απομάκρυνε περισσότερο θόρυβο από την εικόνα αλλά θα αλλοίωνε ακόμα περισσότερο τα χαρακτηριστικά της. Ένα επιλέον μειονέκτημα των προτύπων μεγάλου μεγέθους είναι ότι χρειάζονται μεγάλο υπολογιστικό κόστος για να εφαρμοστούν σε μία εικόνα. Οπότε για λόγους ταχύτητας χρησιμοποιούνται πιο συχνά πρότυπα μεγέθους 3×3, 5×5 και 7×7. Αν υποθέσουμε ότι έχουμε ένα πρότυπο μεγέθους 9×9, θα πρέπει να πραγματοποιηθούν 81 υπολογισμοί για να υπολογιστεί ένα μόνο σημείο στην νέα εικόνα και αυτό θα απαιτούσε μεγάλο υπολογιστικό κόστος αν το μέγεθος της εικόνας ήταν μεγάλο. Βέβαια το μέγεθος του προτύπου καθορίζεται κάθε φορά από τη φύση του προβλήματος. Άν το αντικείμενο που μας ενδιαφέρει έχει μεγάλο μέγεθος χωρίς πολλές λεπτομέρειες τότε η χρήση ενός μεγάλου προτύπου θα ήταν χρήσιμη. Στο Σχήμα 14 παρουσιάζεται η χρήση τριών προτύπων μεγέθους 5×5, 7×7 και 9×9, αντίστοιχα, στην ίδια εικόνα. Το πρότυπο με μέγεθος 5×5 διατηρεί πιο πολλές λεπτομέρειες από αυτό με μέγεθος 7×7 και ακόμα πιο πολλές από αυτό με μέγεθος 9×9. Στο Σχήμα 14c έχουν διατηρηθεί θολά μόνο τα πολύ έντονα χαρακτηριστικά όπως η περιοχή των ματιών. Πρέπει να σημειωθεί ότι όσο μεγαλύτερης διάστασης πρότυπο χρησιμοποιούμε, τόσο πιο πολλές πληροφορίες χάνουμε από το σύνορο της εικόνας και καμία από τις τρείς πιο πάνω τεχνικές που αναφέραμε δεν μπορούν να αναπληρώσουν με ακρίβεια τα χαμένα δεδομένα.



Σχήμα 14: Χρήση τριών διαφορετικών προτύπων

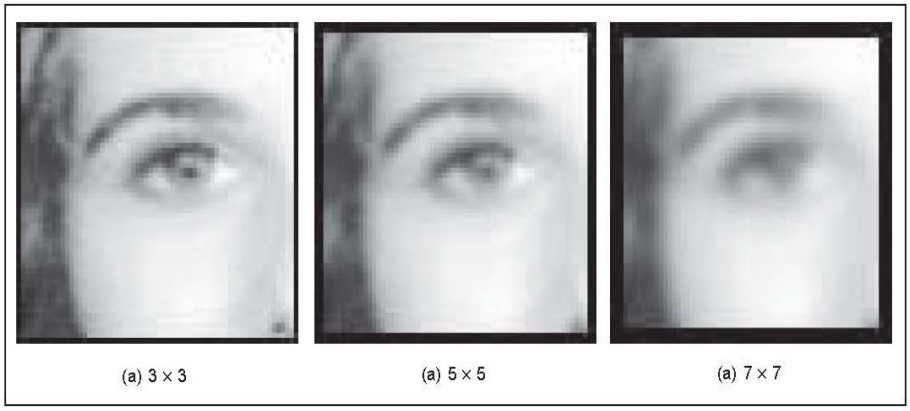
1.4.2 Πρότυπα του Gauss

Τα πρότυπα του Gauss έχουν θεωρηθεί ότι είναι βέλτιστα για την λείανση της εικόνας. Αυτά παίρνουν τιμές από την πιο κάτω σχέση του Gauss . Η συνάρτηση του Gauss g στις συντεταγμένες x, y σε σχέση με τη διασπορά σ^2 δίνεται ως εξής :

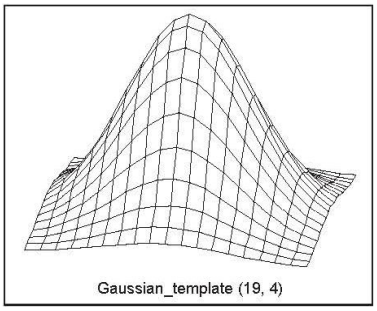
$$g(x, y) = e^{-\frac{(x^2+y^2)}{2\sigma^2}} \qquad (1.18)$$

Η Εξίσωση 1.18 δίνει ένα τρόπο για αν υπολογίσουμε τα βάρη στα πρότυπα του Gauss . Τα αποτελέσματα της χρήσης προτύπων του Gauss διαφορετικού μεγέθους παρουσιάζονται στο Σχήμα 15. Στο Σχήμα 15α το 3×3 πρότυπο του Gauss διατηρεί πιο πολλά χαρακτηριστικά στην εικόνα σε σχέση με αυτό στο Σχημα 14α. Όσο το μέγεθος του προτύπου μεγαλώνει τόσο απομακρύνονται και χαρακτηριστικά γνωρίσματα της εικόνας, όμως σε μικρότερο βαθμό σχέση με τα πρότυπα που χρησιμοποιούσαν τον μέσο όρο του προτύπου. Η γραφική παράσταση της συνάρτησης του Gauss (1.18) έχει τη μορφή της γνωστής καμπάνας (Σχήμα 16). Οι τιμές της συνάρτησης σε διακεκριμένα σημεία είναι οι τιμές των βαρών του προτύπου του Gauss . Η συνέλιξη αυτού του προτύπου με μια εικόνα δίνουν το μέσο όρο του Gauss . Ένα διαδεδομένο μέγεθος για το πρότυπο του Gauss

είναι το 5×5 με διασπορά $\sigma = 1$ (Σχήμα 17). [1,5]



Σχήμα 15: Χρήση τριών προτύπων του Gauss



Σχήμα 16: Η καμπάνα του Gauss

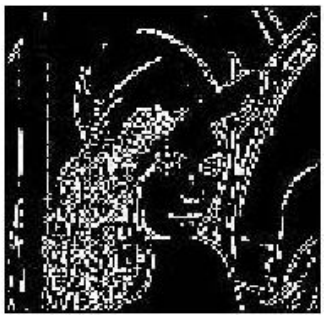
0.02	0.08	0.14	0.08	0.02
0.08	0.37	0.61	0.37	0.08
0.14	0.61	1.0	0.61	0.14
0.08	0.37	0.61	0.37	0.08
0.02	0.08	0.14	0.08	0.02

Σχήμα 17: 5×5 πρότυπο Gauss

2 Εντοπισμός ακμών

2.1 Γενική επισκόπηση

Σκοπός είναι να κατασκευαστούν τεχνικές οι οποίες θα μπορούν να εξάγουν χαρακτηριστικά και διάφορες λεπτομέρειες από μια εικόνα. Η μέθοδος του κατώτατου ορίου είναι μια τέτοια τεχνική η οποία μπορεί να βρει με επιτυχία αντικείμενα σε μια εικόνα, αρκεί να γνωρίζουμε την αντιθεσή τους. Με τον εντοπισμό των ακμών μιας εικόνας μπορούμε να εντοπίσουμε λεπτομέρειες αυτής οι οποίες δεν θα μπορούσαν να εντοπισθούν με την μέθοδο του κάτω ορίου. Η τεχνική του εντοπισμού των ακμών επιστρέφει μια εικόνα με γραμμές οι οποίες εμφανίζονται σε εκείνες τις περιοχές όπου η αντίθεση είναι πιο έντονη (Σχήμα 18). Υπάρχουν πολλές τεχνικές που πετυχαίνουν τον εντοπισμό των ακμών σε μία εικόνα. Σε αυτή την ενότητα θα περιγραφθούν οι δημοφιλέστερες προσεγγίσεις.



Σχήμα 18:

2.2 Βασικοί ανιχνευτές ακμών

2.2.1 Βασικοί μηχανισμοί

Πολλές προσεγγίσεις που έχουν σχέση με την ερμηνεία της εικόνας είναι βασισμένες στις ακμές. Ο εντοπισμός των ακμών εστιάζει πάνω στην αντίθεση της εικόνας. Εξετάζοντας την αντίθεση της εικόνας μπορούμε να δώσουμε έμφαση στα όρια των χαρακτηριστικών της γνωρισμάτων, όπου εκεί η αντίθεση της εικόνας είναι πιο έντονη. Αυτό είναι λογικό αφού με την αντίθεση των χρωμάτων μπορούμε να αντιληφθούμε το σύνορο ενός αντικειμένου, από τη στιγμή που το αντικείμενο έχει διαφορετική ένταση χρωμάτων από το περιβάλλον του. Ουσιαστικά το σύνορο ενός αντικειμένου βρίσκεται ένα βήμα πριν την αλλαγή της έντασης των χρωμάτων. Οι ακμές βρίσκονται σε εκείνες τις θέσεις όπου έχουμε έντονη αλλαγή χρωμάτων. Η αλλαγή στην αντίθεση μπορεί να εντοπιστεί σε περιοχές της εικόνας που έχουν διφορετικά επίπεδα χρωμάτων. Η ανίχνευση των ακμών γίνεται κάθετα και οριζόντια. Η οριζόντια ανίχνευση θα εντοπίσει τις κάθετες ακμές και οι απαιτούμενοι ανιχνευτές καλούνται οριζόντιοι ανιχνευτές ακμών. Αντίστοιχα η κάθετη ανίχνευση θα εντοπίσει τις οριζόντιες ακμές. Όταν σε μια εικόνα P εφαρμόζεται ένας οριζόντιος ανιχνευτής ακμών τότε αυτός σημειώνει τη διαφορά Ex δυο οριζόντιων γειτονικών σημείων :

$$Ex_{x,y} = |P_{x,y} - P_{x+1,y}| \qquad \forall\, x \,\epsilon\, 1, N-1; \forall\, y \,\epsilon\, 1, N \qquad (2.1)$$

Για να εντοπίσουμε τις οριζόντιες ακμές χρειαζόμαστε έναν κάθετο ανιχνευτή ακμών που θα εντοπίσει την αντίθεση στα κάθετα συνεχόμενα σημεία. Για τον κάθετο ανιχνευτή έχουμε την σχέση :

$$Ey_{x,y} = |P_{x,y} - P_{x,y+1}| \qquad \forall\, x \,\epsilon\, 1, N; \forall\, y \,\epsilon\, 1, N-1 \qquad (2.2)$$

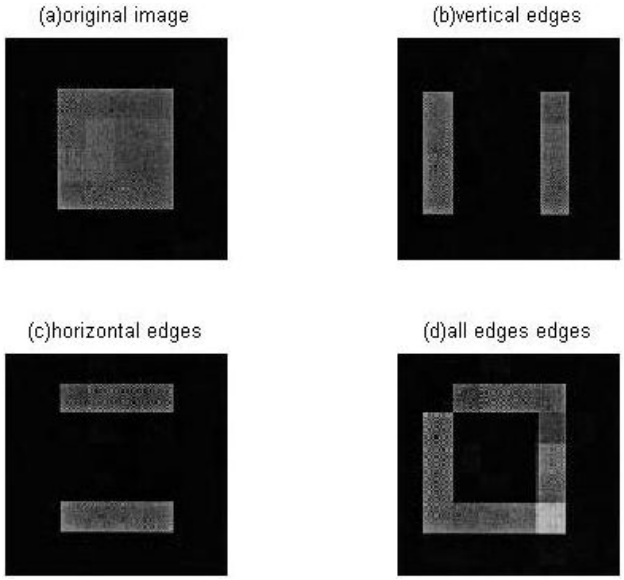
Τα Σχήματα 19b και 19c δείχνουν το αποτέλεσμα του οριζόντιου και του κάθετου μετασχηματισμού που πραγματοποιήθηκαν στην Σχήμα 19α. Συνδυάζοντας τον κάθετο και τον οριζόντιο ανιχνευτή έχουμε τον ανιχνευτή E ο οποίος μπορεί να εντοπίσει κάθετες και οριζόντιες ακμές μαζί. Έτσι έχουμε :

$$E_{x,y} = |P_{x,y} - P_{x+1,y} + P_{x,y} - P_{x,y+1}| \qquad \forall\, x, y \,\epsilon\, 1, N-1 \qquad (2.3)$$

Η πιο πάνω σχέση γίνεται :

$$E_{x,y} = |2 \cdot P_{x,y} - P_{x+1,y} - P_{x,y+1}| \qquad \forall\, x, y \,\epsilon\, 1, N-1 \qquad (2.4)$$

Η Σχέση 2.4 μπορεί να δώσει τους συντελεστές ενός προτύπου (Σχήμα 20) που με τη διαδικασία της συνέλιξης με την εικόνα μπορεί να εντοπίσει όλες τις ακμές της εικόνας, όπως φαίνεται στο Σχήμα 19d . Παρατηρούμε



Σχήμα 19:

ότι το φωτεινό σημείο κάτω δεξιά του τετραγώνου στο Σχήμα 19d είναι πιο φωτεινό απο τα άλλα σημεια. Αυτό συμβαίνει επειδή είναι το μόνο σημείο που εντοπίστηκε σαν ακμή και από τον κάθετο και από τον οριζόντιο ανιχνευτή. Αυτός είναι ο λόγος που είναι πιο φωτεινό από τις υπόλοιπες ακμές. Το πάνω αριστερό σημείο δεν εντοπίστηκε από κανένα ανιχνευτή οπότε δεν εμφανίστηκε καθόλου στην τελική εικόνα. Το πρότυπο στο Σχήμα

2	–1
–1	0

Σχήμα 20:

20 είναι αυτό που χρησιμοποιείται στη διαδικασία της συνέλιξη της εικόνας για τον εντοπισμό των ακμών της. Ύστερα από την συνέλιξη εφαρμόζεται ο μηχανισμός του κάτω ορίου για την επιλογή των πιο φωτεινών σημείων. Η τιμή του κάτω ορίου θα καθορίσει πόσα σημεία θα παραμείνουν στην εικόνα. Αν η τιμή του είναι πολύ υψηλή και πλησιάζει το 255 τοτε θα παραμείνουν πολύ λίγα σημεία. Αντίθετα αν η τιμή του είναι χαμηλή τότε θα

παραμείνουν πιο πολλά σημεία. Η τιμή του κάτω ορίου σε αυτές τις περιπτώσεις είναι καλύτερο να επιλέγεται με βάση τις ανάγκες του προβλήματος και ύστερα από αρκετές δοκιμές. Όμως και η μέθοδος της αυτόματης επιλογής, με μερικές παρεμβάσεις, μπορεί να φανεί ιδιαίτερα αποδοτική. Στο Σχήμα 21 φαίνεται το αποτέλεσμα της διαδικασίας με την οποία βρίσκονται οι ακμές μιας εικόνας Σχημα 21α. Στο Σχήμα 21b παρουσιάζεται το αποτέλεσμα της εφαρμογής του οριζόντιου και του κάθετου εντοπισμού των ακμών. Το Σχήμα 21c είναι η τελική εικόνα η οποία προκύπτει όταν το κάτω όριο είναι ίσο με 15. [1,4]

Σχήμα 21:

2.2.2 Ανάλυση των βασικών μηχανισμών

Με την ανάλυση των βασικών μηχανισμών εντοπισμού των ακμών θα προσπαθήσουμε να βρούμε πρότυπα τα οποία θα αποδίδουν καλύτερα. Για τον σκοπό αυτό θα χρησιμοποιήσουμε σειρές Taylor και θα εξετάσουμε τα εικονοκύτταρα σε πολύ μικρές περιοχές της εικόνας. Αν υποθέσουμε ότι ένα εικονοκύτταρο βρίσκεται σε μια θεση x και η απόσταση δύο εικονοκυττάρων είναι Δx. Τότε, συμβολίζοντας με $f(x)$ την ένταση του εικονοκυττάρου και τη μεταβολή της εντασής του με $f'(x)$, για διαδοχικά σημεία έχουμε.

$$f(x+\Delta x) = f(x) + \Delta x \cdot f'(x) + \frac{\Delta x^2}{2!} \cdot f''(x) + \frac{\Delta x^3}{3!} \cdot f'''(x) \qquad (2.5)$$

Λύνουμε ως προς την πρώτη παράγωγο $f'(x)$ και έχουμε :

$$f'(x) = \frac{f(x+\Delta x) - f(x)}{\Delta x} - (\frac{\Delta x}{2!} \cdot f''(x) + \frac{\Delta x^2}{3!} \cdot f'''(x)) \qquad (2.6)$$

όπου το σφάλμα είναι

$$O(\Delta x^2) = \frac{\Delta x}{2!} \cdot f''(x) + \frac{\Delta x^2}{3!} \cdot f'''(x) \qquad (2.7)$$

Το σφάλμα εξαρτάται από την τιμή του Δx καθώς και από τις τιμές των ανώτερων παραγώγων. Όσο μεγαλύτερο είναι το Δx ή οι τιμές των παραγώγων τόσο μεγαλύτερο θα είναι και το σφάλμα. Στην πράξη όταν τα εικοκύταρρα είναι γειτονικά τότε το σφάλμα δίνεται από την Σχέση (2.7). Στη συνέχεια υπολογίζουμε τον οριζόντιο ανιχνευτή της Σχέσης (2.1) όταν τα εικονοκυττάρα που συμετέχουν στον υπολογισμό δεν είναι διαδοχικά αλλά έχουν απόσταση ενός εικονοκυττάρου. Τότε ο νέος οριζόντιος ανιχνευτής θα είναι Exx και υπολογίζεται από τη σχέση :

$$Exx_{x,y} = |P_{x+1,y} - P_{x-1,y}| \qquad \forall\, x \,\epsilon\, 2, N-1; \forall\, y \,\epsilon\, 1, N \qquad (2.8)$$

Για να δούμε την συμπεριφορά του σφάλματος όταν τα σημεία δεν είναι διαδοχικά υπολογίζουμε το $f(x-\Delta x)$:

$$f(x-\Delta x) = f(x) - \Delta x \cdot f'(x) + \frac{\Delta x^2}{2!} \cdot f''(x) - \frac{\Delta x^3}{3!} \cdot f'''(x) \qquad (2.9)$$

Εξισώνοντας τις Σχέσεις (2.5) και (2.9) έχουμε :

$$f'(x) = \frac{f(x+\Delta x) - f(x-\Delta x)}{2 \cdot \Delta x} - \frac{\Delta x^2}{3!} \cdot f'''(x) \qquad (2.10)$$

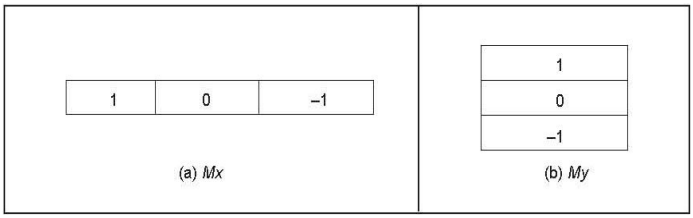
Το σφάλμα της σχέσης 2.10 είναι :

$$O(\Delta x^2) = \frac{\Delta x^2}{3!} \cdot f'''(x) \qquad (2.11)$$

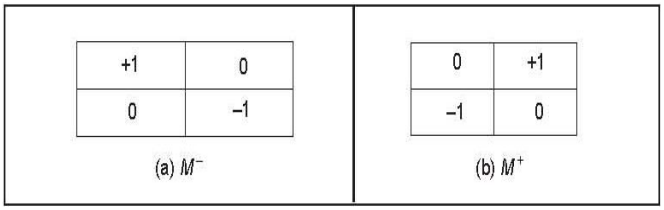
Παρατηρούμε ότι το σφάλμα της (2.10) είναι μικρότερο από αυτό της (2.6). Δηλαδή όταν ο υπολογισμός γίνεται με εικονοκύτταρα που έχουν απόσταση ένα εικονοκύτταρο έχουμε μικρότερο σφάλμα στην προσέγγιση. Το πρότυπο του οριζόντιου και του κάθετου ανιχνευτή φαίνονται στο Σχήμα 22 και συμβολίζονται με Mx και My αντίστοιχα.

Η μέθοδος του Robert (Robert, 1965) ήταν μια από τις πρώτες μεθόδους για τον εντοπισμό ακμών. Είναι μια έκδοση που χρησιμοποιεί δύο πρότυπα, ένα για οριζόντιο και ένα για κάθετο εντοπισμό, και ελέγχει εικονοκύτταρα που βρίσκονται σε διαγώνια διάταξη. Τα δύο πρότυπα ονομάζονται M^- και M^+ και φαίνονται στο Σχήμα 23. Στην πράξη, η μεγαλύερη τιμή που προκύπτει από την χρήση των δύο παραπάνω προτύπων αποθηκεύεται ως η τιμή της ακμής σε εκείνο το σημείο. Οπότε οι ακμές $E_{x,y}$ στη μέθοδο του Robert δίνονται από τη σχέση :

$$E_{x,y} = max(|M^+ \cdot P_{x,y}|, |M^- \cdot P_{x,y}|) \qquad \forall\, x, y \,\epsilon\, 1, N-1 \qquad (2.12)$$

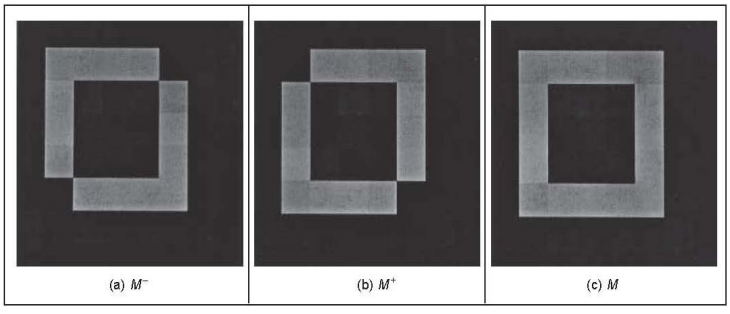


Σχήμα 22: Οριζόντιο και κάθετο πρότυπο



Σχήμα 23: Τα πρότυπα του Robert

Η εφαρμογή της μεθόδου του Robert σε μια εικόνα με ένα τετράγωνο απεικονίζεται στο Σχήμα 24. Το αποτέλεσμα της εφαρμογής των δύο προτύπων προβάλονται στα Σχήματα 24(α) και 24(b) και το τελικό αποτέλεσμα της μεθόδου φαίνεται στο Σχήμα 24c . Παρατηρούμε ότι με τη μέθοδο του Robert οι γωνίες του τετραγώνου υπάρχουν όλες στο τελικό σχήμα σε αντίθεση με την προηγούμενη μέθοδο στην οποία έλειπε μια γωνία από το τελικό σχήμα. Ένας εναλλακτικός τρόπος από το να πάρουμε το μέγιστο είναι κρατήσουμε τα αποτελέσματα

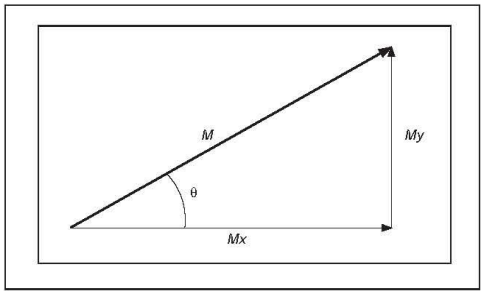


Σχήμα 24: Εφαρμογή της μεθόδου του Robert

των δύο προτύπων και να συνδυάσουμε τις κάθετες και τις οριζόντιες ακμές. Υπάρχουν και άλλοι τρόποι να χρησιμοποιήσουμε τα αποτελέσματα των δύο προτύπων και να έχουμε διαφορετικές προσεγγίσεις των ακμών. Οι πιο κοινές προσεγγίσεις είναι οταν τα αποτελεσματα των δύο προτύπων χρησιμοποιούνται σαν συνιστώσες ενός διανύσματος το οποίο δείχνει την κατεύθυνση της ακμής (Σχήμα 25). [1]

2.2.3 Η μέθοδος του Prewitt

Ο εντοπισμός των ακμών είναι μια διαδικασία που επηρεάζεται από τον θόρυβο που υπάρχει στην εικόνα. Ο Prewitt για να αποφύγει μέρος του θορύβου θεώρησε οτι έπρεπε να εξετάσει μεγαλύτερες περιοχές για τον εντοπισμό μιας ακμής. Για το σκοπό αυτό χρησιμοποίησε μεγαλύτερα πρότυπα εντοπισμού ακμών. Επέκτεινε το κάθετο πρότυπο Mx κατά μήκος τριών γραμμών και το οριζόντιο πρότυπο My κατά μήκος τριών στηλών. Αυτά τα πρότυπα (Σχήμα 26) αποτελούν το μηχανισμό εντοπισμού ακμών του Prewitt .



Σχήμα 25: Διάνυσμα M που προκύπτει από τα πρότυπα Mx και My

1	0	–1
1	0	–1
1	0	–1

(a) *Mx*

1	1	1
0	0	0
–1	–1	–1

(b) *My*

Σχήμα 26: Τα πρότυπα του Prewitt

Με συνδυασμό των δύο προτύπων μπορούμε να εντοπίσουμε τις ακμές που προκύπτουν από τη μέθοδο του Prewitt . Αυτές είναι το διάνυσμα M στο Σχήμα 25. Το μέγεθος της ακμής είναι το μήκος του διανύσματος M και η κατευθυνσή της είναι η γωνία θ του διανύσματος.

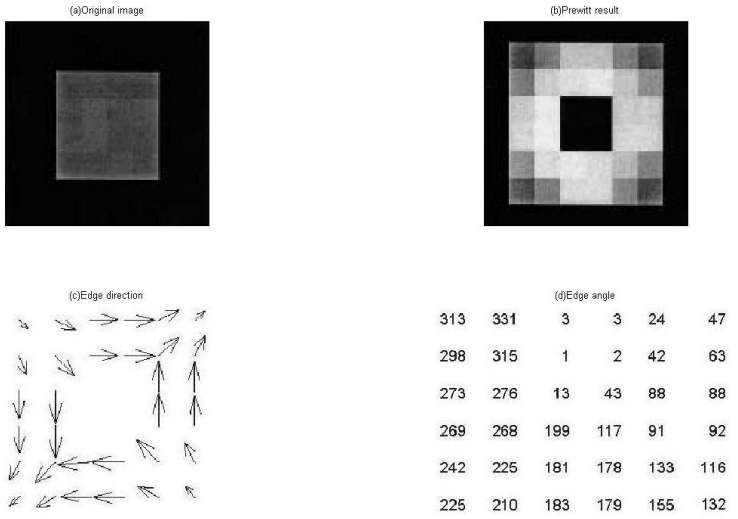
$$M = \sqrt{Mx(x,y)^2 + My(x,y)^2} \qquad (2.13)$$

$$\theta(x,y) = tan^{-1}(\frac{My(x,y)}{Mx(x,y)}) \qquad (2.14)$$

Στο Σχήμα 27 παρουσιάζεται ένα παράδειγμα όπου εφαρμόζεται η μέθοδος του Prewitt . Το Σχήμα 27α είναι το αρχικό τετράγωνο διάστασης 6×6. Στο Σχήμα 27b φαίνεται το αποτέλεσμα της εφαρμογής της μεθόδου. Στο Σχήμα 27d φαίνονται οι γωνίες των ακμών που εντοπίστικαν ενώ στο Σχήμα 27c είναι σχεδιασμένες οι κατευθύνσεις τους.

2.2.4 Η μέθοδος του Sobel

Όταν στη μέθοδο του Prewitt τα κεντρικά βάρη των προτύπων στις μη μηδενικές γραμμές και στήλες είναι διπλασιασμένα (Σχήμα 28) τότε έχουμε τη μέθοδο του Sobel για τον εντοπισμό ακμών, στην οποία υπάρχουν πάλι δύο πρότυπα τα οποία βρίσκουν ακμές σε μορφή διανύσματος. Η μέθοδος του Sobel ήταν η πιο δημοφιλής μέθοδος εντοπισμού ακμών μέχρι την ανάπτυξη μεθόδων με θεωρητική βάση. Αποδείχθηκε δημοφιλής διότι είχε



Σχήμα 27: Εφαρμογή του Prewitt

1	0	−1
2	0	−2
1	0	−1

(a) *Mx*

1	2	1
0	0	0
−1	−2	−1

(b) *My*

Σχήμα 28: 3×3 πρότυπα του Sobel

καλύτερη απόδοση από άλλες σύγχρονες μεθόδους όπως αυτή του Prewitt . Τα πρότυπα στο Σχήμα 28 είναι η πιο συνηθισμένη έκδοση των προτύπων του Sobel . Τα πρότυπα του Sobel μπορούν να είναι μεγαλύτερου μεγέθους, όπως για παράδειγμα 5×5 , 7×7 ή και μεγαλύτερου. Για να υπολογίσουμε τους συντελεστές μεγαλύτερων προτύπων πρέπει να στηριχθούμε σε κάποια θεωρητική βάση. Μια προσέγγιση είναι να βασιστούμε πάνω σε θεωρία που έχει σχέση με τον υπολογισμό του μέσου όρου με βέλτιστους τρόπους. Η μέθοδος του Gauss που έχει ήδη αναφερθεί δίνει τέτοια αποτελέσματα. Τα τρίγωνα του Pascal δίνουν ένα σύνολο συντελεστών που προσεγγίζει αρκετά το αποτέλεσμα των προτύπων του Gauss . Τα τρίγωνα του Pascal είναι :

Triangle A

Window size					
2		1	1		
3		1	2	1	
4	1	3	3	1	
5	1	4	6	4	1

Triangle B

Window size					
2		1	−1		
3	1	0	−1		
4	1	1	−1	−1	
5	1	2	0	−2	−1

Η δεύτερη γραμμή του τριγώνου A μπορεί να χρησιμοποιηθεί για να δώσει τα βάρη του προτύπου στο Σχήμα 28. Το βέλτιστο πρότυπο προκύπτει με συνδυασμό των δύο τριγώνων. Για να πάρουμε το βέλτιστο πρότυπο Mx διάστασης 5×5 χρησιμοποιούμε την τέταρτη γραμμή των τριγώνων. Έστω a η τέταρτη γραμμή του πρώτου

τριγώνου και b η τέταρτη γραμμή του τριγώνου διαφορών. Τότε έχουμε :

$$Mx = a^T \cdot b \qquad (2.15)$$

όπου a^T το ανάστροφο διάνυσμα του a. Για το πρότυπο My ακολουθούμε την ίδια διαδικασία και έχουμε :

$$My = b^T \cdot a \qquad (2.16)$$

Στη συνέχεια παρουσιάζονται τα πρότυπα Mx και My του Sobel διάστασης 5×5 τα οποία συνδέονται με τη σχέση :

$$M^T x = My \qquad (2.17)$$

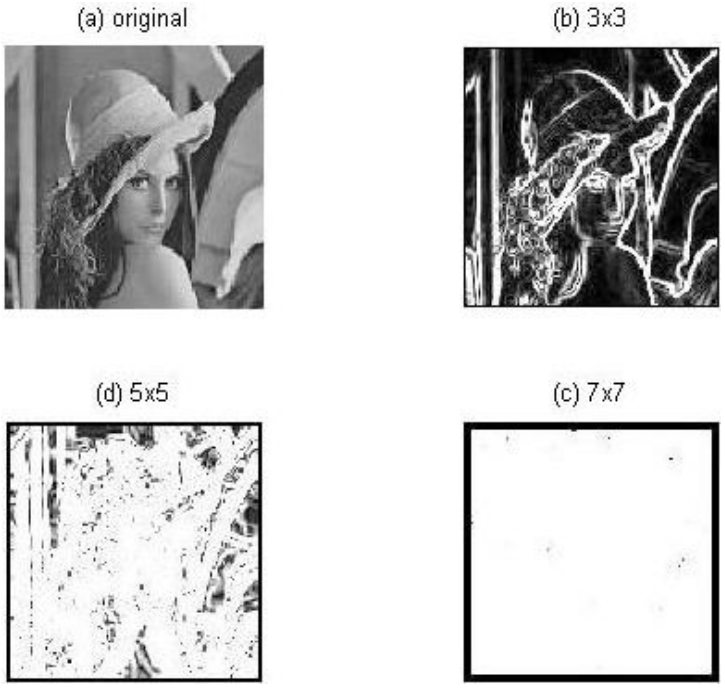
$Mx =$

1	2	0	-2	-1
4	8	0	-8	-4
6	12	0	-12	-6
4	8	0	-8	-4
1	2	0	-2	-1

$My =$

1	4	6	4	1
2	8	12	8	2
0	0	0	0	0
-2	-8	-12	-8	-2
-1	-4	-6	-4	-1

Απο τα πιο πάνω είναι ξεκάθαρο ότι μπορούμε να κατασκευάσουμε πρότυπα με οποιοδήποτε μέγεθος. Η χρήση μεγαλύτερου προτύπου βοηθάει στο να μειωθεί ο θόρυβος στο τελικό αποτέλεσμα. Όμως ταυτόχρονα με την αποφυγή του θορύβου χάνουμε και πληροφορίες για τις ακμές. Αυτό έχει σαν αποτέλεσμα ότι όσο μεγαλύτερο πρότυπο χρησιμοποιούμε τόσο λιγότερες ακμές εντοπίζονται (Σχήμα 29).



Σχήμα 29:

Στο Σχήμα 29b έχει χρησιμοποιηθεί το 3×3 πρότυπο του Sobel . Το πρότυπο αυτό έχει πολύ καλή απόδοση αφού στην τελική εικόνα έχουν σχηματιστεί τα περιγράμματα σε εκείνες τις περιοχές οπου υπάρχει αλλαγή των χρωμάτων. Επίσης έχουν διατηρηθεί και αρκετές λεπτομέρειες όπως τα μάτια και μαλλιά. Στο Σχήμα 29c έχει χρησιμοποιηθεί το 5×5 πρότυπο του Sobel . Με τη χρήση αυτού του προτύπου έχουν εμφανιστεί ακμές μόνο στις περιοχές όπου η αλλαγή των χρωμάτων είναι πολύ έντονη. Στο Σχήμα 29d έχει χρησιμοποιηθεί το 7×7 πρότυπο του Sobel . Στην τελική εικόνα έχουν εμφανιστεί λίγες μαύρες τελείες οι οποίες δεν μπορούν να δώσουν καμία πληροφορία για την αρχική εικόνα. Πρέπει να σημειωθεί και εδώ ότι η χρήση μεγάλου προτύπου

έχει σαν αποτέλεσμα την μείωση του συνόρου της τελικής εικόνας όπως συμβαίνει και στα πρότυπα μείωσης του θορύβου. Το μέγεθος του προτύπου καθορίζεται κάθε φορά από το είδος του προβλήματος και από το πλήθος των πληροφοριών που θέλουμε να πάρουμε. Αφού ολοκληρωθεί η συνέλιξη της εικόνας με το πρότυπο εφαρμόζουμε την τεχνική του κατώτατου ορίου για να πάρουμε μια εικόνα που θα αποτελείται από λευκά και μαύρα εικονοκύτταρα. Το σημείο αυτό είναι ιδιαίτερα σημαντικό αφού η επιλογή του κάτω ορίου θα καθορίσει ποιές πληροφορίες θα μείνουν στην τελική εικόνα (Σχήμα 30). Στη μέθοδο του Sobel μπορούμε να ορίσουμε από την αρχή ποιά θα είναι η κατεύθυνση των ακμών. Αυτό γίνεται με την αντιστροφή των προτύπων. Αν αντιστρέψουμε και τα δύο πρότυπα τότε η κατεύθυνση των ακμών θα αντιστραφεί γύρο και από τους δύο άξονες. Αν αντιστρέψουμε μόνο το ένα πρότυπο τότε η κατεύθυνση των ακμών θα αντιστραφεί γύρο από τον αντίστοιχο άξονα. Πιο κάτω φαίνεται η αντιστροφή και των δύο προτύπων του Sobel διάστασης 3×3.

$$-Mx = \begin{array}{|c|c|c|}\hline -1 & 0 & 1 \\ \hline -2 & 0 & 2 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \qquad -My = \begin{array}{|c|c|c|}\hline -1 & -2 & -1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 2 & 1 \\ \hline \end{array}$$

Σχήμα 30:

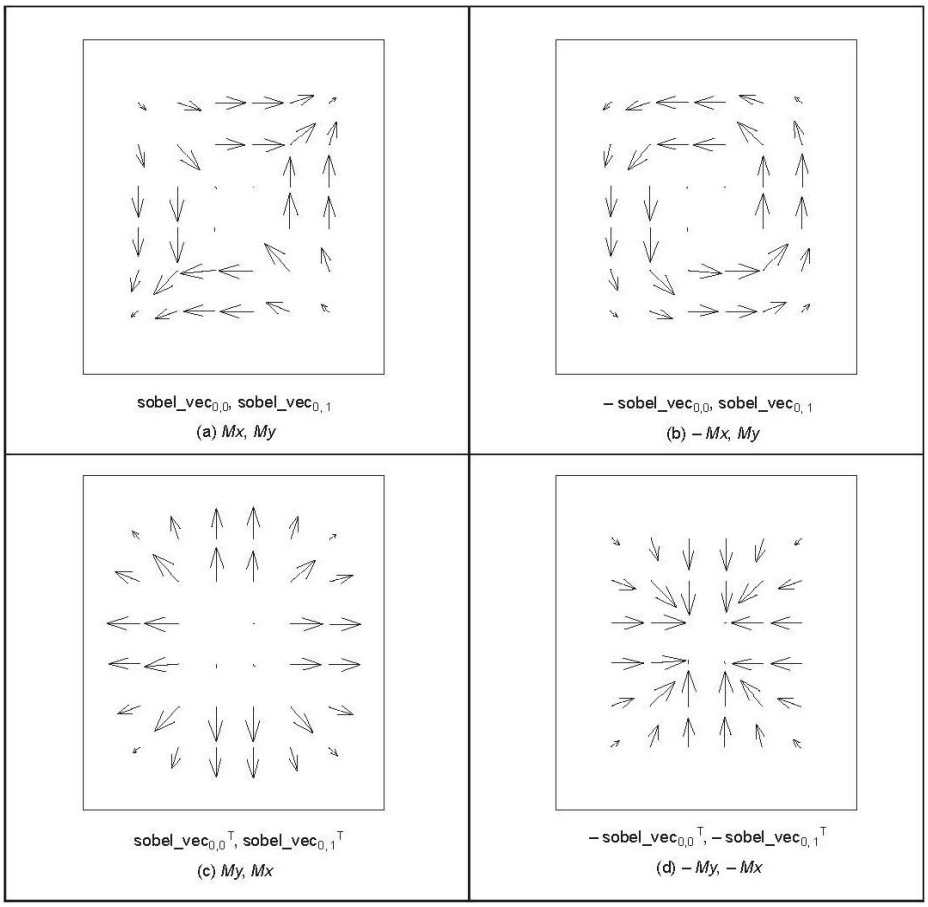
Με την αντιστροφή των προτύπων έχουμε τέσσερις πιθανές κατευθύνσεις των ακμών (Σχήμα 31). Η επιλογή των πιο πάνω προτύπων γίνεται με βάση τη μορφή του σχήματος που εξετάζουμε και με βάση χαρακτηριστικών όπως η επιφάνειά του ή άλλες λεπτομέρειες με κάποιο ιδιαίτερο σχήμα. Αν για παράδειγμα το σχήμα χαρακτηρίζεται από επιφάνεια χωρίς καμπύλες τότε είναι καλύτερη η επιλογή των προτύπων c ή d αλλιώς επιλέγουμε κάποιο από τα a ή b. [1,4]

2.2.4 Η μέθοδος του Canny

Η μέθοδος του Canny [7] είναι ίσως η πιο δημοφιλής μέθοδος εντοπισμού ακμών. Διατυπώθηκε με τρεις κύριους στόχους :

1. Βέλτιστος εντοπισμός με χαμηλό ποσοστό λάθους.
2. Καλός εντοπισμός με ελάχιστη απόσταση μεταξύ της ανιχνευμένης και της αληθινής θέσης των ακμών.
3. Ο εντοπισμός κάθε ακμής θα είναι μοναδικός χωρίς επαναλήψεις.

Σύμφωνα με τον πρώτο στόχο οι ακμές που υπάρχουν σε μια εικόνα θα πρέπει να εντοπίζονται και αυτές που δεν υπάρχουν να αγνοούνται. Ακμές που στην πραγματικότητα δεν υπάρχουν είναι πιθανόν να εντοπιστούν λόγο του θορύβου. Αυτό μπορεί να πραγματοποιηθεί με βέλτιστη λείανση της εικόνας. Ο Canny ήταν ο πρώτος που χρησιμοποίησε τα φίλτρα του Gauss για τον εντοπισμό ακμών αφού τα θεώρησε βέλτιστα για το σκοπό αυτό. Το δεύτερο κριτήριο στοχεύει στην ακρίβεια: οι ακμές πρέπει να εντοπίζονται στην θέση στην οποία βρίσκονται. Ο τρίτος περιορισμός αφορά την τοποθεσία μιας μοναδικής ακμής. Υπάρχουν περιπτώσεις που λόγο του θορύβου και της έντονης αλλαγής των χρωμάτων σε κάποιες περιοχές να εντοπιστούν ακμές που στην πραγματικότητα δεν υπάρχουν. Βασισμένει σε αυτά τα τρία κτιτήρια η μέθοδος του Canny μπορεί να πραγματοποιηθεί με την



Σχήμα 31:

εφαρμογή των παρακάτω έξι βημάτων :

Βήμα 1

Το πρώτο βήμα της μεθόδου είναι το φιλτράρισμα της αρχικής εικόνας για την απομάκρυνση του θορύβου που υπάρχει σε αυτή πρίν ξεκινήσει η διαδικασία εντοπισμού των ακμών. Όπως αναφέραμε και πρίν για αυτό το σκοπό ο Canny χρησιμοποίησε τα φίλτρα του Gauss . Αφού υπολογίσουμε ένα πρότυπο του Gauss στην συνέχεια με συνέλιξη με την εικόνα επιτυγχάνουμε την λείανση αυτής. Το πρότυπο που θα χρησιμοποιήσουμε θα είναι πολύ μικρότερο από την εικόνα. Όσο μεγαλύτερο πρότυπο χρησιμοποιήσουμε τόσες περισσότερες πληροφορίες χάνουμε για την εικόνα με αποτέλεσμα να εντοπίζονται λιγότερες ακμές. Ένα πρότυπο του Gauss που χρησιμοποιείται συχνά είναι όταν $\sigma = 1.4$.

$$\frac{1}{115}\begin{array}{|c|c|c|c|c|}\hline 2 & 4 & 5 & 4 & 2 \\ \hline 4 & 9 & 12 & 9 & 4 \\ \hline 5 & 12 & 15 & 12 & 5 \\ \hline 4 & 9 & 12 & 9 & 4 \\ \hline 2 & 4 & 5 & 4 & 2 \\ \hline \end{array}$$

Βήμα 2

Αφού έχουμε λειάνει την εικόνα και έχουμε απομακρύνει τον θόρυβο βρίσκουμε την κλίση της χρησιμοποιώντας τα 3×3 πρότυπα του Sobel . Το ένα θα υπολογίσει την κλίση της κατά μήκος του άξονα των x και το άλλο

κατά μήκος του άξονα των y.

$$Mx = \begin{array}{|c|c|c|}\hline 1 & 0 & -1 \\ \hline 2 & 0 & -2 \\ \hline 1 & 0 & -1 \\ \hline \end{array} \qquad My = \begin{array}{|c|c|c|}\hline 1 & 2 & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -2 & -1 \\ \hline \end{array}$$

Η κλίση της εικόνας δίνεται από την σχέση :

$$|G| = |Gx| + |Gy| \qquad (2.18)$$

Βήμα 3

Αφού έχουμε υπολογίσει την κλίση των ακών και ως προς τους δύο άξονες στη συνέχεια υπολογίζουμε την κατευθυνσή τους :

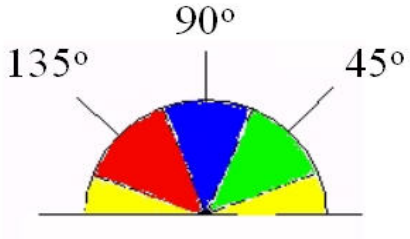
$$\theta = tan^{-1}(\frac{Gy}{Gx}) \qquad (2.19)$$

Βήμα 4

Σε αυτό το βήμα ομαδοποιούμε τις κατευθύνσεις που έχουν υπολογιστεί στο βήμα 3. Υποθέτουμε ότι έχουμε το πιο κάτω 5×5 τμήμα της εικόνας :

$$\begin{array}{ccccc} x & x & x & x & x \\ x & x & x & x & x \\ x & x & a & x & x \\ x & x & x & x & x \\ x & x & x & x & x \end{array}$$

Με την εξέταση του εικονοκύτταρου a υπάρχουν μόνο τέσσερις πιθανές κατευθύνσεις που μπορεί να κινηθεί κοιτάζοντας τα γειτονικά του εικονοκύτταρα: **0 μοίρες** στην οριζόντια κατεύθυνση, **45 μοίρες** κατα μήκος της θετικής διαγωνίου, **90 μοίρες** στην κάθετη κατεύθυνση και **135 μοίρες** κατα μήκος της αρνητικής διαγωνίου. Τώρα έχουμε περιορίσει τις κατευθύνσεις των ακμών σε τέσσερις πιθανές. Ο διαχωρισμός αυτός γίνεται σύμφωνα με το πιο κάτω σχήμα :

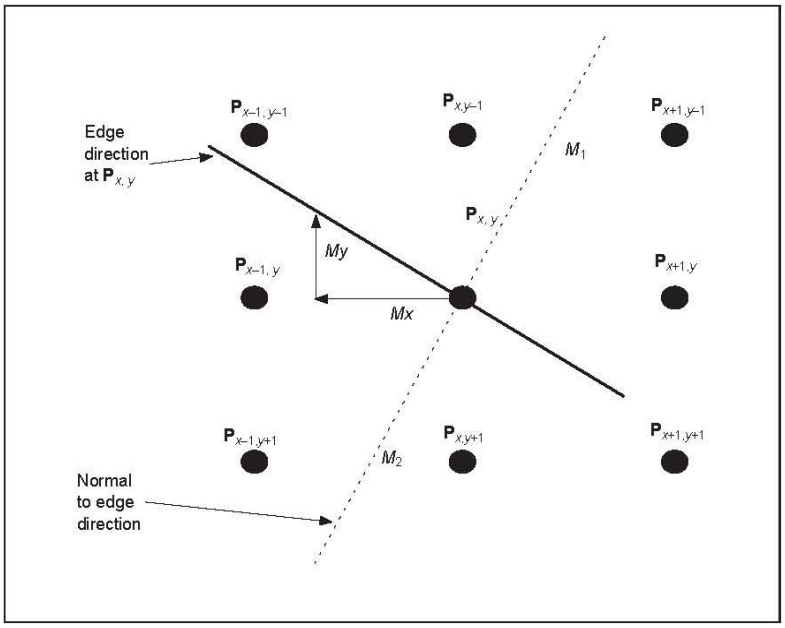


Αν η κατεύθυνση μιας ακμής βρίσκεται στην κίτρινη περιοχή ($0 - 22.5$ και $157.5 - 180$ μοίρες) τότε δίνουμε σε αυτήν νέα τιμή 0 μοίρες. Αν η κατεύθυνση της βρίσκεται στην πράσινη περιοχή ($22.5 - 67.5$ μοίρες) τότε δίνουμε σε αυτήν νέα τιμή 45 μοίρες. Αν η κατεύθυνση της βρίσκεται στην μπλέ περιοχή ($67.5 - 112.5$ μοίρες) τότε δίνουμε σε αυτήν νέα τιμή 90 μοίρες. Τέλος αν η κατεύθυνση της βρίσκεται στην κόκκινη περιοχή ($112.5 - 157.5$ μοίρες) τότε δίνουμε σε αυτήν νέα τιμή 135 μοίρες.

Βήμα 5

Σε συτό το βήμα εντοπίζουμε τα σημεία που τοπικά έχουν την μέγιστη κλίση από τα γειτονικά τους. Για μια γειτονιά διάστασης 3×3 ένα σημείο βρίσκεται σε ένα μέγιστο εάν η κλίση σε καθεμία πλευρά είναι μικρότερη από την κλίση αυτού του σημείου. Αυτό υπονοεί ότι χρειαζόμαστε τις τιμές των κλίσεων κατα μήκος μιας γραμμής που θα είναι κάθετη στην ακμή του σημείου (Σχήμα 32). Στο σχήμα αυτό φαίνονται οι γείτονες ενός σημείου $P_{x,y}$, η κατεύθυνση του $P_{x,y}$ και η κάθετη ευθεία στην κατεύθυνση της ακμής του $P_{x,y}$. Το σημείο $P_{x,y}$ θα σημειωθεί ότι είναι μέγιστο αν η κλίση του $M(x, y)$ είναι μεγαλύτερη από τις κλίσεις $M1$ και $M2$. Οι κλίσεις $M1$ και $M2$ υπολογίζονται από τις σχέσεις :

$$M1 = \frac{My}{Mx} M(x + 1, y - 1) + \frac{Mx - My}{Mx} M(x, y - 1) \qquad (2.20)$$



Σχήμα 32:

και

$$M2 = \frac{My}{Mx} M(x-1, y+1) + \frac{Mx - My}{Mx} M(x, y+1) \qquad (2.21)$$

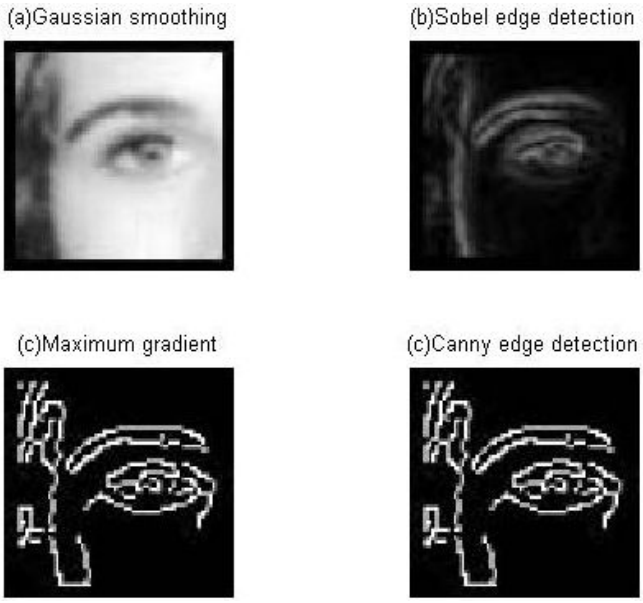
Κρατώντας μόνο τα σημεία με τη μεγαλύτερη κλίση προκύπτει μια εικόνα με λεπτές γραμμές πάνω στα περιγράμματα των σχημάτων της. Αυτό εξυπηρετεί το τελευταίο βήμα της μεθόδου δηλαδή το βήμα 6.

Βήμα 6

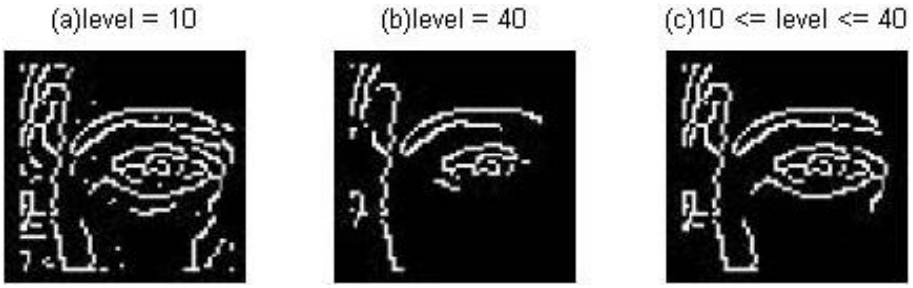
Στο τελευταίο βήμα της μεθόδου το μόνο που έχει μείνει η μετατροπή της εικόνας σε μαύρη-άσπρη. Αυτό το πετυχαίνουμε με τη μέθοδο του κατώτατου ορίου. Η επιλογή ενος κατάλληλου κάτω ορίου είναι ιδιαίτερα σημαντική αφού είναι αυτό που θα διαμορφώσει την τελική μορφή της εικόνας. Πολλές φορές οταν οι συνθήκες το απαιτούν δεν επιλέγεται ένα οριακό σημείο αλλά ένα διάστημα. Αν η τιμή ενός εικονοκυττάρου ανήκει μέσα σε αυτό το διάστημα τότε αυτό γίνεται λευκό αλλιώς το διώχνουμε με την αλλαγή του σε μαύρο. Στο Σχήμα 33 παρουσιάζεται η διαδικασία της μεθόδου του Canny .

Το Σχήμα 34 δείχνει την εικόνα που προκύπτει από τη μέθοδο του Canny όταν έχουμε επιλέξει κατώτατο όριο το 10, το 40 και το διάστημα $[10, 40]$.

Απο το πιο πάνω σχήμα είναι φανερό πόσο σημαντική είναι η επιλογή ενός κάτω ορίου ή ενός διαστήματος. Όταν επίλέξαμε χαμηλό κάτω όριο (level=10 , Σχήμα 34α) εμφανίστηκαν πολλές λεπτομέρειες του προσώπου όμως εμφανίστηκαν και κάποιες λευκές τελείες που πιθανόν δεν μας εξυπηρετούν πουθενά. Όταν αυξήσαμε το κάτω όριο (level=40 Σχήμα 34b) τότε διατηρήθηκαν μόνο τα πολύ έντονα χαρακτηριστικά του προσώπου. Τέλος η επιλογή του διαστήματος $[10, 40]$ (Σχήμα 34c) έδωσε ένα πολύ καλό αποτέτελεσμα. Στο Σχήμα 35 φαίνονται ξεκάθαρα οι διαφορές των δύο μεθόδων Sobel και Canny . Στη μέθοδο του Sobel (Σχήμα 35c) οι ακμές έχουν εντοπιστεί με αρκετή ακρίβεια στην πραγματική τους θέση και μπορούμενα να διακρίνουμε έυκολα τα περιγράμματα των σχημάτων που υπάρχουν στην εικόνα. Με την μέθοδο του Canny (Σχήμα 35b) διατηρούνται όλες οι λεπτομέρειες που εντόπισε η μέθοδος του Sobel με την διαφορά οι γραμμές στο τελικό σχήμα είναι πιο λεπτές δίνοντας έτσι καλύτερο οπτικό αποτέλεσμα αλλά και πληροφορίες μεγαλύτερης ακρίβειας οι οποίες μπορούν να χρησιμοποιηθούν για μελλοντική επεξεργασία [1,4,7].



Σχήμα 33:



Σχήμα 34:

3 Επεξεργασία περιοχών συχνότητας

3.1 Γενική επισκόπηση

Σε αυτή την ενότητα θα εξετάσουμε μια διαφορετική αντιπροσώπευση της εικόνας γνωστή και ως περιοχές συχνότητας της εικόνας. Σύμφωνα με αυτό το όνομα θα εξετάσουμε μια εικόνα σαν μια συλλογή από τμήματα σύχνότητας. Μπορούμε να αναπτύξουμε διάφορες δραστηριότητες στις περιοχές συχνότητας των εικόνων και θα εξετάσουμε διαφορετικούς μηχανισμούς επεξεργασίας. Αυτό μας επιτρέπει να αναπτύξουμε νέες ιδέες και τεχνικές επεξεργασίας οι οποίες θα μας δίνουν νέα νοήματα για την εικόνα. Όταν μια εικόνα μετασχηματιστεί σε εικόνα συχνοτήτων τότε προκύπτει μια εικόνα που δύσκολα κανείς μπορεί να καταλάβει τι απεικονίζει. Όμως θα δούμε ότι αυτή η μορφή είναι πολύ χρήσιμη.

3.2 Μετασχηματισμοί Fourier

Ο ματασχηματισμός Fourier είναι ένας τρόπος να χαρτογραφήσουμε ένα σήμα στις συστατικές του συχνότητες. Η μονάδα μέτρησης των συχνοτήτων είναι το hertz (Hz) και είναι ο ρυθμός επανάληψης με το χρόνο σε δευτερόλεπτα (s). Γενικά έχουμε τις σχέσεις (hertz=1/second) και (second=1/Hz). Ο ήχος έρχεται από ένα

Σχήμα 35:

CD player και καταλήγει στα ηχεία αφού πρώτα έχει επεξεργαστεί από τον ενισχυτή. Στον ενισχυτή μπορούμε να κάνουμε διάφορες αλλαγές στο αρχικό σήμα όπως να αλλάξουμε τα μπάσα ή τα πρίμα. Τα μπάσα καλύπτουν τις χαμηλές συχνότητες και τα πρίμα τις ψηλές. Ο μετασχηματισμός Fourier είναι ένας τρόπος να χαρτογραφήσουμε το σήμα που έρχεται από το CD player , το οποίο είναι ένα σήμα που ποικίλει συνεχώς με το χρόνο, στις συστατικές του συχνότητες. Όταν έχουμε μετασχηματίσει το σήμα τότε ξέρουμε ποιές σχνότητες αποτελούν τον αρχικό ήχο. Με αυτή τη διαδικασία δεν αλλάζουμε το σήμα, μόνο την αντιπροσώπευσή του. Μπορούμε τώρα να το δούμε σε γραφική μορφή συχνοτήτων και όχι σαν τάση που αλλάζει με το χρόνο. Αφού μπορούμε να το δούμε γραφικά τότε μπορούμε να αλλάξουμε και τον ήχο που παράγεται. Εάν για παράδειγμα υπάρχει ένα σφύριγμα στο αρχικό σήμα, τότε αυτό το σφύριγμα θα καλύπτει κάποιες υψηλές συχνότητες τις οποίες μπορούμε να τις εντοπίσουμε και να τις απομακρύνουμε. Πιο συγκεκριμένα ο μετασχηματισμός Fourier Fp ενός σήματος p δίνεται από την πιο κάτω σχέση :

$$Fp(\omega) = \int_{-\infty}^{\infty} p(t)e^{-i\omega t}dt \qquad (3.1)$$

Το $Fp(\omega)$ είναι ο μετασχηματισμός Fourier .

Το ω είναι η γωνιακή συχνότητα, $\omega = 2\pi f$ μετριέται σε rad/sec όπου η συχνότητα $f = 1/time$.

Το i είναι το μιγαδικό i.

Το $p(t)$ είναι το συνεχόμενο σήμα που μεταβάλεται συνεχώς με το χρόνο.

Το $e^{-i\omega t}$ δίνει τις συστατικές συχνότητες πάνω στον άξονα των x συνάρτηση του χρόνου t και είναι ίσο με $e^{-i\omega t} = cos(\omega t) - i \cdot sin(\omega t)$.

Μπορούμε να παραγάγουμε το μετασχηματισμό Fourier με την εφαρμογή της σχέσης 3.1 στο σήμα που μας ενδιαφέρει. Μπορούμε να δούμε πως δουλέυει περιοριζοντας την ανάλυσή μας σε απλά σήματα. Αυτό που θα παρατηρήσουμε είναι οτι τα περίπλοκα σήματα κατασκευάζονται αν προσθέσουμε πολλά διαφορετικά σήματα μαζί. Αν πάρουμε ένα παλμό που έχει μέγεθος A από την χρονική στιγμή $t = -T/2$ μέχρι τη χρονική στιγμή $t = T/2$ και είναι μηδέν σε όλα τα άλλα σημεία τότε ο παλμός δίνεται από τη σχέση :

$$\begin{cases} A & , \quad \frac{-T}{2} \leq t \leq \frac{T}{2} \\ 0 & , \quad else \end{cases} \qquad (3.2)$$

Για να πετύχουμε το μετασχηματισμό Fourier , κάνουμε αντικατάσταση στη σχέσης 3.1 με $p(t) = A$ μόνο για καθορισμένο χρόνο οπότε επιλέξαμε τα όρια του ολοκληρώματος να είναι η αρχή και το τέλος του παλμού. Έτσι ο μετασχηματισμός Fourier δίνεται από τη σχέση :

$$Fp(\omega) = \int_{\frac{-T}{2}}^{\frac{T}{2}} Ae^{-i\omega t}dt \qquad (3.3)$$

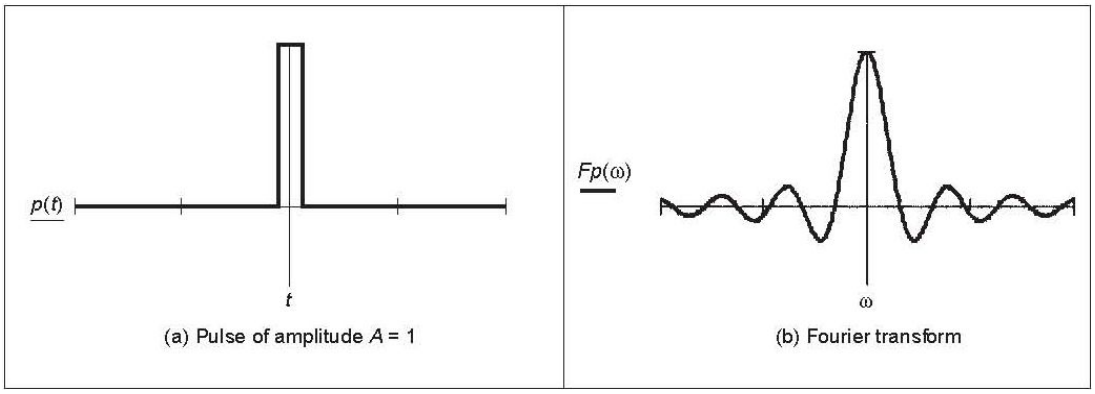
Με την επιλυσή του προκύπτει η πιο κάτω ισότητα :

$$Fp(\omega) = -\frac{Ae^{-i\omega\frac{T}{2}} - Ae^{i\omega\frac{T}{2}}}{i\omega} \qquad (3.4)$$

Χρησιμοποιούμε τη σχέση $sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i}$ και έχουμε :

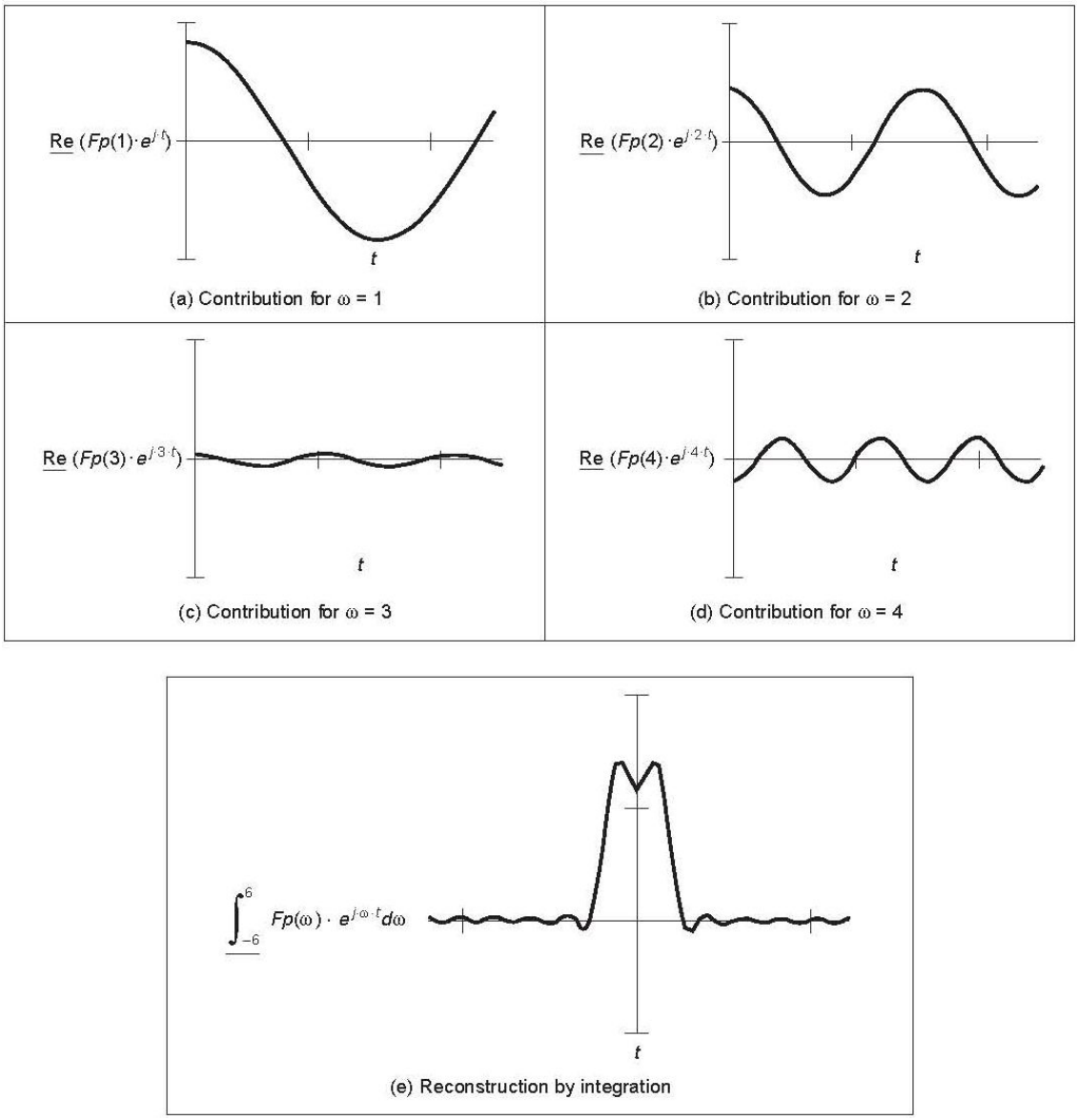
$$Fp(\omega) = \begin{cases} \frac{2A}{\omega} sin(\frac{\omega T}{2}), & \omega \neq 0 \\ AT \quad , & \omega = 0 \end{cases} \qquad (3.5)$$

Το αρχικό σήμα και ο μετασχηματισμός του φαίνονται στο Σχήμα 36. Η Σχέση 3.5 που έχει σχεδιαστεί στο Σχήμα 36a υποδεικνύει ότι ο παλμός σχηματίζεται από πολλές χαμηλές συχνότητες, που αποτελούν το κυρίως σώμα του παλμού, και από λίγες υψηλές που μας δίνουν τις κορυφές του παλμού. Οι περιοχές που καλύπτουν οι συχνότητες είναι συμμετρικές γύρο από την μηδενική συχνότητα (Σχήμα 36b). Ο σχεδιασμός του μετασχηματισμού Fourier καλείται το φάσμα του σήματος το οποίο μπορεί να θεωρηθεί συγγενικό με το φάσμα του φωτός.



Σχήμα 36: Ενας παλμός και ο μετασχηματισμός Fourier

Ο μετασχηματισμός Fourier μας λέει ποιές συχνότητες αποτελούν ένα σήμα σε μια χρονική περιόδο. Εάν αθροίσουμε μαζί ημιτονοειδή σήματα, όπως αυτό της Σχέσης 3.5 που προσδιορίζονται από το μετασχηματισμό Fourier τότε θα προκύψει το αρχικό μετασχηματισμένο σήμα. Αυτή η διαδικασία απεικονίζεται στο Σχήμα 37. Ο μετασχηματισμός Fourier δίνει μιγαδικό αριθμό οπότε εμείς σχεδιάζουμε μόνο το πραγματικό μέρος. Μια χαμηλή συχνότητα, για $\omega = 1$, στο Σχήμα a συνεισφέρει μια μεγάλη συνιστώσα του αρχικού σήματος. Μια πιο υψηλή συχνότητα, για $\omega = 2$, συνεισφέρει λιγότερο (Σχήμα 37b). Ομοίος βλέπουμε την συμπεριφορά του μετασχηματισμού στα Σχήματα 37c και 37d για $\omega = 3$ και $\omega = 4$. Όταν τα σήματα ενοποιούνται τότε προκύπτει ένα σήμα (Σχήμα 37e) που είναι παρόμοιο με το αρχικό σήμα (Σχήμα 36b). Στο Σχήμα 37 έχουμε εξετάσει μόνο τις συχνότητες στο διάστημα $\omega = -6$ και $\omega = 6$. Εάν το διάστημα ήταν μεγαλύτερο τότε θα περιλάμβανονταν πιο πολλές υψηλές συχνότητες και το τελικό αποτέλεσμα θα ηταν πιο κοντά στον πραγματικό παλμό.



Σχήμα 37: Κατασκευή σήματος από μικρότερα σήματα

Το αποτέλεσμα του μετασχηματισμού Fourier είναι μιγαδικός αριθμός. Οπότε η σχέση 3.1 γράφεται :

$$\int_{-\infty}^{\infty} p(t)e^{-i\omega t}dt = Re[Fp(\omega)] + i \cdot Im[Fp(\omega)] \qquad (3.6)$$

όπου $Re[Fp(\omega)]$ και $Im[Fp(\omega)]$ είναι το πραγματικό και το φανταστικό μέρος του μετασχηματισμού αντίστοιχα. Το μέγεθος του μετασχηματισμού δίνεται από τη σχέση :

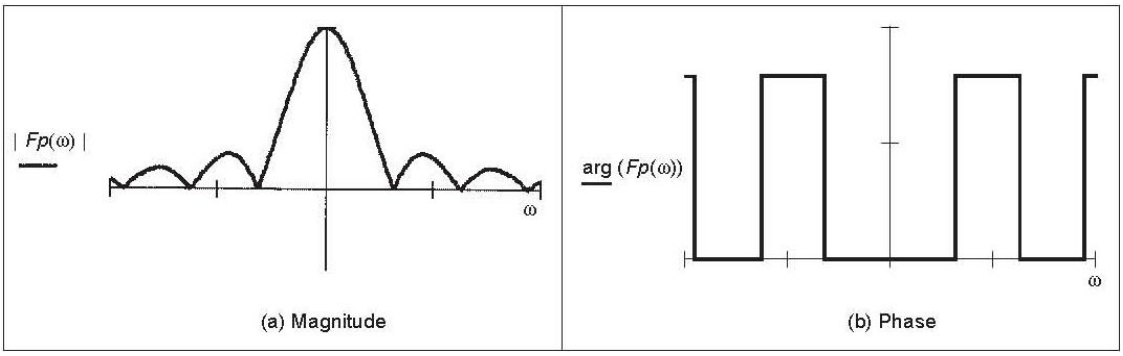
$$\left|\int_{-\infty}^{\infty} p(t)e^{-i\omega t}dt\right| = \sqrt{Re[Fp(\omega)]^2 + Im[Fp(\omega)]^2} \qquad (3.7)$$

Η φάση του μετασχηματισμού δίνεται από τη σχέση :

$$\langle\int_{-\infty}^{\infty} p(t)e^{-i\omega t}dt = tan^{-}1\frac{Im[Fp(\omega)]}{Re[Fp(\omega)]} \qquad (3.8)$$

Τα πρόσημα του πραγματικού και του φανταστικού μέρους μπορούν να χρησιμοποιηθούν για να καθορίσουμε σε πιο τεταρτημόριο βρίσκεται η φάση η οποία μπορεί να πάρει τιμές μεταξύ 0 και 2π. Το μέγεθος περιγράφει

την ποσότητα από κάθε τμήμα συχνότητας και η φάση περιγράφει το συγχρονισμό όταν τα τμήματα συχνότητας εμφανίζονται. Το μέγεθος και η φάση του μετασχηματισμού του παλμού παρουσιάζονται στο Σχήμα 38 όπου το μέγεθος επιστρέφει ένα θετικό μετασχηματισμό και η φάση είναι είτε 0 είτε 2π.



Σχήμα 38: Το μέγεθος και η φάση του μετασχηματισμού

Για να επιστρέψουμε από το μετασχηματισμένο σήμα στο αρχικό σήμα χρησιμοποιούμε τον αντίστροφο μετασχηματισμό Fourier. Αυτή είναι η διαδικασία στην οποία ανακατασκευάζουμε τον αρχικό παλμό από τις συστατικές του συχνότητες. Ο αντίστροφος μετασχηματισμός Fourier δίνεται από τη σχέση :

$$p(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Fp(\omega)e^{i\omega t}d\omega \qquad (3.9)$$

Οι Εξιδώσεις 3.1 και 3.9 δίνουν μια σχέση γνωστή και σαν ζευγάρι μετατροπής που μας επιτρέπει να μετασχηματίσουμε ένα σήμα σε συχνότητες και αντίστροφα. Με αυτή τη διαδικασία μπορούμε να εκτελέσουμε διαδικασίες στις συστατικές συχνότητες και το αποτέλεσμα να το χρησιμοποιήσουμε στο αρχικό σήμα και αντίστροφα. Μια σημαντική διαδικασία είναι η συνέλιξη. Η συνέλιξη δύο σημάτων $p_1(t)$ και $p_2(t)$ συμβολίζεται με $p_1(t) * p_2(t)$ και δίνεται από τη σχέση :

$$p_1(t) * p_2(t) = \int_{-\infty}^{\infty} p_1(\tau)p_2(t-\tau)d\tau \qquad (3.10)$$

όπου τ είναι η χρονική διαφορά των δύο σημάτων. Ο μετασχηματισμός Fourier της συνέλιξης είναι :

$$\begin{aligned} F[p_1(t) * p_2(t)] &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} p_1(\tau)p_2(t-\tau)d\tau\right)e^{-i\omega t}dt \\ &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} p_2(t-\tau)e^{-i\omega t}dt\right)p_1(\tau)d\tau \qquad (3.11) \end{aligned}$$

Αποδεικνύεται (όπως στην Παράγραφο 3.4.3) ότι :

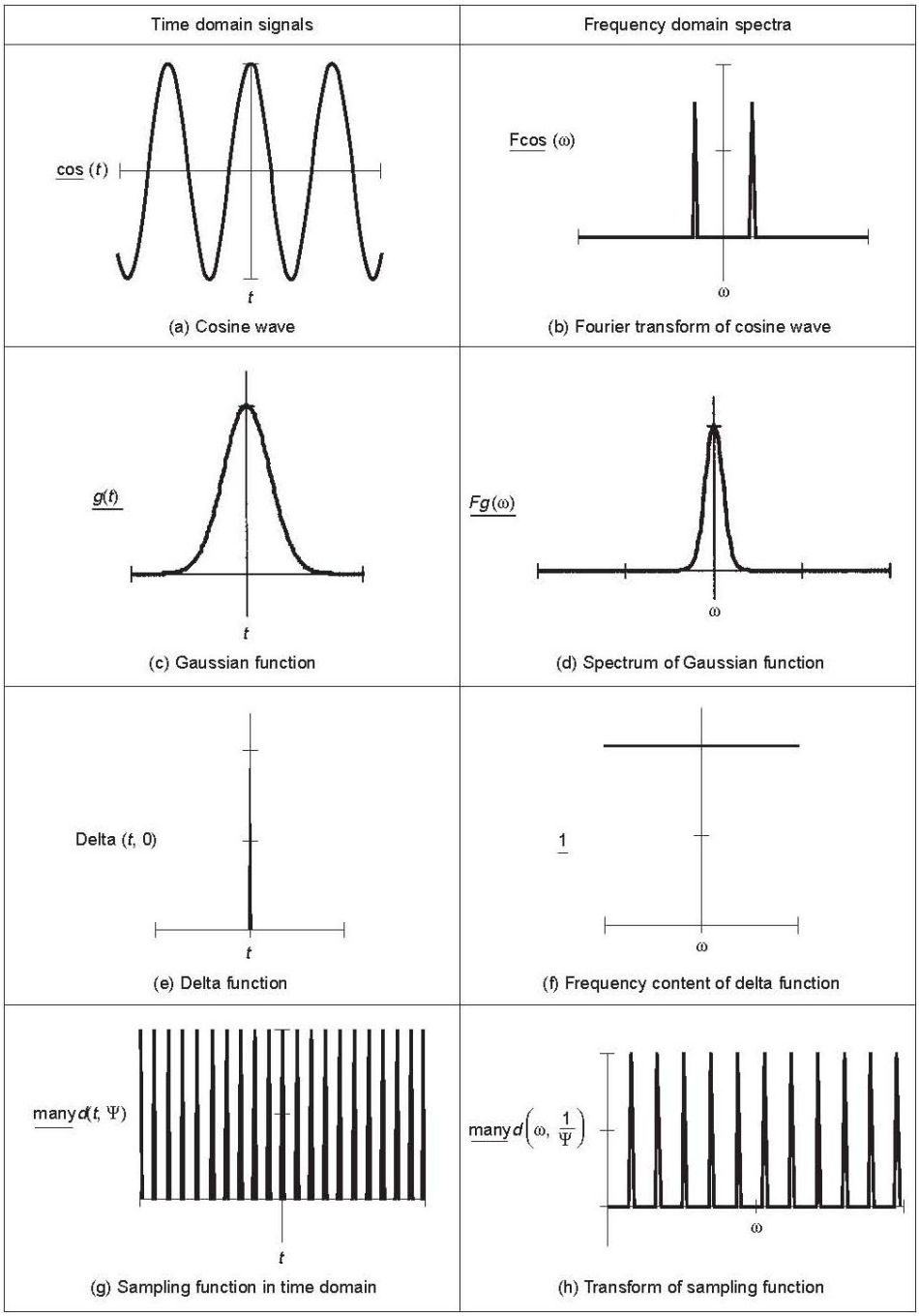
$$F[p_2(t-\tau)] = e^{-i\omega\tau}Fp_2(\omega)$$

Τότε έχουμε :

$$\begin{aligned} F[p_1(t) * p_2(t)] &= \int_{-\infty}^{\infty} Fp_2(\omega)p_1(\tau)e^{-i\omega\tau}d\tau \\ &= Fp_2(\omega)\int_{-\infty}^{\infty} p_1(\tau)e^{-i\omega\tau}d\tau \\ &= Fp_2(\omega)\cdot Fp_1(\omega) \qquad (3.12) \end{aligned}$$

Μπορούμε να συσχετίσουμε τη συνέλιξη δύο σημάτων με τη συνέλιξη εικόνας και προτύπου που έχουμε ήδη αναφέρει. Η διαδικασία της συνέλιξης δύο σημάτων μπορεί να χρησιμοποιηθεί και στις εικόνες, αφού αυτές πρώτα

μετασχηματιστούν σε μορφή συχνοτήτων όπως θα δούμε πιο κάτω, και παρουσιάζει ανάλογα αποτελέσματα με την συνέλιξη προτύπου και εικόνας. Το πλεονέκτημα της συνέλιξης με σήματα είναι η ταχύτητα αφού σε αυτή την περίπτωση η συνέλιξη πραγματοποιείται με πολλαπλασιασμό κάθε γραμμής και κάθε στήλης με το ενδιαφερόμενο σήμα. Στο Σχήμα 39 φαίνεται ο μετασχηματισμός Fourier γνωστών συναρτήσεων.



Σχήμα 39: Ζευγάρια μετασχηματισμών Fourier

Στο Σχήμα 39a έχουμε την συνάρτηση συνημιτόνου και στο Σχήμα 39b τον μετασχηματισμό Fourier αυτής. Ο μετασχηματισμός αυτής είναι δύο σημεία συχνότητας με την ίδια τιμή στον θετικό και τον αρνητικό άξονα. Τα Σχήματα 39c και 39d δείχνουν ότι ο μετασχηματισμός Fourier της συνάρτησης Gauss είναι μια άλλη συνάρτηση

Gauss . Το Σχήμα 39e είναι η συνάρτηση δέλτα :

$$delta(t-\tau) = \begin{cases} 1 & ,t=\tau \\ 0 & ,otherwise \end{cases} \qquad (3.13)$$

Στο Σχήμα 39f ο μετασχηματισμός της συνάρτησης δέλτα είναι μια ευθεία που περιλαμβάνει ίδιες τιμές για όλες τις συχνότητες. Αυτό μπορεί να εξηγηθεί από τη Σχέση 3.5 που αν ο παλμός είναι σύντομης διάρκειας τότε ο ημιτονοειδής μετασχηματισμός είναι ευρύς. Όταν ο παλμός γίνεται πιο λεπτός τότε ο μετασχηματισμός αρχίζει και γίνεται οριζόντιος. Τέλος το Σχήμα 39g είναι ένα σύνολο συναρτήσεων δέλτα και ο μετασχηματισμός τους που φαίνεται στο Σχήμα 39h είναι πάλι ένα σύνολο συναρτήσεων δέλτα αλλά με διαφορετική απόσταση μεταξύ τους [1,4].

3.3 Ο διακριτός μετασχηματισμός Fourier

3.3.1 Ο διακριτός μετασχηματισμός Fourier σε μια διάσταση

Στην επεξεργασία εικόνας ασχολούμαστε με διακριτά σημεία τα οποία είναι τοποθετημένα σε ακέραιες θέσεις πάνω σε κάποιον πίνακα. Οπότε ο μετασχηματιμός Fourier που έχουμε ορίσει μέχρι τώρα, ο οποίος αναφέρεται σε διαστήματα πραγματικών αριθμών, δεν μπορεί να εφαρμοστεί σε αυτή την περίπτωση. Θα πρέπει να ορίσουμε τον μετασχηματισμό Fourier για διακριτά σημεία. Αυτός είναι γνωστός σαν DFT (Discrete Fourier Transform) και ορίζεται από την σχέση :

$$Fp_u = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} p_x e^{-i\left(\frac{2\pi}{N}\right)xu} \qquad (3.14)$$

p_x είναι το αρχικό διάνυσμα, το οποίο είναι συνήθος δείγμα κάποιου συνεχόμενου σήματος.

N είναι το πλήθος των σημείων του διανύσματος p_x.

Fp_u είναι το μετασχηματισμένο διάνυσμα Fourier .

Ο πιο πάνω μετασχηματισμός είναι ο ανόλογος του συνεχή μετασχηματισμού Fourier . Το συνεχόμενο σήμα αντικαθίσταται από διακριτά σημεία, οι συνεχόμενες συχνότητες από διακριτές και το ολοκλήρωμα αντικαθίσταται από ένα άθροισμα. Εαν εφαρμόζαμε τον DFT σε ένα δείγμα ενός παλμού από το πρώτο μέχρι το $(N/2-1)$ στοιχείο όπου ο παλμός παύει τότε η πιο πάνω σχέση γίνεται:

$$Fp_u = \frac{1}{\sqrt{N}} \sum_{x=0}^{\frac{N}{2}-1} A e^{-i\left(\frac{2\pi}{N}\right)xu} \qquad (3.14)$$

Έχουμε τη σχέση από τη γεωμετρική πρόοδο :

$$\sum_{k=0}^{n} \alpha_0 r^k = \frac{\alpha_0(1-r^{n+1})}{1-r} \qquad (3.15)$$

Η σχέση 3.14 γίνεται :

$$Fp_u = \frac{A}{\sqrt{N}} \left(\frac{1-e^{-i\left(\frac{2\pi}{N}\right)\left(\frac{N}{2}\right)u}}{1-e^{-i\left(\frac{2\pi}{N}\right)u}} \right) \qquad (3.16)$$

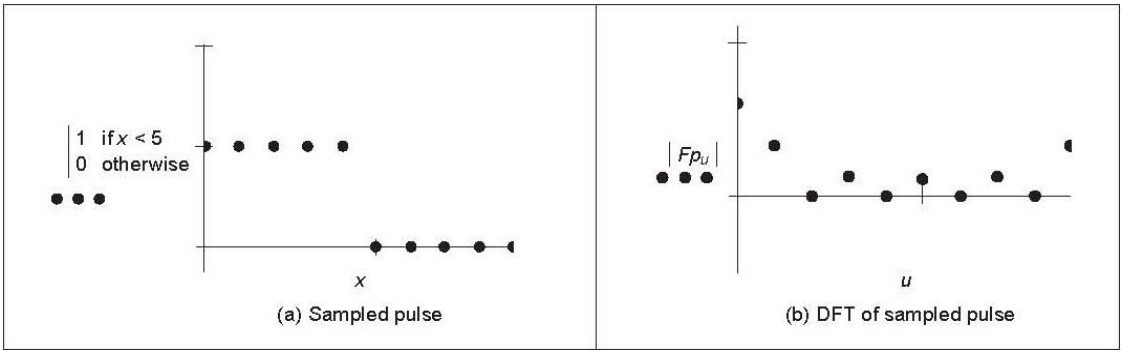
Με την απλοποίηση της σχέσης 3.16 έχουμε :

$$Fp_u = \frac{A}{\sqrt{N}} e^{-i\left(\frac{\pi u}{2}\right)\left(1-\frac{2}{N}\right)} \frac{sin(\pi u/2)}{sin(\pi u/N)} \qquad (3.17)$$

Το μέγεθος του μετασχηματισμού είναι :

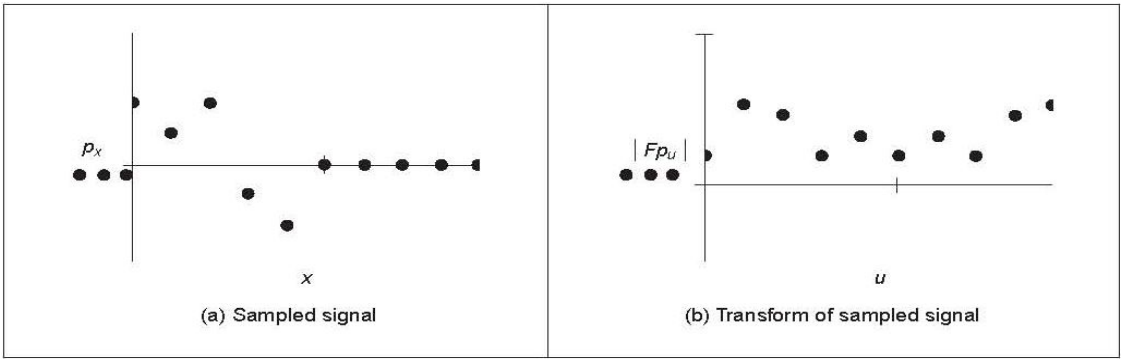
$$|Fp_u| = \frac{A}{\sqrt{N}} \left| \frac{sin(\pi u/2)}{sin(\pi u)/N} \right| \qquad (3.18)$$

Ο αρχικός παλμός, εκφρασμένος σε διακριτά σημεία, απεικονίζεται στο Σχήμα 40α και ο μετασχηματισμός του στο Σχήμα 40b .



Σχήμα 40:

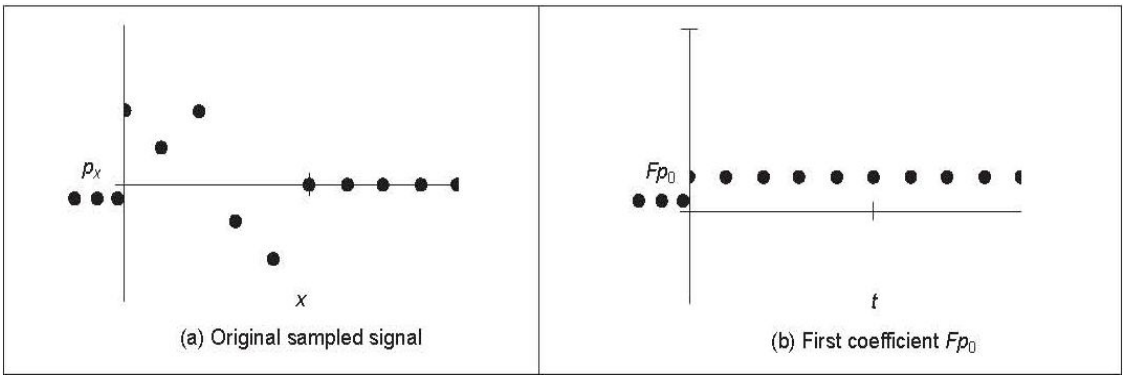
Το αποτέλεσμα που φαίνεται από το Σχήμα 40, μπορεί εύκολα να συγκριθεί με αυτό του συνεχή μετασχηματισμού Fourier Σχήμα 36, αφού και το σήμα που παράγεται από τον διακριτό μετασχηματισμό έχει μια ημιτονοειδή μορφή. Το διάνυσμα p_x μπορεί να μην έχει πάντα την μορφή του Σχήματος 40α. Πολλές φορές έχει διάταξη που δεν μπορούμε να την καθορίσουμε (Σχήμα 41a). Ο DFT ενός τέτοιου διανύσματος φαίνεται στο Σχήμα 41b .

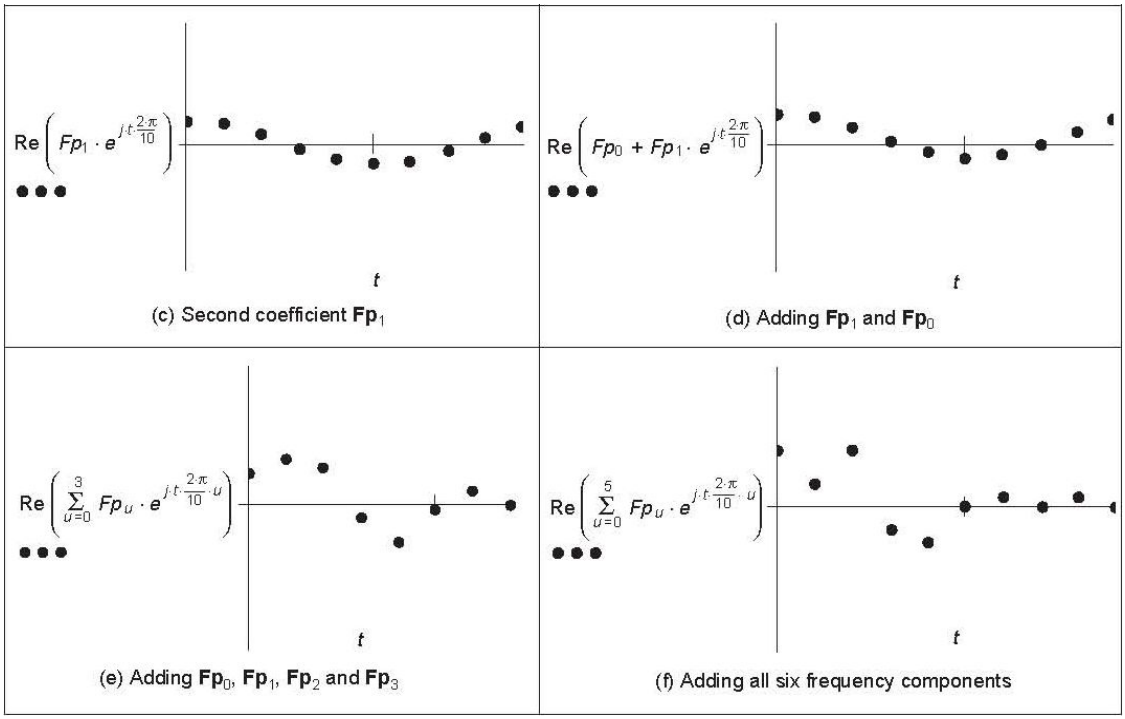


Σχήμα 41:

Όπως στον συνεχή έτσι και στον διακριτό μετασχηματισμό Fourier μπορούμε να κατασκευάσουμε το αρχικό σήμα απο το τροποποιημένο. Η διαδικασία ακατασκευής του αρχικού σήματος φαίνεται στο Σχήμα 42. Αυτό μπορούμε να το πετύχουμε με τον αντίστροφο DFT που δίνεται από την πιο κάτω σχέση :

$$p_x = \sum_{u=0}^{N-1} Fp_u e^{i\left(\frac{2\pi}{N}\right)ux} \qquad (3.19)$$





Σχήμα 42: Διαδικασία ανακατασκευής αρχικού σήματος

Ο μετασχηματισμός Fourier στο Σχήμα 41b μπορεί να χρησιμοποιηθεί για να κατασκευαστεί το αρχικό σήμα στό Σχήμα(41α). Αρχικά υπολογίζουμε τη συχνότητα Fp_0 (Σχήμα 42b). Αυτή είναι μια οριζόντια γραμμή. Στη συνέχεια υπολογίζουμε την συχνότητα Fp_1 (Σχήμα 42c) και την αθροίζουμε στην Fp_0 (Σχήμα 42c). Με τον ίδιο τρόπο συνεχίζεται αυτή η διαδικασία μέχρι να αθροίσουμε τις συχνότητες Fp_0 , Fp_1 $\cdots$, Fp_5. Η διαδικασία ολοκληρώνεται στο Σχήμα 42f όπου το τελικό αποτέλεσμα είναι σχεδόν ίδιο με το αρχικό σήμα. [1,4]

3.3.2 Ο διακριτός μετασχηματισμός Fourier σε δύο διαστάσεις

Η σχέση 3.14 δίνει τον διακριτό μετασχηματισμό Fourier σε μια διάσταση. Για να τον εφαρμόσουμε σε εικόνες θα πρέπει να τον ορίσουμε σε δύο διαστάσεις. Οπότε τώρα εργαζόμαστε πάνω και στους δύο άξονες x , y οι οποίοι χρειάζονται για να ορίσουμε την θέση ενός εικονοκύτταρου. Αυτό υπονοεί ότι χρειαζόμαστε δύο διαστάσεις u, v για να ορίσουμε τις συχνότητες του μετασχηματισμού. Δηλαδή θα έχουμε οριζόντιες και κάθετες συστατικές συχνότητες. Εάν μια εικόνα αποτελείται μόνο από κάθετες γραμμές τότε ο DFT σε δύο διαστάσεις ($2D\ DFT$) θα δώσει μόνο τις οριζόντιες συστατικές συχνότητες. Οι κάθετες συχνότητες θα είναι μηδέν. Αυτό συμβαίνει γιατί δεν υπάρχει καμία κάθετη παραλλαγή κατα μήκος του y άξονα. Ο διακριτός μετασχηματισμός Fourier σε δύο διαστάσεις δίνεται από τη σχέση :

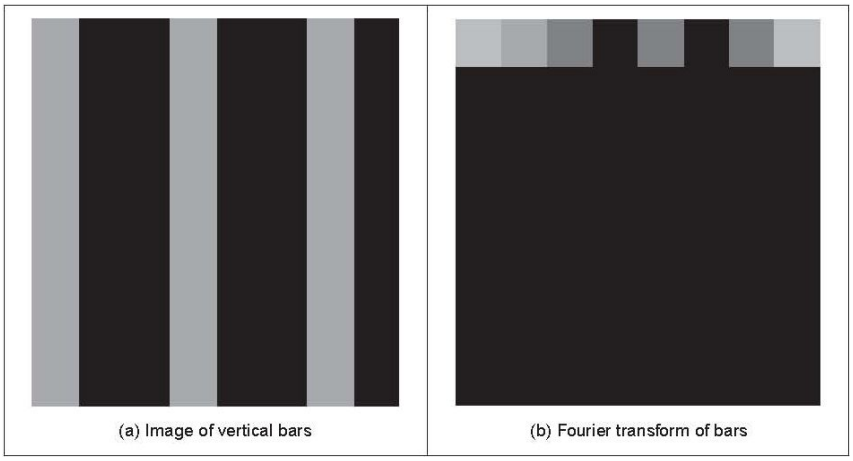
$$FP_{u,v} = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P_{x,y} e^{-\frac{2\pi}{N} i(ux+vy)} \qquad (3.20)$$

$P_{x,y}$ είναι η αρχική εικόνα.

$N \times N$ είναι το πλήθος των εικονοκυττάρων της $P_{x,y}$.

$FP_{u,v}$ είναι η μετασχηματισμένη εικόνα.

Για παράδειγμα εφαρμόζουμε τον 2D DFT σε μία εικόνα που έχει μόνο κάθετες γραμμές (Σχήμα 43α). Το αποτέλεσμα του μετασχηματισμού φαίνεται στο Σχήμα 43b . Το αποτέλεσμα έχει τη μορφή που περιμέναμε,



Σχήμα 43: Εφαρμογή του $2D\ DFT$ σε εικόνα με κάθετες γραμμές

δηκαδή έχει μόνο οριζόντιες συχνότητες και όχι κάθετες αφού δεν υπάρχει καμία κάθετη παραλλαγή κατα μήκος του y άξονα.

Όπως στον DFT έτσι και στον 2D DFT μπορούμε να ορίσουμε τον αντίστροφο του. Ο αντίστροφος 2D DFT ορίζεται από τη σχέση :

$$P_{x,y} = \sum_{u=0}^{N-1}\sum_{v=0}^{N-1} FP_{u,v} e^{\frac{2\pi}{N} i(ux+vy)} \qquad (3.21)$$

Μια σημαντική ιδιότητα των μετασχηματισμών Fourier είναι η αναπαραγωγή. Αυτό σημαίνει ότι ο μετασχηματισμός επαναλαμβάνει τις συχνότητες μέχρι το άπειρο. Για να το δείξουμε αυτό στον 2D DFT αρκεί να δείξουμε ότι $FP_{u+mN,v+nN} = FP_{u,v}$ όπου n , m ακέραιοι. Έχουμε ότι :

$$FP_{u+mN,v+nN} = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} P_{x,y} e^{-\frac{2\pi}{N} i((u+mN)x+(v+nN)y)} \qquad (3.22)$$

Η 3.22 γίνεται :

$$FP_{u+mN,v+nN} = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} P_{x,y} e^{-\frac{2\pi}{N} i(ux+vy)} \cdot e^{-2\pi i(mx+ny)} \qquad (3.23)$$

Αφού $mx + ny$ =ακέραιος, συνεπώς ο εκθέτης του e είναι πάντα ακέραιος που πολλαπλασιάζεται με το 2π.Άρα ισχύει :

$$e^{-2\pi i(mx+ny)} = 1 \qquad (3.24)$$

Απο την 3.24 η 3.23 γίνεται :

$$FP_{u+mN,v+nN} = \frac{1}{N}\sum_{x=0}^{N-1}\sum_{y=0}^{N-1} P_{x,y} e^{-\frac{2\pi}{N} i(ux+vy)} \Longrightarrow$$

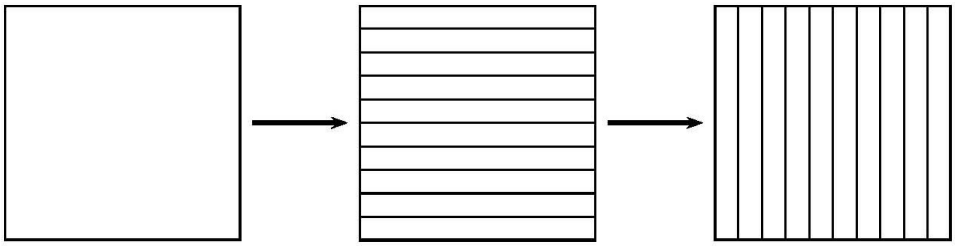
$$FP_{u+mN,v+nN} = FP_{u,v} \qquad (3.25)$$

Οπότε δείξαμε ότι η αναπαραγωγή ισχύει για τον 2D DFT . Οι πιο πάνω σχέσεις μπορούν να επεκταθούν και σε εικόνες που δεν είναι τετραγωνικές $N \times N$. Ο 2D DFT και ο αντίστροφος του για εικόνες $N \times M$ δίνονται από τις πιο κάτω σχέσεις :

$$FP_{u,v} = \frac{1}{\sqrt{NM}}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1} P_{x,y} e^{-\frac{2\pi}{NM} i(uxN+vyM)} \qquad (3.26)$$

$$P_{x,y} = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} FP_{u,v} e^{\frac{2\pi}{NM} i(uxN+vyM)} \qquad (3.27)$$

Οι Σχέσεις 3.26 και 3.27 έχουν ένα μειονέκτημα. Απαιτούν πολύ χρόνο για να εκτελεστούν, ιδιαίτερα για μεγάλες εικόνες. Αυτό έρχεται να διορθώσει η ιδιότητα του διαχωρισμού. Ο 2D DFT μπορεί να υλοποιηθεί με την εφαρμογή μόνο του DFT. Πρώτα εφαρμόζουμε τον DFT σε όλες τις γραμμές της εικόνας και στην συνέχεια στο αποτέλεσμα που προκύπτει εφαρμόζουμε πάλι τον DFT αλλά αυτή τη φορά σε όλες τις στήλες. Στο Σχήμα 40 φαίνεται αυτή η διαδικασία.



(a) Original image (b) DFT of each row of (a) (c) DFT of each column of (b)

Σχήμα 44: Υλοποίηση του 2D DFT με εφαρμογή του DFT

Το πρόβλημα με τον μετασχηματισμό Fourier είναι ότι παράγεται μια εικόνα που είναι δύσκολο να ερμηνεύσουμε τι απεικονίζει. Ο μετασχηματισμός Fourier μιας εικόνας τη μετατρέπει σε μορφή συχνοτήτων. Ο $2D\ DFT$ τοποθετεί τις χαμηλές συχνότητες στις γωνίες του μετασχηματισμού Σχήμα 45.

Σχήμα 45:

Απο το σχήμα μπορούμε να παρατηρήσουμε ότι οι λευκές τελείες οι οποίες είναι οι χαμηλές συχνότητες έχουν συγκεντρωθεί στις άκρες του μετασχηματισμού. Όταν πρόκειται να δουλέψουμε με εικόνες που έχουν μετασχηματιστεί από τον $2D\ DFT$ είναι πιθανό να θέλουμε να έχουμε μια καλύτερη όψη αυτών. Μια μετατροπή που θα τοποθετούσε τις χαμηλές συχνότητες στο κέντρο της εικόνας είναι πιθανό να εξυπηρετούσε. Ένας τρόπος είναι να περιστρέψουμε κάθε ένα από τα τέσσερα τεταρτημόρια του μετασχηματισμού Fourier κατά 180 μοίρες. Μια εναλλακτική λύση είναι να τροποποιήσουμε την αρχική εικόνα έτσι ώστε οι χαμηλές συχνότητες του μετασχηματισμού να μετατοπιστούν προς το κέντρο. Για να το πετύχουμε αυτό πολλαπλασιάζουμε κάθε εικονοκύτταρο $P_{x,y}$ της αρχικής εικόνας με το -1^{x+y}.Ο νέος μετασχηματισμός είναι :

$$FP_{u,v} \times (-1^{x+y}) = FP_{u,v} = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P_{x,y} e^{-\frac{2\pi}{N} i(ux+vy)} \times (-1^{x+y})$$

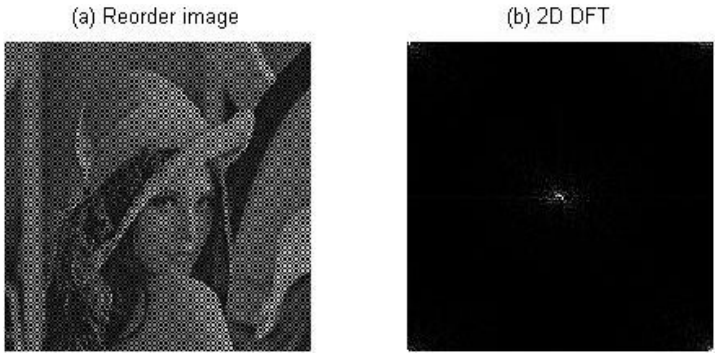
Έχουμε ότι $cos(-\pi) = -1$ τότε $e^{-i\pi} = -1$.Άρα η πιο πάνω σχέση γίνεται :

$$FP_{u,v} \times (-1^{x+y}) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P_{x,y} e^{-\frac{2\pi}{N} i(ux+vy)} \cdot e^{-i\pi(x+y)} \Longrightarrow$$

$$FP_{u,v} \times (-1^{x+y}) = \frac{1}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} P_{x,y} e^{-i(\frac{2\pi}{N})((u+\frac{N}{2})x+(v+\frac{N}{2})y)} \Longrightarrow$$

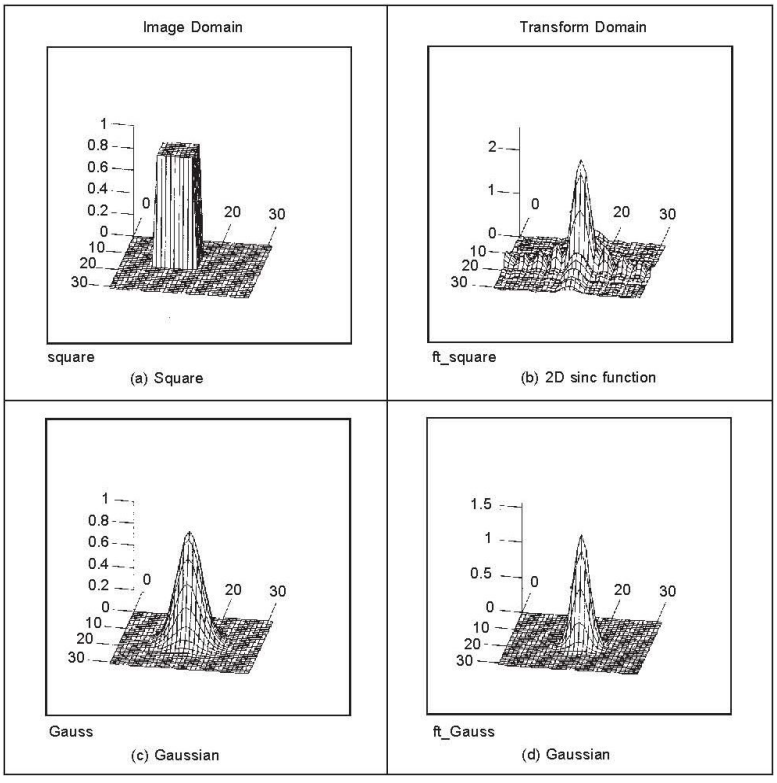
$$FP_{u,v} \times (-1^{x+y}) = FP_{u+\frac{N}{2},v+\frac{N}{2}} \qquad (3.28)$$

Το αποτέλεσμα αυτού του μετασχηματισμού φαίνεται στο Σχήμα 46.



Σχήμα 46:

Στο πιο πάνω σχήμα παρατηρούμε ότι οι χαμηλές συχνότητες έχουν μετατοπιστεί προς το κέντρο του μετασχηματισμού. Τώρα η μετασχηματισμένη εικόνα είναι πιο εύκολο να παρατηρηθεί. Θα πρέπει να τονίσουμε ότι αυτός ο μετασχηματισμός μπορεί να βελτιώσει μόνο την απεικόνιση και δεν επηρεάζει τις συστατικές συχνότητες της εικόνας. Στο Σχήμα 47 φαίνονται τα αποτελέσματα της εφαρμογής του 2D DFT σε ένα ορθογώνιο και στην συνάρτηση Gauss . Ο 2D DFT του ορθογωνίου είναι μια συνάρτηση ημιτόνου, ενώ ο 2D DFT της συνάρτησης Gauss είναι πάλι συνάρτηση Gauss [1,4].



Σχήμα 47:

3.4 Ιδιότητες του μετασχηματισμού Fourier

3.4.1 Μετατόπιση

Ο μετασχηματισμός Fourier είναι ανεξάρτητος από τη θέση των χαρακτηριστικών γνωρισμάτων της αρχικής εικόνας. Αν μετατοπίσουμε τα χαρακτηριστικά της αρχικής εικόνας κατα ένα σταθερό αριθμό τότε ο μετασχηματισμός Fourier παραμένει ίδιος. Αυτή η ιδιότητα του μετασχηματισμού ονομάζεται ιδιότητα της μετατόπισης. Αφού μπορούμε να υλοποιήσουμε τον $2D\ DFT$ με πολλαπλή εφαρμογή του DFT θα ασχοληθούμε με την μετατόπιση των στοιχείων ενός διανύσματος. Έστω το αρχικό διάνυσμα $p(x)$ και a μια σταθερή μετατόπιση των στοιχείων του. Τότε το νέο διάνυσμα θα είναι η $p(x+a)$. Ο μετασχηματισμός Fourier του αρχικού διανύσματος είναι $Fp(u)$ ενώ του μετατοπισμένου είναι $\bar{F}p(u)$. Για το μετασχηματισμό Fourier του αρχικού διανύσματος έχουμε :

$$Fp(u) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} p(x) e^{-i\left(\frac{2\pi}{N}\right)xu}$$

Οπότε ο μετασχηματισμός Fourier του μετατοπισμένου διανύσματος θα είναι :

$$\bar{F}p(u) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} p(x+a) e^{-i\left(\frac{2\pi}{N}\right)(x+a)u} \Longrightarrow$$

$$\bar{F}p(u) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} p(x+a) e^{-i\left(\frac{2\pi}{N}\right)(xu+au)} \Longrightarrow$$

$$\bar{F}p(u) = e^{-i\left(\frac{2\pi}{N}\right)au} \cdot \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} p(x+a) e^{-i\left(\frac{2\pi}{N}\right)xu} \quad (3.29)$$

Το $x+a$ δειλώνει κάποια θέση μέσα στο διάστημα [1,N]. Οπότε θέτω $k = x+a$ όπου $1 \leq k \leq N$. Αφού έχουμε υποθέσει ότι το διάνυσμα x έχει μήκος N μπορούμε να πούμε ότι ο μετασχηματισμός Fourier στην πρώτη περίπτωση ξεκινάει από την αρχή του διανύσματος ενώ στην δεύτερη περίπτωση ξεκινάει από την θέση κ. Γραφικά αυτό μπορεί να φανεί πιο κάτω:

$$1\ ,\ 2\ ,\ 3\ , \cdots ,\ k-1\ ,\ k\ ,\ k+1\ , \cdots ,\ N$$

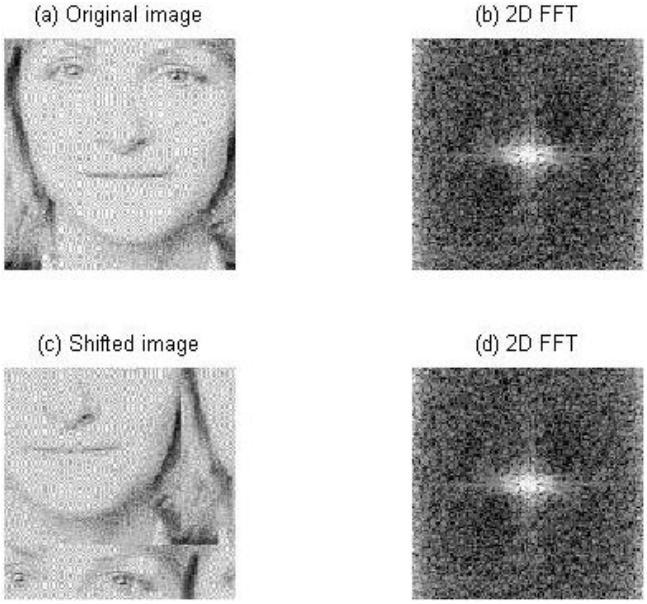
Ο $Fp(u)$ ξεκινάει από το 1 και καταλήγει στο N ενώ ο $\bar{F}p(u)$ ξεκινάει από το k φθάνει μέχρι το N και στη συνέχει πάει από το 1 μέχρι το $k-1$. Σύμφωνα με αυτά και με τη σχέση $x = k-a$ η 3.29 γίνεται :

$$\bar{F}p(u) = e^{-i\left(\frac{2\pi}{N}\right)au} \cdot \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} p(k) e^{-i\left(\frac{2\pi}{N}\right)(k-a)u} \Longrightarrow$$

$$\bar{F}p(u) = e^{-i\left(\frac{2\pi}{N}\right)au} \cdot e^{i\left(\frac{2\pi}{N}\right)au} \cdot \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} p(k) e^{-i\left(\frac{2\pi}{N}\right)ku} \Longrightarrow$$

$$\bar{F}p(u) = Fp(u) \qquad (3.30)$$

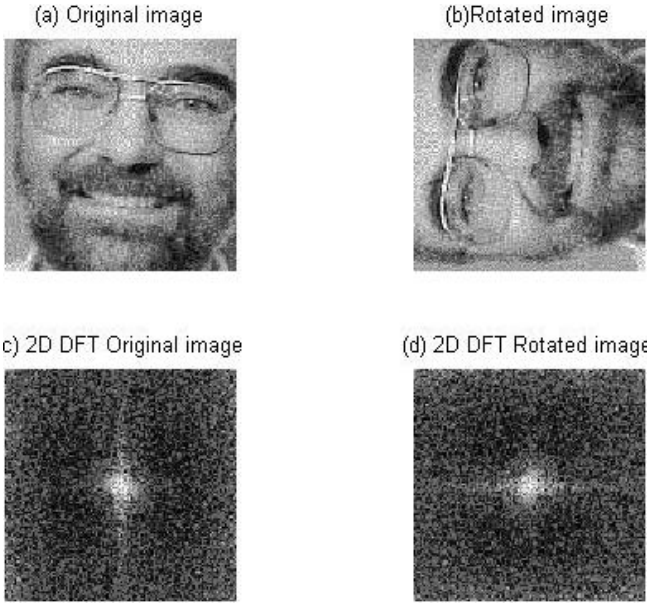
Αυτό δείχνει ότι ο μετασχηματισμός Fourier του αρχικού και του μετατοπισμένου διανύσματος είναι ο ίδιος. Αυτό επεκτείνεται και στις εικόνες. Στο Σχήμα 48 φαίνεται η ιδιότητα της μετατόπισης.



Σχήμα 48: Ιδιότητα της μετατόπισης

3.4.2 Περιστροφή

Ο μετασχηματισμός Fourier περιστρέφεται όταν περιστρέφεται η αρχική εικόνα. Με το μετασχηματισμό Fourier μπορούμε να προσδιορίσουμε την κατεύθυνση των χαρακτηριστικών γνωρισμάτων που υπάρχουν σε μία εικόνα, ακόμα και σε περιπτώσεις που είναι δύσκολο να προσδιοριστεί το περιεχόμενο της εικόνας. Επίσης γνωρίζοντας την αρχική κατεύθυνση μπορούμε να ορίσουμε μια νέα κατεύθυνση που μας εξυπηρετεί. Η αλλαγή της κατεύθυνσης στον μετασχηματισμό Fourier φαίνεται στο Σχήμα 49.



Σχήμα 49: Ιδιότητα της περιστροφής

Το πρόσωπο στο Σχήμα 49α περιστρέφεται κατά 90 μοίρες και δίνει το Σχήμα 49b . Συγκρίνοντας το μετασχηματισμό της πρώτης εικόνας (Σχήμα 49c) και το μετασχηματισμό της δεύτερης (Σχήμα 49d) παρατηρούμε ότι

ο μετασχηματισμός έχει και αυτός περιστραφεί κατά 90 μοίρες. Ο γενικός άξονας στο Σχήμα 49a είναι σχεδόν κάθετος και αυτός έρχεται σε σχεδόν οριζόντια θέση όταν περιστρέψουμε την εικόνα.

3.4.3 Αλλαγή κλίμακας

Απο τον ορισμό, ο χρόνος είναι αμοιβαίος με τη συχνότητα. Υποθέτουμε ότι συμπιέζουμε μια εικόνα, που είναι αντίστοιχο με τη μείωση του χρόνου, τότε οι συστατικές της συχνότητες θα εξαπλωθούν. Αντίστοιχα με την παράγραφο 3.3.1 θα ασχοληθούμε με διάνυσμα (παλμός). Υποθέτουμε ότι συμπιέζουμε χρονικά ένα παλμό $p(t)$ πολλαπλασιάζοντας το χρόνο t με μια σταθερή ποσότητα λ οπότε προκύπτει ο παλμός $p(\lambda t)$. Ο μετασχηματισμός Fourier για τον αρχικό παλμό είναι $Fp(\omega)$ και για τον τελικό είναι $\bar{F}p(\omega)$. Ο μετασχηματισμός Fourier του αρχικού παλμού είναι :

$$Fp(\omega) = \int_{-\infty}^{\infty} p(t)e^{-i\omega t}dt$$

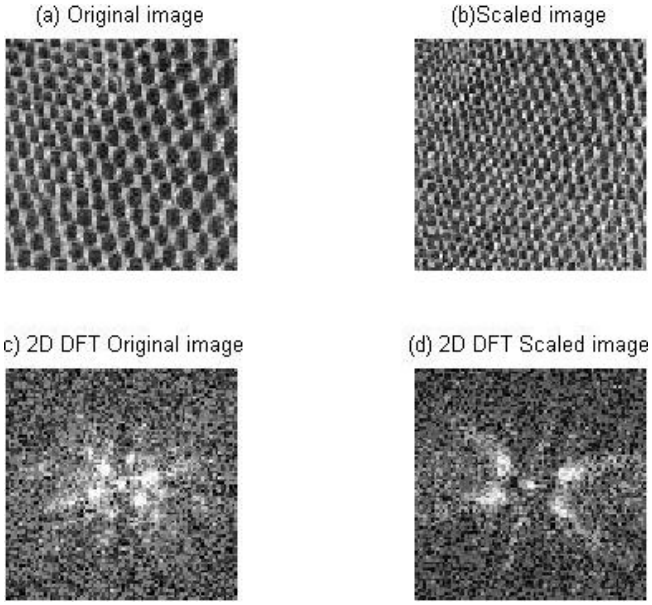
Για τον συμπιεσμένο παλμό θέτω $k = \lambda t$. Τότε $dt = 1/(\lambda)dk$. Τα όρια ολοκλήρωσης δεν αλλάζουν οπότε για τον μετασχηματισμό $\bar{F}p(\omega)$ έχουμε :

$$\bar{F}p(\omega) = \int_{-\infty}^{\infty} p(k)e^{-i\omega k}(1/\lambda)dk \Longrightarrow$$

$$\bar{F}p(\omega) = (1/\lambda) \cdot \int_{-\infty}^{\infty} p(k)e^{-i\omega k}dk \Longrightarrow$$

$$\bar{F}p(\omega) = (1/\lambda)Fp(\omega) \qquad (3.31)$$

Στο Σχήμα 50 φαίνεται η συμπίεση μιας εικόνας και η συμπεριφορά του μετασχηματισμού Fourier .



Σχήμα 50: Αλλαγή κλίμακας

Στο Σχήμα 50α έχουμε ένα φράκτη ο οποίος μειώνεται στην κλίμακα Σχήμα 50b . Ο μετασχηματισμός Fourier της αρχικής εικόνας φαίνεται στο Σχήμα 50c και αποκαλύπτει ότι οι μεγάλες συστατικές συχνότητες της αρχικής εικόνας είναι τακτοποιημένος σε ένα αστεροειδές σχέδιο. Συνέπεια της μείωσης της κλίμακας της αρχικής εικόνας είναι η εξάπλωση των συχνοτήτων της συμπιεσμένης εικόνας (Σχήμα 50d). Με αυτό το μετασχηματισμό μπορούμε να πάρουμε πληροφορίες για την απόσταση που βλέπουμε ένα αντικείμενο και μας επιτρέπει να κάνουμε τροποποιήσεις στην κλιμακά του. [1,10]

4 Περιγραφή αντικειμένων

4.1 Γενική επισκόπηση

Τα αντικείμενα αντιπροσωπεύονται σαν μια συλλογή από εικονοκύτταρα μέσα σε μια εικόνα. Έτσι θα πρέπει να περιγράψουμε τις ιδιότητες των ομάδων από εικονοκύτταρα. Η περιγραφή είναι συχνά ένα σύνολο από αριθμούς που είναι γνωστοί σαν αριθμοί περιγραφής του αντικειμένου. Από αυτούς μπορούμε να συγκρίνουμε και να αναγνωρίσουμε αντικείμενα με απλό ταίριασμα των αριθμών περιγραφής ενός αντικειμένου με τους αριθμούς περιγραφής ενός γνωστού αντικειμένου. Για να μπορούν όμως να φανούν χρήσιμοι για την περιγραφή ενός αντικειμένου οι αριθμοί περιγραφής θα πρέπει να έχουν τέσσερις σημαντικές ιδιότητες. Πρώτον θα πρέπει να ορίζουν ένα πλήρες σύνολο. Αυτό σημαίνει ότι τα δύο αντικείμενα έχουν τους ίδιους αριθμούς περιγραφής αν και μόνο αν έχουν το ίδιο σχήμα. Δεύτερον είναι η ιδιότητα της ομοιότητας. Δηλαδή θα πρέπει να μπορούμε να αναγνωρίσουμε όμοια ανικείμενα όταν έχουν όμοιους αριθμούς περιγραφής. Τρίτον θα πρέπει να έχουν αναλλοίωτες ιδιότητες. Μια τέτοια είναι η ιδιότητα της περιστροφής. Οι αριθμοί περιγραφής θα πρέπει να μην επηρεάζονται από την περιστροφή των αντικειμένων. Επίσης μια ακόμα αναλλοίωτη ιδιότητα είναι αυτή της τοποθεσίας. Ανεξάρτητα με την θέση του αντικειμένου οι αριθμοί περιγραφής θα πρέπει να μένουν ίδιοι. Η τέταρτη ιδιότητα είναι ότι οι αριθμοί περιγραφής θα πρέπει να είναι ένα περιεκτικό σύνολο. Δηλαδή θα πρέπει να αντιπροσωπεύουν την ουσία του αντικειμένου με έναν αποδοτικό τρόπο. Για παράδειγμα μπορεί να περιέχουν μόνο γενικές πληροφορίες για ένα αντικείμενο που θα πρέπει να ξεχωρίζει από ένα μεγάλο σύνολο αντικειμένων. Το πλήθος των πληροφοριών που μπορούμε να αντλήσουμε για να περιγράψουμε ένα αντικείμενο είναι συνήθως μικρό σε σχέση με το σύνολο των πραγματικών του χαρακτηριστικών. Η κατασκευή ενός τέτοιου συνόλου είναι πολύ δύσκολη. Έτσι σκοπός κάθε φορά είναι η άντληση εκείνων των χαρακτηριστικών που θα μας εξυπηρετούν καλύτερα σε κάθε περίπτωση. Σε αυτή την παράγραφο θα ασχοληθούμε με τον χαρακτηρισμό του συνόρου ενός σχήματος και με την περιοχή που καλύπτει.

4.2 Περιγραφή συνόρου

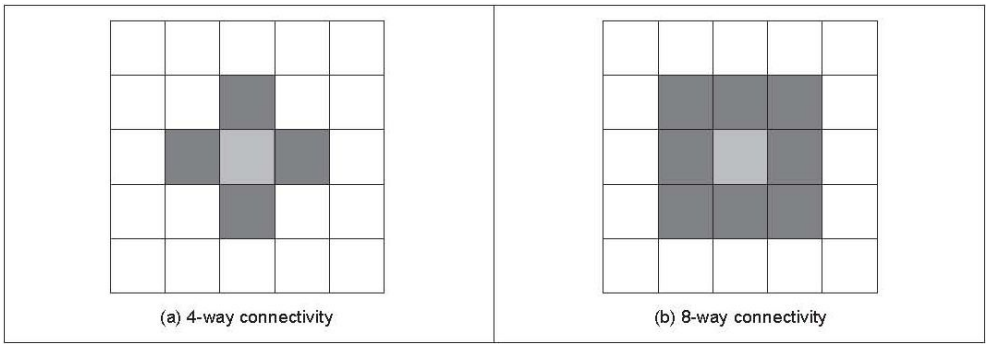
4.2.1 Σύνορο και περιοχή

Μια περιοχή περιγράφει συνήθως εσωτερικά σημεία τα οποία περιβάλλονται από ένα σύνορο το οποίο καλείται συχνά το περίγραμμα της περιοχής. Η μορφή του περιγράμματος περιγράφει το σχήμα του αντικειμένου. Ορίζουμε οτι ένα σημείο ανήκει στο σύνορο ενός αντικειμένου εάν είναι μέρος της περιοχής που καλύπτει το αντικείμενο και υπάρχει τουλάχιστον ένα γειτονικό του εικονοκύτταρο που δεν είναι μέρος της περιοχής που καλύπτει το αντικείμενο. Το περίγραμμα συνήθως βρίσκεται ακολουθώντας το σύνορο της περιοχής του αντικειμένου. Πρώτα βρίσκουμε ένα σημείο του συνόρου και στη συνέχεια περπατάμε αριστερόστροφα ή δεξιόστροφα για να βρούμε το επόμενο σημείο που ανήκει στο σύνορο του αντικειμένου. Για να ορίσουμε τα εσωτερικά σημεία και τα σημεία του περιγράμματος, θα πρέπει να εξετάσουμε τις γειτονικές σχέσεις των εικονοκυττάρων. Αυτές οι σχέσεις περιγράφονται με κανόνες συνδετικότητας. Υπάρχουν δύο διαδεδομένοι τρόποι για να ορίσουμε την συνδετικότητα:

(α) Συνδετικότητα με τέσσερις γείτονες : Μόνο οι άμεσοι γείτονες ελέγχονται για συνδετικότητα.

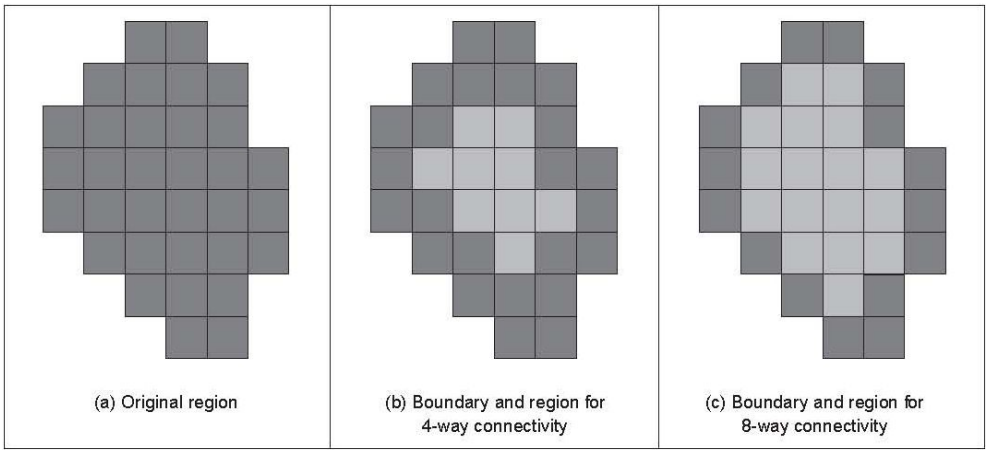
(β) Συνδετικότητα με οκτώ γείτονες : Και τα οκτώ γειτονικά εικονοκύτταρα που περιβάλλουν ένα εικονοκύτταρο ελέγχονται για συνδετικότητα.

Αυτοί οι δύο τύποι συνδετικότητας παρουσιάζονται στο Σχήμα 51. Σε αυτό το σχήμα, τα εικονοκύτταρα που είναι με σκούρο χρώμα είναι οι γείτονες των εικονοκύτταρων που είναι με πιο ανοιχτό χρώμα. Στο Σχήμα 51α έχουμε την περίπτωση της συνδετικότητας με τέσσερις γείτονες. Σε αυτή την περίπτωση οι πιθανοί γείτονες ενός εικονοκύτταρου βρίσκονται βόρεια, ανατολικά, νότια ή δυτικά αυτού. Στην οκτώ συνδετικότητα (Σχήμα 51b) υπάρχουν τέσσερις επιπλέον κατευθύνσεις : βορειοανατολικά, νοτιοανατολικά, νοτιοδυτικά και βορειοδυτικά.



Σχήμα 51: 4 και 8 συνδετικότητα

Το περίγραμμα και η περιοχή μπορούν να καθοριστούν χρησιμοποιώντας και τους δύο τύπους συνδετικότητας και είναι πάντα συμπληρωματικοί. Αν τα εικονοκύτταρα του περιγράμματος συνδέονται με τέσσερις γείτονες, τότε τα εικονοκύτταρα της περιοχής θα συνδέονται με οκτώ γείτονες και αντίστροφα. Αυτή η σχέση παρουσιάζεται στο Σχήμα 52. Στο Σχήμα 52α έχουμε μια περίοχή με σκούρο χρώμα. Στο Σχήμα 52b έχουμε χρησιμοποιήσει την συνδετικότητα με τέσσερις γείτονες. Σε αυτή την περίπτωση το περίγραμμα που προκύπτει έχει κλιμακωτή διάταξη. Στο Σχήμα 52c έχουμε χρησιμοποιήσει την συνδετικότητα με οκτώ γείτονες. Παρατηρούμε ότι το περίγραμμα διατηρεί μια διαγώνια μορφή στην οποία συνδέονται οι γωνίες των γειτόνων.

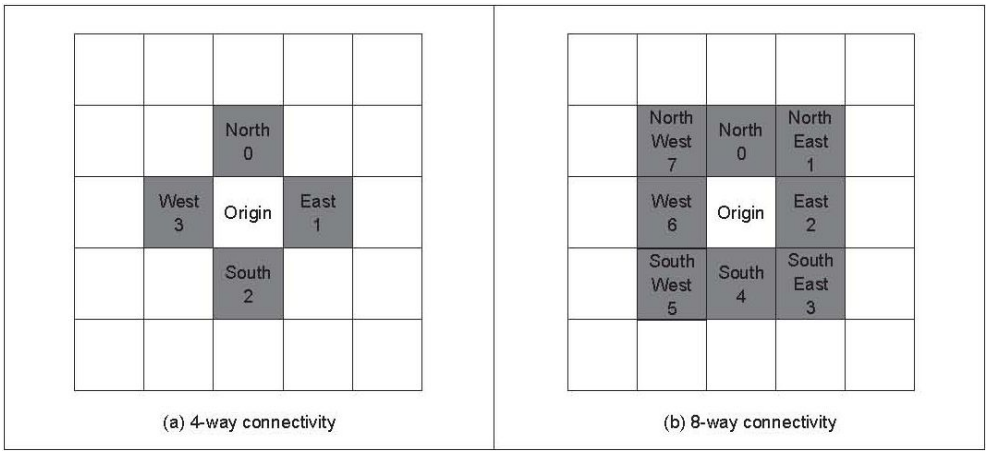


Σχήμα 52:

4.2.2 Κώδικες αλυσίδων

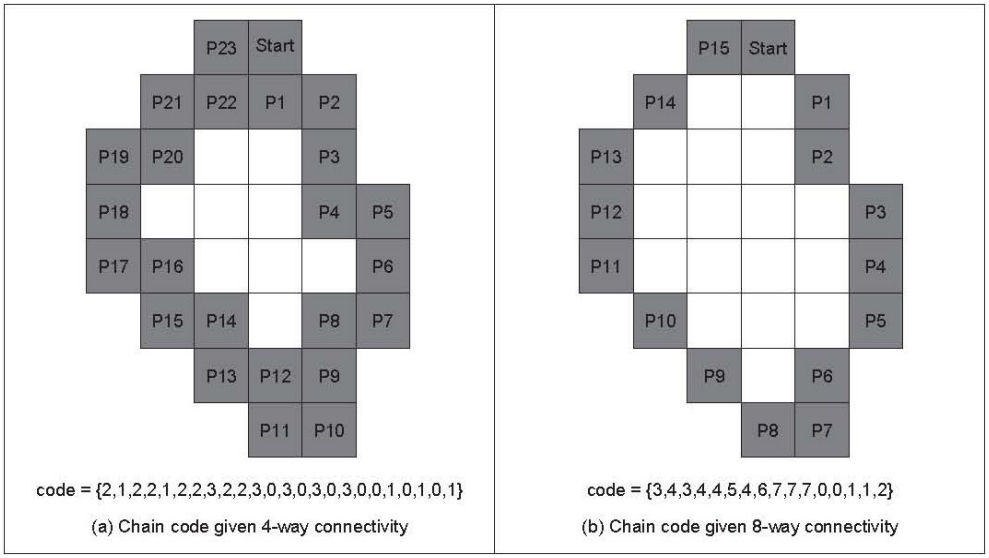
Προκειμένου να ληφθεί μια αντιπροσώπευση ενός περιγράμματος, μπορούμε απλά να αποθηκεύσουμε τις συντεταγμένες των σημείων που αποτελούν το περίγραμμα. Έναλλακτικά μπορούμε να αποθηκεύσουμε τη σχετική θέση των διαδοχικών σημείων που ανήκουν στο περίγραμμα. Αυτή είναι η βασική ιδέα πίσω από τους κώδικες των αλυσίδων. Ουσιαστικά το σύνολο των εικονοκυττάρων που υπάρχουν στο περίγραμμα ενός σχήματος μεταφράζεται σαν ένα σύνολο από συνδέσεις μεταξύ αυτών των εικονοκυττάρων. Για να βρούμε τον κώδικα ενός περιγράμματος ακολουθούμε μια απλή διαδικασία. Πρώτα επιλέγουμε ένα σημείο του περιγράμματος και το ονομάζουμε *start*. Στη συνέχεια θα πρέπει να ορίσουμε μια σχέση που θα λέει σε πιά κατεύθυνση είναι το επόμενο εικονοκύτταρο του περιγράμματος. Το επόμενο εικονοκύτταρο είναι ένα από τα παρακείμενα σημεία που θα βρίσκεται σε μια από τις τέσσερις ή οκτώ κατευθύνσεις που αναφέραμε πρίν. Μπορούμε να ονομάσουμε κάθε κατεύθυνση με ένα αριθμό από το μηδέν μέχρι το επτά. Οπότε η κατεύθυνση του επόμενου εικοκύτταρου θα είναι κάποιος ακέραιος αριθμός μεταξύ μηδέν και επτά. Περπατάμε κυκλικά το περίγραμμα και για κάθε εικονοκύτταρο γράφουμε στον κώδικα την κατεύθυνση του επόμενου εικοκύτταρου. Μόλις συναντήσουμε το

εικονοκύτταρο που έχει χαρακτηριστεί *start* σταματάει η διαδικασία κατασκευής του κώδικα. Αυτό το κριτήριο τερματισμού αναφέρεται σε περιγράμματα που κλείνουν κύκλο. Ανάλογα με το πρόβλημα που έχουμε να αντιμετωπίσουμε διαμορφώνουμε και τα κριτήρια τερματισμού. Οι κατευθύνσεις για τις δύο μορφές συνδετικότητας που έχουμε αναφέρει φαίνονται στο Σχήμα 53.



Σχήμα 53: Κωδικοποίηση των κατευθύνσεων

Στο Σχήμα 54α έχει εφαρμοστεί η μέθοδος με τους τέσσερις γείτονες και έχει εφαρμοστεί η κωδικοποίηση των κατευθύνσεων του Σχήματος 53α. Ο αλγόριθμος ελέγχει πρώτα το εικονοκύτταρο που βρίσκεται στη κατεύθυνση με κωδικό μηδέν. Αν αυτό το εικονοκύτταρο ανήκει στο περίγραμμα το κρατάει γράφει στον κώδικα το μηδέν και πηγαίνει στο εικονοκύτταρο αυτής της κατεύθυνσης. Αν δεν ανήκει στο περίγραμμα ελεγχει το εικονοκύτταρο που βρίσκεται στη κατεύθυνση με κωδικό ένα, και συνεχίζει την ίδια διαδικασία κυκλικά, δηλαδή μετά στην κατεύθυνση δύο και μετά στην τρία. Στο Σχήμα 54b έχει χρησιμοποιηθεί η μέθοδος με τους οκτώ γείτονες και έχει εφαρμοστεί η κωδικοποίηση των κατευθύνσεων του Σχήματος 53(b). Η λογική που ακολουθεί ο αλγόριθμος σε αυτή την περίπτωση είναι ακριβώς η ίδια με πρίν, με την διαφορά εδώ θα πρέπει να ελέγξει οκτώ κατευθύνσεις. Παρατηρούμε ότι το περίγραμμα που προκύπτει με τη μέθοδο της οκτώ συνδετικότητας είναι πιο λεπτό από αυτό της τέσσερα συνδετικότητας. Αυτό συμβαίνει διότι στην οκτώ συνδετικότητα μπορούμε να κινηθούμε διαγώνια ενώ στην τέσσερα συνδετικότητα κινούμαστε κλιμακωτά.



Σχήμα 54: Κώδικες αλυσίδων από διαφορετική συνδετικότητα

Ο κώδικας που προκύπτει κάθε φορά θα είναι διαφορετικός αν ξεκινήσουμε από διαφορετικό σημείο. Χρειαζόμαστε ένα σταθερό σημείο εκκίνησης. Θεωρούμε ότι αυτό θα είναι πάντα το σημείο που βρίσκεται αριστερά

με την μικρότερη συντεταγμένη ως προς y. Αυτό το σημείο ή θα βρίσκεται πρίν από την διαδικασία εντοπισμού της καμπύλης ή θα το βρίσκουμε αφού έχουμε εντοπίσει την καμπύλη ξεκινώντας από ένα τυχαίο σημείο. Αυτό μπορεί να επιτευχθεί θεωρώντας ότι τα στοιχεία του κώδικα ανήκουν σε μια διπλά διαδυνδεδεμένη λίστα. Μπορούμε να μετατοπίσουμε τα στοιχεία κυκλικά αντικαθιστώντας το στοιχείο που βρίσκεται στην αρχή της λίστας με αυτό που βρίσκεται στο τέλος της, με μετατόπιση όλων των στοιχείων μια θέση αριστερά. Μόλις στην αρχή της λίστας τοποθετηθεί το σημείο που βρίσκεται αριστερά με την μικρότερη συντεταγμένη ως πορς y τότε αυτό επιστρέφεται σαν η αρχή του κώδικα. Αυτή η διαδικασία φαίνεται στο Σχήμα 55. Το Σχήμα 55a έχει τον κώδικα που προκύπτει από το Σχήμα 54b . Στο Σχήμα 55b έχει γίνει αριστερή μετατόπιση στα στοιχεία του κώδικα. Αυτό είναι ισοδύναμο με τον κώδικα που θα προέκυπτε αν είχαμε ορίσει σημείο εκκίνησης το $P1$. Στο Σχήμα 55c έχει γίνει και δεύτερη μετατόπιση των στοιχείων του κώδικα. Τώρα είναι σαν να είχαμε ορίσει σημείο εκκίνησης το $P2$. Ύστερα από διαδοχικές μετακινήσεις των στοιχείων φέρνουμε στην αρχή του κώδικα το στοιχείο μηδέν (Σχήμα 55d). Αυτός ο κώδικας είναι αυτός που θα προέκυπτε αν είχαμε θεωρήση αρχή του αλγορίθμου το σημείο $P8$. Από άποψη ταχύτητας είναι πιο γρήγορο να κάνουμε διαδοχικές μετακινήσεις πάνω στον κώδικα παρά να ψάξουμε να βρούμε από την αρχή το $P8$.

code = {3,4,3,4,4,5,4,6,7,7,7,0,0,1,1,2} (a) Initial chain code	code = {4,3,4,4,4,5,4,6,7,7,7,0,0,1,1,2,3} (b) Result of one shift
code = {3,4,4,5,4,6,7,7,7,0,0,1,1,2,3,4} (c) Result of two shifts	code = {0,0,1,1,2,3,4,3,4,4,4,5,4,6,7,7,7} (d) Minimum integer chain code

Σχήμα 55: Περιστροφή του κώδικα

Είναι φανερό πως ο κώδικας αλυσίδων είναι διαφορετικός για διαφορετικά αντικείμενα. Βέβαια αυτό δεν μας εξασφαλίζει οτι θα είναι ο ίδιος για ίδια αντικείμενα. Είναι πολλοί οι παράγοντες που μπορούν να τον μεταβάλλουν. Ένας σημαντικός παράγοντας είναι ο θόρυβος. Αν ένα αντικείμενο φωτογραφήθηκε δυό φορές την ίδια στιγμή και η μια φωτογραφία έχει πιο πολύ θόρυβο από την άλλη, τότε είναι σχεδόν σίγουρο ότι το αντικείμενο θα έχει διαφορετικό κώδικα στην κάθε φωτογραφία. Επίσης ένας σημαντικός παράγοντας είναι η κλίμακα που έχει το αντικείμενο. Αν συμπιέσουμε το αντικείμενο τότε πάλι αλλάζει ο κωδικάς του. Ακόμα και μία μικρή αλλαγή στην γωνία της φωτογραφικής μηχανής είναι ικανή να μεταβάλλει τον κώδικα ίδιων αντικειμένων. Γενικά οτιδήποτε μπορεί να επηρεάσει την αρχική εικόνα του αντικειμένου μπορεί να επηρεάσει και τον κώδικα αυτού. Οι κώδικες αλυσίδων είναι μία τεχνική που παρόλη την αστάθεια που παρουσιάζει, χρησιοποιείται έντονα σε περιπτώσεις που το σύνολο των αντικειμένων που εξετάζουμε είναι καθορισμένο και σταθερό. [1]

4.2.3 Συντελεστές Fourier

Είναι φανερό ότι οι κώδικες αλυσίδων δεν μπορούν να χρησιμοποιηθούν πάντα για να περιγράψουν το σχήμα ενός αντικειμένου, παρά μόνο σε πολύ ιδιαίτερες περιπτώσεις. Για την περιγραφή του περιγράμματος χρησιμοποιούνται άλλες τεχνικές όπως είναι τα προσεγγιστικά πολυώνυμα που περιγράφουν καμπύλες. Σε αυτή την παράγραφο θα ασχοληθούμε με τους περιγραφείς Fourier με τους οποίους μπορούμε να περιγράψουμε το σχήμα ενός αντικειμένου. Η βασική ιδέα είναι να χαρακτηρίσουμε το περίγραμμα ενός σχήματος με ένα σύνολο αριθμών οι οποίοι θα το αντιπροσωπεύουν. Μπορούμε να χρησιμοποιήσουμε μικρό ή μεγάλο σύνολο αριθμών περιγραφής, ανάλογα με το σχήμα που θέλουμε να περιγράψουμε, οι οποίοι θα ονομάζονται συντελεστές Fourier . Για να μπορέσουμε να υλοποιήσουμε μια περιγραφή Fourier σε μια καμπύλη θα πρέπει να ακολουθήσουμε δύο βασικά βήματα :

(1) Πρώτα θα πρέπει να αποκτήσουμε μια αντιπροσώπευση της καμπύλης.

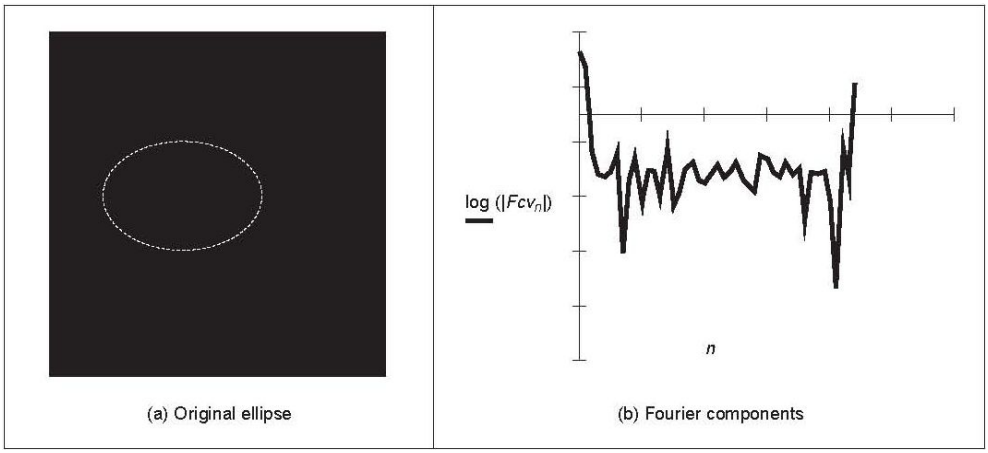
(2) Ύστερα την περιγράφουμε χρησιμοποιώτας θεωρία Fourier .

Οι περιγραφείς Fourier στους οποίους θα αναφερθούμε μπορούν να περιγράψουν περίπλοκα περιγράμματα που παρουσιάζουν πολλές γωνίες. Ανάλογα με το είδος της καμπύλης που ασχολούμαστε μπορούμε να αναπτύξουμε και διαφορετικούς περιγραφείς Fourier . Εκτός από το είδος της καμπύλης που θέλουμε να προσεγγίσουμε, ένας παράγοντας που επηρεάζει την ανάπτυξη και τις ιδιότητες της περιγραφής είναι η επιλογή της επέκτασης της μεθόδου Fourier . Εάν θεωρούμε ότι μια καμπύλη προκύπτει από μια περιοδική συνάρτηση, τότε μπορούμε

να επιλέξουμε τις ημιτονοειδής σειρές Fourier . Εδώ θα αναπτύξουμε μια αντιπροσώπευση που βασίζεται πάνω στις σειρές Fourier αφού οι καμπύλες που θέλουμε να προσεγγίσουμε είναι ένα σύνολο από διακεκριμένα εικονοκύτταρα.

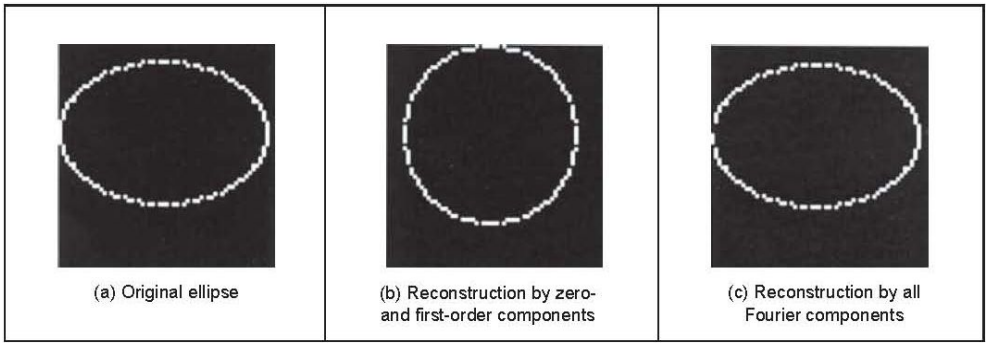
4.2.3.1 Η βάση των συντελεστών Fourier

Στην πιό βασική μορφή, οι συντεταγμένες των εικονοκυττάρων του περιγράμματος είναι ένα σύνολο (x, y) σημείων. Μια περιγραφή Fourier αυτών ουσιαστικά δίνει το σύνολο των συστατικών συχνοτήτων που ταιριάζουν στα σημεία του περιγράμματος. Το πρώτο στοιχείο των συνιστωσών Fourier που μπορούμε να πάρουμε για το περίγραμμα ενός σχήματος είναι η μεση τιμή των συντεταγμένων του x και y. Αυτό θα δώσει τις συντεταγμένες του κέντρου του περιγράμματος. Η δεύτερη συνιστώσα είναι η ακτίνα του κύκλου που ταιριάζει καλύτερα στα σημεία του περιγράμματος. Ένας κύκλος μπορεί να περιγραφεί με τις δυο συνιστώσες που μόλις αναφέραμε. Οι συνιστώσες υψηλότερης τάξης περιγράφουν λεπτομέρεια. Στο Σχήμα 56α έχουμε μια έλλειψη και στο Σχήμα 56b έχουμε τον μετασχηματισμό Fourier του περιγράμματος της σε λογαριθμική μορφή. Η περιγραφή Fourier της έλλειψης είναι οι συστατικές συχνότητες των συντεταγμένων του περιγραμματός της. Με ένα τέτοιο τρόπο μπορούμε να βρούμε ένα σύνολο από αριθμούς οι οποίοι θα μας περιγράφουν τη μορφή ενός σχήματος. Μια παρόμοια έλλειψη θα δώσει παρόμοιους αριθμούς και ένα τελείως διαφορετικό σχήμα θα δώσει τελείως διαφορετικούς αριθμούς.



Σχήμα 56: Μια έλλειψη και η Fourier περιγραφή της

Πολλές φορές χρειάζεται να ελέγξουμε το αποτέλεσμα. Ο καλύτερος τρόπος είναι να ανακατασκευάσουμε τη μορφή του σχήματος από τους αριθμούς περιγραφής του. Αν η ανακατασκευή ταιριάζει με το αρχικό σχήμα τότε η περιγραφή του απο τους αριθμούς θα είναι αποτελεσματική. Όπως έχουμε αναφέρει στην θεωρία σχετικά με τους μετασχηματισμούς Fourier , όταν ένα διάνυσμα μετασχηματιστεί κατά Fourier τότε μπορούμε να χρησιμοποιήσουμε τον αντίστροφο μετασχηματισμό Fourier και να ξαναπάρουμε το αρχικό διάνυσμα με κάποιες αποκλίσεις. Η έλλειψη μπορεί να ξανασχηματιστεί αν εφαρμόσουμε τον αντίστροφο μετασχηματισμό Fourier στις συστατικές της συχνότητες οι οποίες είναι οι αριθμοί περιγραφής της. Στο Σχήμα 57α έχουμε μία έλλειψη. Αν χρησιμοποιήσουμε μόνο το κέντρο και την ακτίνα της, δηλαδή μόνο την πρώτη και την δεύτερη συνιστώσα περιγραφής Fourier , θα κατασκευάσουμε ένα κύκλο (Σχήμα 57b). Αν χρησιμοποιήσουμε όλους τους αριθμούς περιγραφής της έλλειψης, τότε θα κατασκευάσουμε μια έλλειψη που θα ταιριάζει αρκετά στην αρχική (Σχήμα 57c). Αυτά είναι μόνο μια περίληψη της βάσης σχετικά με τους περιγραφείς Fourier . Στη συνέχεια θα ορίσουμε περιγραφείς που δίνουν την ίδια περιγραφή ανεξάρτητα από τη θέση, την περιστροφή και το μέγεθος του αντικειμένου. Εδώ εξετάσαμε την περιγραφή ενός αντικειμένου που επιτυγχάνεται με έναν τρόπο που επιτρέπει την ανακατασκευή του. Προκειμένου να αναπτυχθούν πρακτικά χρήσιμοι περιγραφείς θα πρέπει να εξετάσουμε περισσότερες βασικές ιδιότητες [1,11].



Σχήμα 57:

4.2.3.2 Επέκταση Fourier

Για να ορίσουμε μια επέκταση Fourier , μπορούμε να αρχίσουμε λαμβάνοντας υπόψη ότι μια συνεχής καμπύλη $c(t)$ μπορεί να εκφραστεί σαν ένα άθροισμα της μορφής :

$$c(t) = \sum_k c_k f_k(t) \qquad (4.1)$$

όπου c_k είναι οι συντελεστές της επέκτασης και $f_k(t)$ είναι οι βασικές συναρτήσεις, δηλαδή οι συναρτήσεις που χρησιμοποιούνται στις τεχνικές μοντελοποίησης καμπυλών ή επιφανειών που προσεγγίζουν ή παρεμβάλουν σημεία. Η δυσκολία βρίσκεται στο να βρεθούν οι συντελεστές c_k δεδομένου ενός συνόλου βασικών συναρτήσεων. Η σχέση 4.1 είναι πολύ γενική και μπορούν να χρησιμοποιηθούν διάφορες βασικές συναρτήσεις. Για παράδειγμα οι $f_k(t)$ μπορούν να επιλεχθούν έτσι ώστε το άθροισμα της σχέσης 4.1 να δίνει πολυώνυμα όπως είναι τα παρεμβολικά πολυώνυμα *Lagrange*, *Newton* και *Spline*. Η επέκταση Fourier αντιπροσωπεύει περιοδικές συναρτήσεις. Έτσι έχουμε :

$$c(t) = \sum_{k=-\infty}^{\infty} c_k \cdot e^{ik\omega t} \qquad (4.2)$$

Στην πιο πάνω σχέση το ω είναι η στοιχειώδη συχνότητα και είναι ίση με $T/2\pi$ όπου T είναι η περίοδος της συνάρτησης. Το κύριο χαρακτηριστικό της επέκτασης Fourier είναι ότι ορίζει μια ορθογώνια βάση. Αυτό σημαίνει οτι :

$$\int_0^T f_k(t) f_l(t) dt = 0 \qquad (4.3)$$

όταν $k \neq l$. Αυτή η ιδιότητα είναι σημαντική επειδή απλοποιεί τον υπολογισμό των συντελεστών. Έτσι στη σχέση 4.1 αντί να λύσουμε ως προς c_k πολλαπλασιάζουμε και τα δύο μέρη με $f_k(t)$ και στη συνέχεια ολοκληρώνουμε. Οι συντελεστές τώρα δίνονται από τη σχέση :

$$c_k = \frac{1}{T} \int_0^T c(t) e^{-ik\omega t} dt \qquad (4.4)$$

Εκτός από την εκθετική μορφή της σχέσης 4.2 η επέκταση Fourier μπορεί να εκφραστεί και σε τριγωνομετρική μορφή. Αυτή η μορφή δείχνει ότι η επέκταση Fourier αντιστοιχεί στο άθροισμα τριγωνομετρικών συναρτήσεων που αυξάνονται στη συχνότητα. Αυτό μπορεί να ληφθεί από την πιο κάτω σχέση :

$$c(t) = c_0 + \sum_{k=1}^{\infty} (c_k e^{ik\omega t} + c_{-k} e^{-ik\omega t}) \qquad (4.5)$$

Στη σχέση 4.5 οι τιμές των $e^{ik\omega t}$ και $e^{-ik\omega t}$ ορίζουν ζευγάρια μιγαδικών που είναι συζυγής. Κατά συνέπεια οι συντελεστές c_k και c_{-k} περιγράφουν ένα μιγαδικό αριθμό με τον συζυγή του. Ορίζουμε αυτούς τους αριθμούς :

$$c_k = c_{k,1} - ic_{k,2} \quad \text{και} \quad c_{-k} = c_{k,1} + ic_{k,2} \qquad (4.6)$$

Με αντικατάσταση στην σχέση 4.5 έχουμε :

$$c(t) = c_0 + 2\sum_{k=1}^{\infty}\left(c_{k,1}\left(\frac{e^{ik\omega t} + e^{-ik\omega t}}{2}\right) + ic_{k,2}\left(\frac{-e^{ik\omega t} + e^{-ik\omega t}}{2}\right)\right) \qquad (4.7)$$

Έχουμε ότι:

$$cos(k\omega t) = \frac{e^{ik\omega t} + e^{-ik\omega t}}{2} \quad , \quad -i\, sin(k\omega t) = \frac{-e^{ik\omega t} + e^{-ik\omega t}}{2} \qquad (4.8)$$

Άρα η 4.7 γίνεται :

$$c(t) = c_0 + 2\sum_{k=1}^{\infty}(c_{k,1}cos(k\omega t) + c_{k,2}sin(k\omega t)) \qquad (4.9)$$

Αν θέσουμε :

$$\alpha_k = 2c_{k,1} \ \text{ και } \ b_k = 2c_{k,2} \qquad (4.10)$$

τότε έχουμε την τριγωνομετρική μορφή που δίνεται από τη σχέση :

$$c(t) = \frac{\alpha_0}{2} + \sum_{k=1}^{\infty}(\alpha_k cos(k\omega t) + b_k sin(k\omega t)) \qquad (4.11)$$

Οι συντελεστές αυτής της επέκτασης, α_k και b_k, είναι γνωστεί σαν συντελεστές Fourier . Αυτοί οι αριθμοί μπορούμε να πούμε ότι περιγράφουν μια καμπύλη αφού θα έχουν διαφορετικές τιμές για διαφορετικές καμπύλες. Σύμφωνα με τις σχέσεις 4.6 και 4.10 έχουμε :

$$c_k = \frac{\alpha_k - ib_k}{2} \quad , \quad c_{-k} = \frac{\alpha_k + ib_k}{2} \qquad (4.12)$$

Οι συντελεστές της σχέσης 4.11 μπορούν να υπολογιστούν χρησιμοποιώντας την ορθογώνια ιδιότητα της σχέσης 4.3. Έτσι για τον υπολογισμό των τιμών των συντελεστών έχουμε :

$$\alpha_k = \frac{2}{T}\int_0^T c(t)cos(k\omega t)dt \quad , \quad b_k = \frac{2}{T}\int_0^T c(t)sin(k\omega t)dt \qquad (4.13) \qquad [1]$$

4.2.3.3 Ιδιότητα μετατόπισης

Αυτό που θα εξετάσουμε είναι πόσο μεταβάλλονται οι περιγραφείς Fourier οταν η αρχική καμπύλη μετακινηθεί. Εκτός από τις ισότητες 4.2 και 4.11 μια επέκταση Fourier μπορεί να γραφτεί σε μια άλλη τριγωνομετρική μορφή. Εάν θεωρήσουμε ότι:

$$|c_k| = \sqrt{\alpha_k^2 + b_k^2} \quad \text{και} \quad \phi_k = tan^{-1}(b_k/\alpha_k) \qquad (4.14)$$

τότε η επέκταση Fourier γίνεται :

$$c(t) = \frac{\alpha_0}{2} + \sum_{k=0}^{\infty}|c_k|cos(k\omega t + \phi_k) \qquad (4.15)$$

Μια σημαντική ιδιότητα της επέκτασης Fourier είναι ότι το μέτρο $|c_k|$ δεν αλλάζει οταν η καμπύλη $c(t)$ μετακινηθεί. Υποθέτουμε ότι η αρχική καμπύλη $c(t)$ μετακινείται κατά μια σταθερά τ. Τότε έχουμε την καμπύλη $c(t-\tau)$. Οι περιγραφείς Fourier της νέας καμπύλης είναι :

$$\alpha'_k = \frac{2}{T}\int_0^T c(t'+\tau)cos(k\omega t')dt \quad , \quad b'_k = \frac{2}{T}\int_0^T c(t'+\tau)sin(k\omega t')dt \qquad (4.16)$$

Χρησιμοποιούμε τη σχέση $t = t' + \tau$ και έχουμε :

$$\alpha'_k = \frac{2}{T}\int_0^T c(t)cos(k\omega t - k\omega\tau)dt \quad , \quad b'_k = \frac{2}{T}\int_0^T c(t)sin(k\omega t - k\omega\tau)dt \qquad (4.17)$$

Τελικά προκύπτουν οι σχέσεις:

$$\alpha'_k = \alpha_k\, cos(k\omega\tau) + b_k\, sin(k\omega\tau) \quad , \quad b'_k = b_k\, cos(k\omega\tau) - \alpha_k\, sin(k\omega\tau) \qquad (4.18)$$

Οπότε το μέτρο $|c'_k|$ δίνεται από τη σχέση :

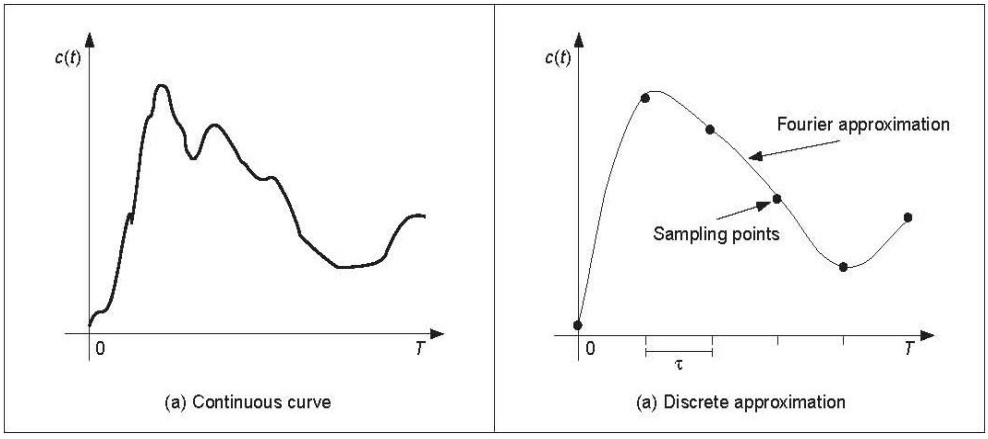
$$|c'_k| = \sqrt{(\alpha_k\, cos(k\omega\tau) + b_k\, sin(k\omega\tau))^2 + (b_k\, cos(k\omega\tau) - \alpha_k\, sin(k\omega\tau))^2} \qquad (4.19)$$

Τελικά προκύπει :

$$|c'_k| = \sqrt{\alpha_k^2 + b_k^2} = |c_k| \qquad (4.20) \qquad [1]$$

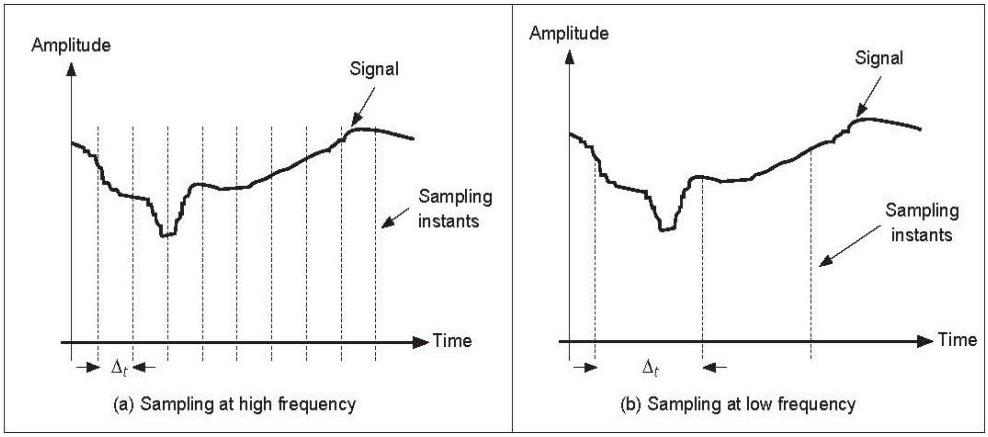
4.2.3.4 Διακεκριμένοι υπολογισμοί

Σε αυτό το σημείο θα ορίσουμε την αριθμητική διαδικασία που είναι απαραίτητη για να υπολογίσουμε τους περιγραφείς Fourier μιας καμπύλης. Το πρόβλημα που υπάρχει στις σχέσεις 4.11 και 4.13 είναι οτι έχουν οριστεί για συνεχείς καμπύλες. Όμως από τη στιγμή που η εικόνα αποτελείται από εικονοκύτταρα, τότε και η καμπύλη που θέλουμε να περιγράψουμε θα σχηματίζεται από συνεχόμενα διακεκριμένα εικονοκύτταρα. Αυτό έχει δύο σημαντικά αποτελέσματα. Πρώτον ο αριθμός των συχνοτήτων που παράγονται από την επέκταση Fourier είναι περιορισμένος. Δεύτερον επιτρέπει τον υπολογισμό των ολοκληρωμάτων των συντελεστών με αριθμητικούς τρόπους. Στο Σχήμα 58 υπάρχει ένα παράδειγμα όπου μια καμπύλη χωρίζεται σε διακεκριμένα σημεία. Στο Σχήμα 58α υπάρχει μια συνεχόμενη καμπύλη σε ένα διάστημα T. Το Σχήμα 58b δείχνει την προσέγγιση της καμπύλης με ένα σύνολο από διακεκριμένα σημεία. Εάν προσπαθήσουμε να πάρουμε την καμπύλη από τα επιλεγέντα σημεία, θα διαπιστώσουμε ότι η διαδικασία δειγματοληψίας μειώνει τη λεπτομέρεια του αρχικού σχήματος.



Σχήμα 58:

Ένα σημαντικό θέμα είναι να μπορέσουμε να καθορίσουμε το μέγεθος του δείγματος που θα αντιπροσωπεύει το σχήμα. Το μέγεθος του δείγματος που επιλέγουμε κάθε φορά, δηλαδή αν θα πάρουμε πολλά ή λίγα σημεία από το διάνυσμα των εικονοκυττάρων που σχηματίζουν την καμπύλη, έχει άμεση σχέση με τη μορφή της καμπύλης. Αν η καμπύλη έχει μορφή όπως στο Σχήμα 59 τότε θα πρέπει να επιλέξουμε σημεία που θα έχουν μικρή απόσταση μεταξύ τους, προκυμένου να έχουμε μια καλή αντιπροσώπευση της αρχικής καμπύλης.



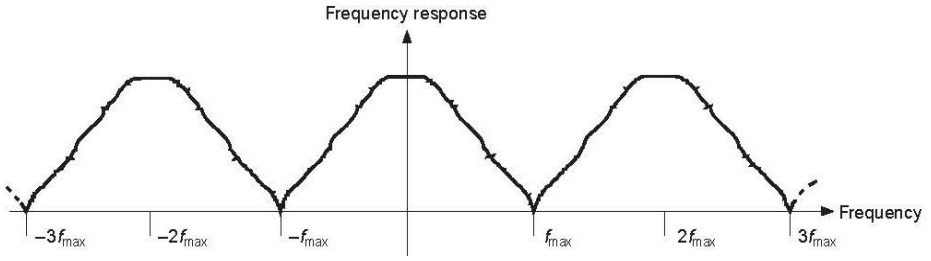
Σχήμα 59:

Αν το χρονικό διάστημα Δ_t είναι μικρό (Σχήμα 59α), τότε διατηρούμε πολλές λεπτομέρειες από την αρχική καμπύλη με αποτέλεσμα η καμπύλη που θα προσδιορίσουμε να είναι όμοια με την αρχική. Βέβαια οταν το δείγμα είναι μεγάλο τότε αυτό έχει συνέπεια την αύξηση των συναρτησιακών υπολογισμών κάτι που θέλουμε να αποφύγουμε. Αν το δείγμα είναι πολύ μικρό τότε σίγουρα οι συναρτησιακοί υπολογισμοί είναι πολύ λιγότεροι, όμως ταυτόχρονα έχουμε χάσει πολλές λεπτομέρειες απο την αρχική καμπύλη. Έτσι προκύπτει ο κίνδυνος, η καμπύλη που παράγεται από το δείγμα να έχει μικρή ομοιότητα με την αρχική, οπότε και οι περιγραφείς Fourier της νέας καμπύλης δεν θα αντιπροσωπεύουν την αρχική. Στο Σχήμα 59b παρατηρούμε ότι επίλέγοντας μεγάλο Δ_t έχουμε χάσει πολλές λεπτομέρειες του αρχικού σχήματος, όπως τη μικρή βύθιση λίγο πριν τη μέση του σχήματος, με αποτέλεσμα το δείγμα να μας παρέχει μια πολύ γενική εκτίμηση του αρχικού σχήματος. Αφού έχουμε επιλέξει το μέγεθος του δείγματος, πρέπει να εξετάσουμε πόσοι συντελεστές Fourier αντιστοιχούν σε αυτό. Την απάντηση την παίρνουμε από το κριτήριο δειγματοληψίας N σημείων :

Κριτήριο δειγματοληψίας N σημείων

Προκειμένου να αναδημιουργηθεί ένα σήμα από τα δείγματά του, η συχνότητα δειγματοληψίας πρέπει να είναι τουλάχιστον δύο φορές όσο η υψηλότερη συχνότητα του επιλεγέντος σήματος.

Έτσι αν υποθέσουμε ότι από ένα χρονικό σήμα, όπως στο Σχήμα 59, παίρνουμε δείγμα κάθε Δ_t και με f_{max} συμβολίσουμε την υψηλότερη συχνότητα του επιλεγέντος σήματος τότε το φάσμα της συχνότητας δειγματοληψίας που προκύπτει από τον μετασχηματισμό Fourier είναι το διάστημα $[-f_{max},f_{max}]$ (Σχήμα 60).



Σχήμα 60:

Αυτό μας δείχνει ότι αν έχουμε επίλέξει m ισαπέχοντα σημεία τότε μπορούν να προκύψουν το πολύ $m/2$ συντελεστές Fourier .Έτσι η επέκταση Fourier γίνεται :

$$c(t) = \frac{\alpha_0}{2} + \sum_{k=1}^{m/2}(\alpha_k cos(k\omega t) + b_k sin(k\omega t)) \qquad (4.21)$$

Στην πράξη δεν είναι χρήσιμο να υπολογίσουμε τον μέγιστο αριθμό των συντελεστών Fourier . Στις πιο πολλές περιπτώσεις το περίγραμμα του σχήματος που θέλουμε να περιγράψουμε αποτελείται από εκατοντάδες ή ακόμα και χιλιάδες εικονοκύτταρα, οπότε ο αριθμός των συντελεστών που θα προκύψουν θα είναι πολύ μεγάλος. Ένα τόσο μεγάλο διάνυσμα είναι πολύ δύσκολο να το χειριστούμε και ακόμα πιο δύσκολα μπορούμε να πάρουμε πληροφορίες για το σχήμα που προσεγγίζει. Είναι σχεδόν το ίδιο σαν χειριζόμαστε το διάνυσμα του αρχικού περιγράμματος. Εκτός από αυτό ένα διάνυσμα με το μισό μέγεθος του αρχικού θα είναι και ιδιαίτερα ευαίσθητο στο θόρυβο που περιέχει η εικόνα, αφού θα προσπαθεί να προσδιορίσει και τις μικρές λεπτομέρειες του περιγράμματος. Αυτό που μένει τώρα είναι να προσδιορίσουμε τα ολοκληρώματα της σχέσης 4.13 με αριθμητικό τρόπο, αφού η καμπύλη που εξετάζουμε σχηματίζεται από διακεκριμένα σημεία. Μπορούμε να προσεγγίσουμε τις τιμές αυτών των ολοκληρωμάτων με διάφορους τρόπους. Ενας κοινός τρόπος είναι να χρησιμοποιήσουμε το άθροισμα Riemann (Σχήμα 61b). Το άθροισμα Riemann είναι το συνολικό εμβαδόν m ορθογωνίων που βρίσκονται εν μέρει πάνω και εν μέρει κάτω από την καμπύλη. Το μεσαίο σημείο του κάθε ορθογωνίου αντιστοιχεί σε ένα σημείο του δείγματος. Θεωρώ ότι $t = i\tau$ όπου i είναι ενας ακέραιος από το διάστημα [1,m] και η c_i ορίζει την τιμή της συνάρτησης στο σημείο i, δηλαδή $c_i = c(i\tau)$. Κάθε διάστημα έχει μήκος $\tau = T/m$. Άρα έχουμε :

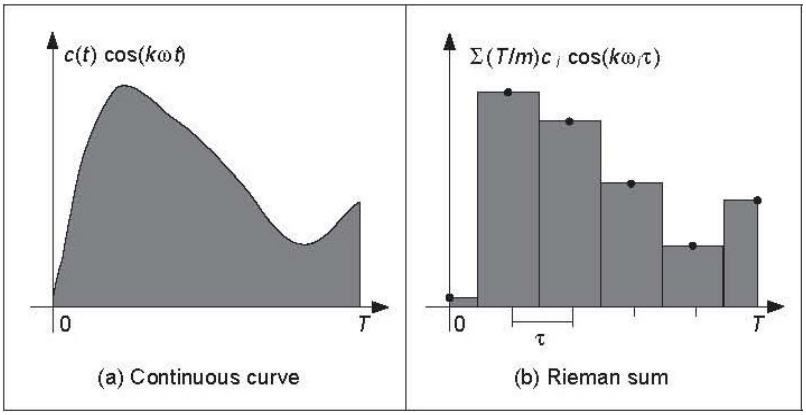
$$\int_0^T c(t)cos(k\omega t)dt \approx \sum_{i=1}^{m} \frac{T}{m} c_i cos(k\omega i\tau)$$

$$\int_0^T c(t)sin(k\omega t)dt \approx \sum_{i=1}^{m} \frac{T}{m} c_i sin(k\omega i\tau) \qquad (4.22)$$

Συνεπώς οι συντελεστές Fourier είναι :

$$\alpha_k = \frac{2}{m}\sum_{i=1}^{m} c_i cos(k\omega i\tau)\ ,\ b_k = \frac{2}{m}\sum_{i=1}^{m} c_i sin(k\omega i\tau) \qquad (4.23)$$

Ο αριθμητικός τρόπος υπολογισμού των ολοκληρωμάτων πάντα έχει σφάλμα. Η επιλογή πολλών σημείων θα έχει σαν αποτέλεσμα το άθροισμα Riemann θα προσεγγίσει με μεγαλύτερη ακρίβεια το ολοκλήρωμα. Ο αριθμός των σημείων που επιλέγουμε όπως έχουμε αναφέρει έχει σχέσει με την μορφή της καμπύλης. Όμως πολλές φορές αυτός ο αριθμός επιλέγεται με βάση το πρόβλημα που έχουμε να λύσουμε και την ακρίβεια που θέλουμε να πετύχουμε. [1]



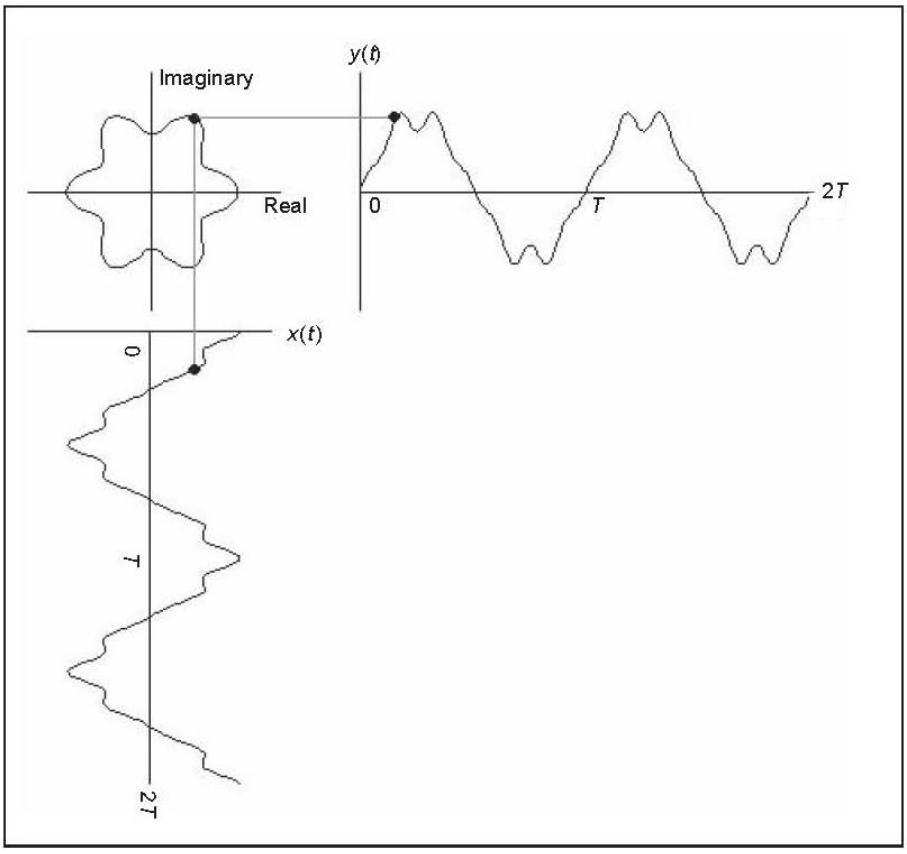
Σχήμα 61:

4.2.3.5 Ελλειπτικοί συντελεστές Fourier

Όπως έχουμε αναφέρει, η καμπύλη που θέλουμε να περιγράψουμε είναι ένα σύνολο από (x, y) συντεταγμένες στο επίπεδο. Έτσι μπορούμε να θεωρήσουμε ότι το κάθε εικονοκύτταρο της καμπύλης αντιπροσωπεύεται από ένα μιγαδικό αριθμό. Η συντεταγμένη x είναι το πραγματικό μέρος και η συντεταγμένη y είναι το φανταστικό. Έτσι η καμπύλη μπορεί να γραφεί :

$$c(t) = x(t) + iy(t) \qquad (4.24)$$

όπου το t ανήκει στο διάστημα $[0, 2T]$. Στο Σχήμα 62 βλέπουμε πως ένα σχήμα μπορεί να αναλυθεί σε δύο καμπύλες οι οποίες εκτείνονται στο διάστημα $[0, 2T]$. Το T είναι η περίοδος και είναι ίση με 2π. Οι δύο καμπύλες εκτείνονται από 0 μέχρι $2T$ αφού η ανάλυση Fourier γίνεται από 0 μέχρι T προς τη θετική και προς την αρνητική κατεύθυνση.



Σχήμα 62:

Αυτό που πρέπει να κάνουμε στη συνέχεια είναι να σχηματίσουμε τους συντελεστές Fourier για την έκφραση της καμπύλης της σχέσης 4.24. Έχουμε ότι :

$$c_k = c_{xk} + ic_{yk} \qquad (4.25)$$

Απο τη σχέση 4.4 έχουμε :

$$c_{xk} = \frac{1}{T}\int_0^T x(t)e^{-ik\omega t}dt$$

$$c_{yk} = \frac{1}{T}\int_0^T y(t)e^{-ik\omega t}dt \qquad (4.26)$$

Απο τη σχέση 4.12 έχουμε :

$$c_{xk} = \frac{\alpha_{xk} - ib_{xk}}{2} \ , \ c_{x-k} = \frac{\alpha_{xk} + ib_{xk}}{2}$$
$$c_{yk} = \frac{\alpha_{yk} - ib_{yk}}{2} \ , \ c_{y-k} = \frac{\alpha_{yk} + ib_{yk}}{2} \tag{4.27}$$

Απο τη σχέση 4.13 έχουμε :

$$\alpha_{xk} = \frac{2}{T}\int_0^T x(t)cos(k\omega t)dt \ , \ b_{xk} = \frac{2}{T}\int_0^T x(t)sin(k\omega t)dt$$
$$\alpha_{yk} = \frac{2}{T}\int_0^T y(t)cos(k\omega t)dt \ , \ b_{yk} = \frac{2}{T}\int_0^T y(t)sin(k\omega t)dt \tag{4.28}$$

Απο τη σχέση 4.23 τα πιο πάνω ολοκληρώματα μπορούν να υπολογιστούν από τις σχέσεις :

$$\alpha_{xk} = \frac{2}{m}\sum_{j=1}^{m} x_j cos(k\omega j\tau) \ , \ b_{xk} = \frac{2}{m}\sum_{j=1}^{m} x_j sin(k\omega j\tau)$$
$$\alpha_{yk} = \frac{2}{m}\sum_{j=1}^{m} y_j cos(k\omega j\tau) \ , \ b_{yk} = \frac{2}{m}\sum_{j=1}^{m} y_j sin(k\omega j\tau) \tag{4.29}$$

όπου τα x_j και y_j είναι οι τιμές των $x(t)$ και $y(t)$ στο σημείο j. από τις σχέσεις 4.25 και 4.27 μπορούμε να εκφράσουμε το c_k σαν το άθροισμα από ένα ζευγάρι μιγαδικών αριθμών.

$$c_k = A_k - iB_k \ , \ c_{-k} = A_k + iB_k \tag{4.30}$$

όπου

$$A_k = \frac{\alpha_{xk} + i\alpha_{yk}}{2} \ , \ B_k = \frac{b_{xk} + ib_{yk}}{2} \tag{4.31}$$

Με τη χρήση της σχέσης 4.5 η καμπύλη $c(t)$ γίνεται :

$$c(t) = c_0 + \sum_{k=1}^{\infty}(A_k - iB_k)e^{ik\omega t} + \sum_{k=-\infty}^{-1}(A_k + iB_k)e^{ik\omega t} \tag{4.32}$$

Απο τη σχέση 4.11 η καμπύλη μπορεί να πάρει την πιο κάτω τριγωνομετρική μορφή :

$$c(t) = \left(\frac{\alpha_{x0}}{2} + \sum_{k=1}^{\infty}(\alpha_{xk}cos(k\omega t) + b_{xk}sin(k\omega t))\right) + i\left(\frac{\alpha_{y0}}{2} + \sum_{k=1}^{\infty}(\alpha_{yk}cos(k\omega t) + b_{yk}sin(k\omega t))\right) \tag{4.33}$$

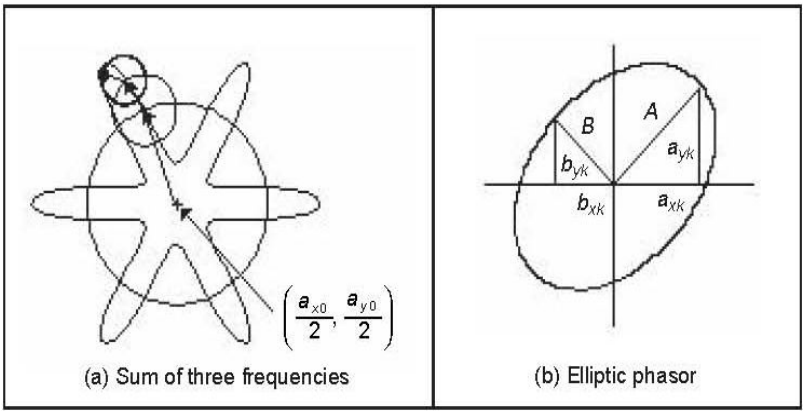
Η σχέση 4.33 μπορεί να εκφραστεί σε μορφή πίκακα :

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} \\ \alpha_{y0} \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} \alpha_{xk} & b_{xk} \\ \alpha_{yk} & b_{yk} \end{bmatrix} \begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \qquad (4.34)$$

Τα σημεία $\alpha_{x0}/2$ και $\alpha_{y0}/2$ αποτελούν το κέντρο της καμπύλης $c(t)$. Το διάνυσμα

$$\begin{bmatrix} \alpha_{xk} & b_{xk} \\ \alpha_{yk} & b_{yk} \end{bmatrix} \begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix}$$

ορίζει το κέντρο της κ-οστής έλλειψης στην κατεύθυνση t αν θεωρήσουμε ότι κινούμαστε κυκλικά γύρο απο το περίγραμμα της καμπύλης $c(t)$. Το μέτρο των μιγαδικών A_k και B_k ορίζουν το μήκος των αξόνων x και y της κ-οστής έλλειψης αντίστοιχα (Σχήμα 63b). Στο Σχήμα 63α υποθέτουμε ότι το k είναι ίσο με 3. Αυτό σημαίνει ότι θα σχηματιστούν τέσσερις ελλείψεις. Αρχικά βρίσκουμε το κέντρο του περιγράμματος που είναι το $(\frac{\alpha_{x0}}{2}, \frac{\alpha_{y0}}{2})$, το οποίο θα είναι και το κέντρο της αρχικής έλλειψης. Πάνω στο περίγραμμα της αρχικής έλλειψης θα βρίσκεται το κέντρο της δεύτερης έλλειψης. Γενικά πάνω στο περίγραμμα της (κ-1)-οστής έλλειψης θα βρίσκεται το κέντρο της κ-οστής έλλειψης. Τελικα στο Σχήμα 63α το κέντρο της τελευταίας έλλειψης, δηλαδή της τέταρτης, το δεχόμαστε ότι ανήκει πάνω στην καμπύλη $c(t)$.



Σχήμα 63:

4.2.3.6 Ιδιότητες των ελλειπτικών συντελεστών Fourier

Αυτό που θα εξετάσουμε σε αυτή την ενότητα είναι η μεταβλητότητα των συντελεστών Fourier . Οταν η αρχική καμπύλη μεταβληθεί θέλουμε να δούμε πόσο μεταβάλονται οι συντελεστές Fourier . Πιο συγκεκριμένα θα εξετάσουμε τις περιπτώσεις που η αρχική καμπύλη μετατοπίζεται, αλλάζει στην κλίμακα και περιστρέφεται.

1. Μετατόπιση

Έχουμε δείξει ότι η καμπύλη $c(t) = x(t) + iy(t)$ μπορεί να εκφραστεί από τη σχέση 4.34. Υποθέτουμε οτι η καμπύλη $c(t)$ μετατοπίζεται κατά t_x και t_y ως προς τον άξονα x και y αντίστοιχα. Έτσι προκύπτει η νέα καμπύλη $c'(t) = x'(t) + iy'(t)$. Σύμφωνα με τη σχέση 4.34 η $c'(t)$ μπορεί να εκφραστεί με την πιο κάτω σχέση :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha'_{x0} \\ \alpha'_{y0} \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} \alpha'_{xk} & b'_{xk} \\ \alpha'_{yk} & b'_{yk} \end{bmatrix} \begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \qquad (4.35)$$

Αφού η $c'(t)$ προκύπτει από τη μετατόπιση της $c(t)$ κατά t_x και t_y ως προς τον άξονα x και y αντίστοιχα τότε έχουμε :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} \\ \alpha_{y0} \end{bmatrix} + \sum_{k=1}^{\infty}\begin{bmatrix} \alpha_{xk} & b_{xk} \\ \alpha_{yk} & b_{yk} \end{bmatrix}\begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix} \qquad (4.36)$$

Η πιο πάνω σχέση γίνεται :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} + 2t_x \\ \alpha_{y0} + 2t_y \end{bmatrix} + \sum_{k=1}^{\infty}\begin{bmatrix} \alpha_{xk} & b_{xk} \\ \alpha_{yk} & b_{yk} \end{bmatrix}\begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \qquad (4.37)$$

Συγκρίνοντας τους συντελεστές Fourier της αρχικής και της μετατοπισμένης καμπύλης έχουμε την πιο κάτω σχέση :

$$\alpha'_{xk} = \alpha_{xk}\ ,\ b'_{xk} = b_{xk}$$
$$\alpha'_{yk} = \alpha_{yk}\ ,\ b'_{yk} = b_{yk}$$

Όταν $k \neq 0$.Για $k = 0$ έχουμε : (4.38)

$$\alpha'_{x0} = \alpha_{x0} + 2t_x$$
$$\alpha'_{y0} = \alpha_{y0} + 2t_y$$

2. Αλλαγή Κλίμακας

Η αλλαγή της κλίμακας μιας καμπύλης $c(t)$ μπορεί να εκφραστεί σαν τη διαστολή ή τη συστολή της καμπύλης από το κέντρο της. Υποθέτουμε ότι η καμπύλη $c(t)$ αλλάζει στην κλίμακα και προκύπτει η καμπύλη $c'(t)$. Αν s είναι ο παράγωντας κλίμακας τότε η νέα καμπύλη $c'(t)$ από τη σχέση 4.34 μπορεί να γραφεί :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} \\ \alpha_{y0} \end{bmatrix} + s\sum_{k=1}^{\infty}\begin{bmatrix} \alpha_{xk} & b_{xk} \\ \alpha_{yk} & b_{yk} \end{bmatrix}\begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \qquad (4.39)$$

Το s μπορεί να μπεί μέσα στο άθροισμα αφού δεν εξαρτάται από το k. Άρα η πιο πάνω σχέση γίνεται :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} \\ \alpha_{y0} \end{bmatrix} + \sum_{k=1}^{\infty}\begin{bmatrix} s\alpha_{xk} & sb_{xk} \\ s\alpha_{yk} & sb_{yk} \end{bmatrix}\begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \qquad (4.40)$$

Συγκρίνοντας τις σχέσεις 4.34 και 4.40 μπορούμε να πάρουμε τη σχέση που έχουν οι συντελεστές Fourier στίς δύο καμπύλες :

$$\alpha'_{xk} = s\alpha_{xk}\ ,\ b'_{xk} = sb_{xk}$$
$$\alpha'_{yk} = s\alpha_{yk}\ ,\ b'_{yk} = sb_{yk}$$

Όταν $k \neq 0$.Για $k = 0$ έχουμε : (4.41)

$$\alpha'_{x0} = \alpha_{x0}$$
$$\alpha'_{y0} = \alpha_{y0}$$

3. Περιστροφή

Σε αυτή τη περίπτωση υποθέτουμε ότι η αρχική καμπύλη περιστρέφεται γύρο από ένα σταθερό άξονα L. Έστω η γωνία περιστροφής ως προς τον άξονα L είναι ρ. Τότε η νέα καμπύλη $c'(t)$ από τη σχέση 4.34 γίνεται :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} \\ \alpha_{y0} \end{bmatrix} + \begin{bmatrix} cos(\rho) & sin(\rho) \\ -sin(\rho) & cos(\rho) \end{bmatrix} \sum_{k=1}^{\infty} \begin{bmatrix} \alpha_{xk} & b_{xk} \\ \alpha_{yk} & b_{yk} \end{bmatrix} \begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \tag{4.42}$$

Με την εκτέλεση των πράξεων προκύπτει η επόμενη σχέση :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} \\ \alpha_{y0} \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} \alpha_{xk}cos(\rho) + \alpha_{yk}sin(\rho) & b_{xk}cos(\rho) + b_{yk}sin(\rho) \\ -\alpha_{xk}sin(\rho) + \alpha_{yk}cos(\rho) & -b_{xk}sin(\rho) + b_{yk}cos(\rho) \end{bmatrix} \begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \tag{4.42}$$

Συγκρίνοντας τώρα τις σχέσεις 4.34 και 4.42 μπορούμε να πάρουμε τη σχέση που έχουν οι συντελεστές Fourier της αρχικής καμπύλης και αυτής που πρέκυψε από την περιστροφή:

$$\alpha'_{xk} = \alpha_{xk}cos(\rho) + \alpha_{yk}sin(\rho)\ ,\ b'_{xk} = b_{xk}cos(\rho) + b_{yk}sin(\rho)$$
$$\alpha'_{yk} = -\alpha_{xk}sin(\rho) + \alpha_{yk}cos(\rho)\ ,\ b'_{yk} = -b_{xk}sin(\rho) + b_{yk}cos(\rho)$$

Όταν $k \neq 0$.Για $k = 0$ έχουμε : (4.43)

$$\alpha'_{x0} = \alpha_{x0}$$
$$\alpha'_{y0} = \alpha_{y0}$$

4. Μετατόπιση - Αλλαγή Κλίμακας - Περιστροφή

Αυτό που μένει είναι να δώσουμε τη σχέση των συντελεστών Fourier της αρχικής καμπύλης και της καμπύλης που προκύπτει αν συμβούν και οι τρείς πιο πάνω μετασχηματισμοί. Όπως και πρίν η καμπύλη $c(t)$ μετατοπίζεται κατά t_x και t_y ως προς τον άξονα x και y αντίστοιχα, αλλάζει στην κλίμακα κατά s και περιστρέφεται κατά γωνία ρ γύρο από ένα σταθερό άξονα L. Η νέα καμπύλη $c'(t)$ δίνεται από τη σχέση :

$$\begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \alpha_{x0} + 2t_x \\ \alpha_{y0} + 2t_y \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} s(\alpha_{xk}cos(\rho) + \alpha_{yk}sin(\rho)) & s(b_{xk}cos(\rho) + b_{yk}sin(\rho)) \\ s(-\alpha_{xk}sin(\rho) + \alpha_{yk}cos(\rho)) & s(-b_{xk}sin(\rho) + b_{yk}cos(\rho)) \end{bmatrix} \begin{bmatrix} cos(k\omega t) \\ sin(k\omega t) \end{bmatrix} \tag{4.43}$$

Η σχέση των συντελεστών Fourier της $c(t)$ και της $c'(t)$ είναι :

$$\alpha'_{xk} = s(\alpha_{xk}cos(\rho) + \alpha_{yk}sin(\rho))\ ,\ b'_{xk} = s(b_{xk}cos(\rho) + b_{yk}sin(\rho))$$
$$\alpha'_{yk} = s(-\alpha_{xk}sin(\rho) + \alpha_{yk}cos(\rho)\ ,\ b'_{yk} = s(-b_{xk}sin(\rho) + b_{yk}cos(\rho))$$

Όταν $k \neq 0$.Για $k = 0$ έχουμε : (4.44)

$$\alpha'_{x0} = \alpha_{x0} + 2t_x$$
$$\alpha'_{y0} = \alpha_{y0} + 2t_y$$

Απο την σχέση 4.44 παρατηρούμε ότι οι συντελεστές Fourier δεν παραμένουν ίδιοι οταν η αρχική καμπύλη μετασχηματιστεί. Ο *Granlund* (1972) όρισε μια άλλη έκφραση των συντελεστών Fourier στην οποία οι συντελεστές δεν μεταβάλονται όταν η αρχική καμπύλη μετασχηματιστεί. Αυτή η έκφραση δίνεται από τον τύπο :

$$FourierDescriptors = \frac{|A_k|}{|A_1|} + \frac{|B_k|}{|B_1|} = \frac{\sqrt{\alpha_{xk}^2 + \alpha_{yk}^2}}{\sqrt{\alpha_{x1}^2 + \alpha_{y1}^2}} + \frac{\sqrt{b_{xk}^2 + b_{yk}^2}}{\sqrt{b_{x1}^2 + b_{y1}^2}} \qquad (4.45)$$

όπου τα A_k και B_k δίνονται από τη σχέση 4.31. Στην νέα έκφραση οι συντελεστές για $k = 0$ δεν χρησιμοποιούνται. Για την μετασχηματισμένη καμπύλη έχουμε :

$$FourierDescriptors' = \frac{|A'_k|}{|A'_1|} + \frac{|B'_k|}{|B'_1|} = \frac{\sqrt{\alpha_{xk}'^2 + \alpha_{yk}'^2}}{\sqrt{\alpha_{x1}'^2 + \alpha_{y1}'^2}} + \frac{\sqrt{b_{xk}'^2 + b_{yk}'^2}}{\sqrt{b_{x1}'^2 + b_{y1}'^2}} \qquad (4.46)$$

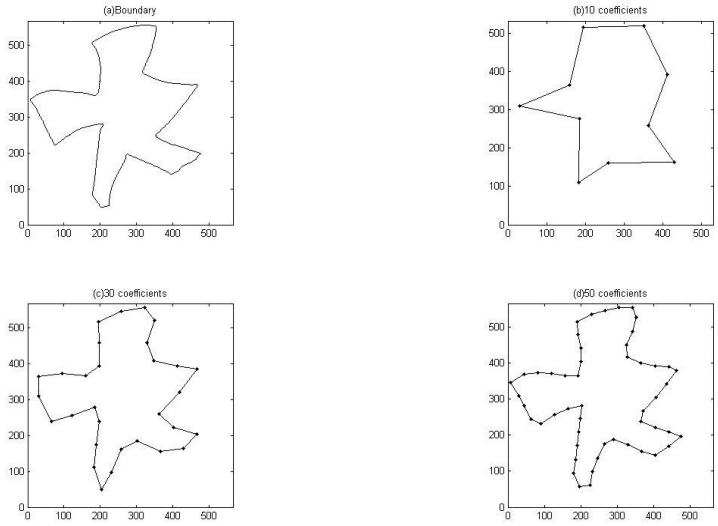
Όμως από τη σχέση 4.44 έχουμε την ισότητα :

$$\frac{\sqrt{\alpha_{xk}'^2 + \alpha_{yk}'^2}}{\sqrt{\alpha_{x1}'^2 + \alpha_{y1}'^2}} + \frac{\sqrt{b_{xk}'^2 + b_{yk}'^2}}{\sqrt{b_{x1}'^2 + b_{y1}'^2}} = \frac{\sqrt{\alpha_{xk}^2 + \alpha_{yk}^2}}{\sqrt{\alpha_{x1}^2 + \alpha_{y1}^2}} + \frac{\sqrt{b_{xk}^2 + b_{yk}^2}}{\sqrt{b_{x1}^2 + b_{y1}^2}} \qquad (4.47)$$

Τελικά ισχύει :

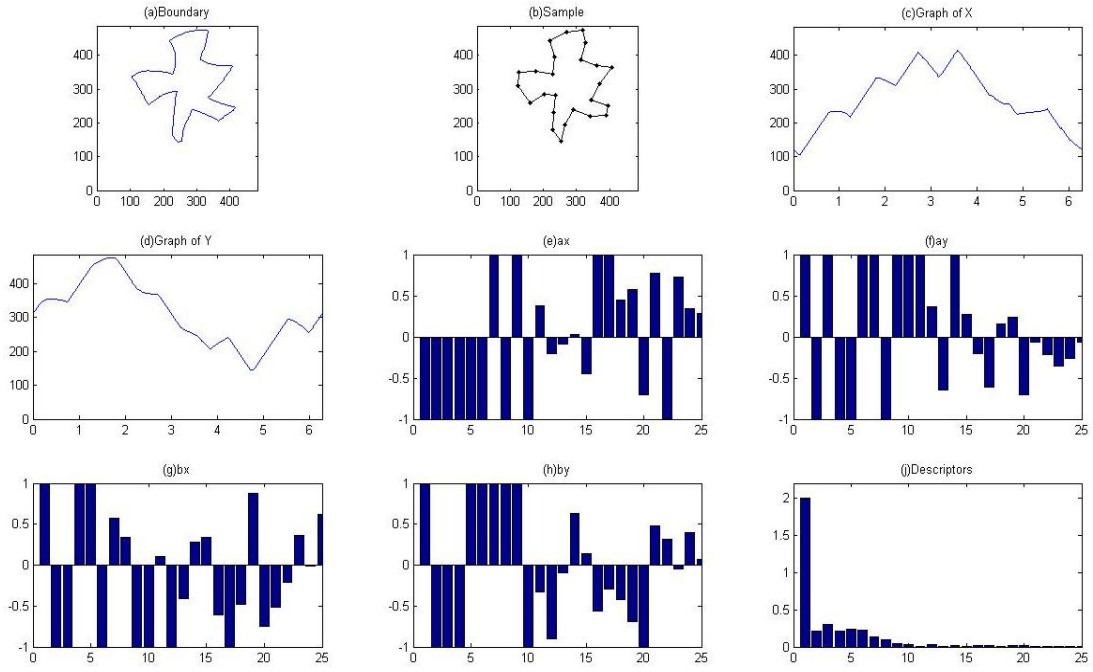
$$FourierDescriptors' = FourierDescriptors \qquad (4.48) \qquad [1, 4, 15]$$

Αφού δείξαμε τις ιδιοτήτες των συντελεστών Fourier , είναι ιδιαίτερα σημαντικό να ασχοληθούμε με τον αριθμό των συντελεστών που πρέπει να επιλέξουμε για να περιγράψουμε ένα σχήμα. Δεν υπάρχει κάποιος κανόνας που θα μας λέει πόσους συντελεστές πρέπει να επιλέξουμε για την περιγραφή ενός σχήματος. Η επιλογή του αριθμού των συντελεστών έχει σχέση με τη μορφή του σχήματος, με την ακρίβεια που θέλουμε να έχουμε και προκύπτει ύστερα από πειραματισμό. Όπως έχουμε αναφέρει, ο αριθμός των συντελεστών μπορεί να είναι το πολύ ο μισός αριθμός των σημείων που αποτελούν το περίγραμμα του σχήματος. Δηλαδή αν το περίγραμμα αποτελείται από N εικονοκύτταρα τότε μπορούμε να έχουμε το πολύ $N/2$ συντελεστές. Ένας τέτοιος αριθμός είναι συνήθως πολύ μεγάλος, αν υποθέσουμε ότι το περίγραμμα θα αποτελείται από μερικές εκατοντάδες εικονοκύτταρα, και είναι δύσκολο να τον χειριστούμε. Γενικά προσπαθούμε να επιλέξουμε έναν αριθμό M σημείων που θα προσδιορίζουν αρκετά καλά το αρχικό σχήμα και θέλουμε το M να μην είναι πολύ μεγάλο. Στό Σχήμα 64 παρουσιάζεται η μορφή που έχει ένα περίγραμμα όταν αυτό σχηματιστεί με λιγότερα ισαπέχοντα σημεία. Στο Σχήμα 64α το αρχικό περίγραμμα ενός τυχαίου σχήματος αποτελείται από 849 εικονοκύτταρα. Στο Σχήμα 64b έχουμε επιλέξει 10 ισαπέχοντα σημεία από το αρχικό περίγραμμα. Το περίγραμμα που προκύπτει δεν έχει καμία ομοιότητα με το περίγραμμα στο Σχήμα 64a , οπότε 10 συντελεστές δεν είναι αρκετοί για να περιγράψουν το αρχικό περίγραμμα. Στο Σχήμα 64c επιλέγουμε 30 σημεία από το αρχικό περίγραμμα. Αυτό που προκύπτει έχει αρκετές ομοιότητες με το αρχικό περίγραμμα και 30 συντελεστές είναι πιθανό αρκετοί για να το περιγράψουν. Στο Σχήμα 64d έχουμε επιλέξει 50 σημεία από το αρχικό σχήμα. Αυτή τη φορά το περίγραμμα που προκύπτει έχει διατηρήσει πολλές λεπτομέρειες του αρχικού και μπορεί να το προσεγγίσει με μεγαλύτερη ακρίβεια .Όπως παρατηρούμε καθώς αυξάνεται ο αριθμός των συντελεστών, έχουμε μια καλύτερη αντιπροσώπευση του αρχικού σχήματος. Επίσης βλέπουμε ότι ο αριθμός των συντελεστών είναι συνήθως πολύ μικρότερος από τον αριθμό των εικονοκυττάρων που αποτελούν το περίγραμμα του σχήματος.



Σχήμα 64:

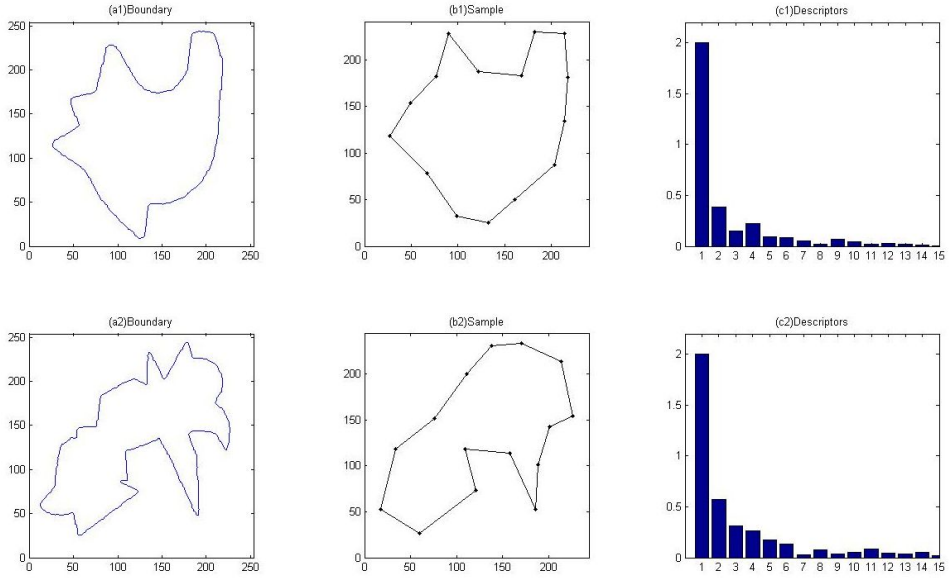
Στο Σχήμα 65 παρουσιάζονται οι συντελεστές Fourier του Σχήματος 64a , όταν έχουμε επιλέξει 25 ισαπέχοντα σημεία.



Σχήμα 65:

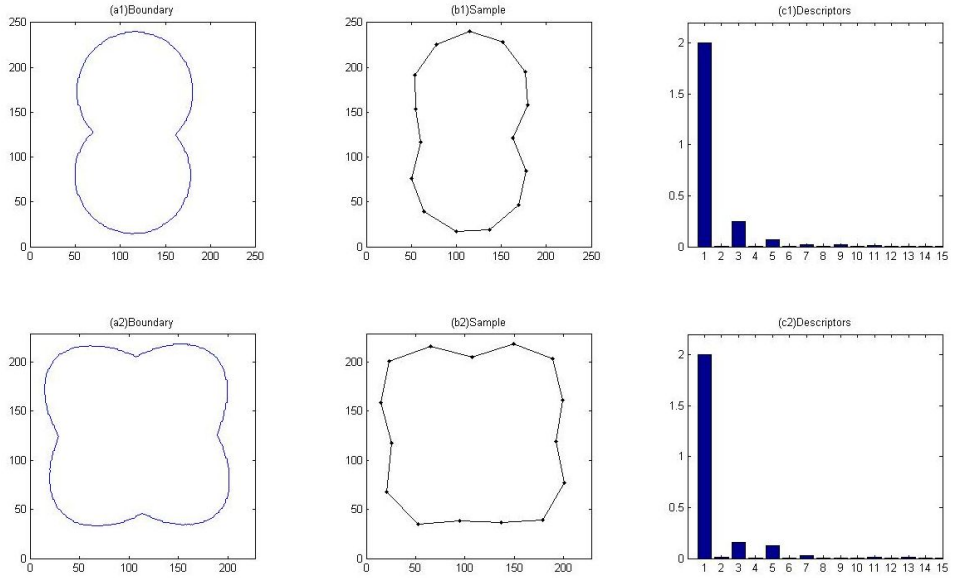
Στο Σχήμα 65a και στο Σχήμα 65b έχουμε το αρχικό περίγραμμα και το περίγραμμα που σχηματίζεται από 25 ισαπέχοντα σημεία αντίστοιχα. Στα Σχήματα 65c και 65d έχουμε την ανάλυση του αρχικού περιγράμματος ως προς την x και y συντεταγμένη. Τα Σχήματα 65e , 65f , 65g και 65h παρουσιάζουν τους συντελεστές Fourier της σχέσης 4.29 αντίστοιχα. Τέλος στο Σχήμα 65j φαίνονται οι αναλλοίωτοι συντελεστές Fourier της σχέσης 4.45.

Στο Σχήμα 66 μπορούμε να παρατηρήσουμε ότι οι συντελεστές Fourier είναι διαφορετικοί για σχήματα που έχουν αρκετές διαφορές μεταξύ τους. Στα Σχήματα 66 a1 και 66 a2 έχουμε το περίγραμμα των αρχικών σχημάτων, τα οποία θέλουμε να τα προσεγγίσουμε με 15 συντελεστές Fourier . Στα Σχήματα 66 b1 και 66 b2 έχουμε πάρει 15 ισαπέχοντα σημεία. Στα Σχήματα 66 c1 και 66 c2 παρουσιάζονται γραφικά οι συντελεστές Fourier των δύο σχημάτων. Έυκολα μπορούμε να παρατηρήσουμε ότι οι συντελεστές των δύο σχημάτων έχουν αρκετές διαφορές μεταξύ τους. Οπότε μπορούμε να ξεχωρίσουμε τα δύο σχήματα μόνο από τους συντελεστές τους.



Σχήμα 66:

Στο Σχήμα 67 βλέπουμε ότι οι συντελεστές Fourier έχουν μικρές διαφορές για όμοια σχήματα. Το περίγραμμα των Σχημάτων 67 a1 και 67 a2 έχουν αρκετές ομοιότητες μεταξύ τους. Παρατηρούμε οτι οι συντελεστές Fourier που προκύπτουν είναι σχεδόν ίδιοι. Αυτό σημαίνει ότι σε περιπτώσεις που είναι όμοια τα σχήματα που θέλουμε να προσδιορίσουμε, οι συντελεστές Fourier μπορεί να μην δίνουν αρκετές πληροφορίες έτσι ώστε να μπορέσουμε να τα ξεχωρίσουμε.



Σχήμα 67:

5 Τεχνητά Νευρωνικά Δίκτυα

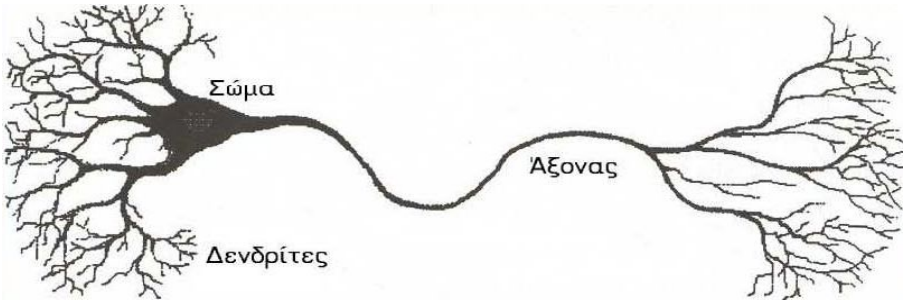
5.1 Γενική επισκόπηση

Με τον όρο Τεχνητό Νευρωνικό Δίκτυο εννούμε ένα μαθηματικό μοντέλο που προσπαθεί να μιμηθεί τη δομή και τη λειτουργεία ενός Βιολογικού Νευρωνικού δικτύου. Αποτελείται από ένα σύνολο διασυνδεδεμένων τεχνητών νευρώνων που είναι οργανωμένοι σε στρώματα και εκτελεί αριθμητικούς υπολογισμούς χρησιμοποιώντας δομή μαζικού παραλληλισμού. Το Τεχνητό Νευρωνικό Δίκτυο είναι ένας προσαρμοστικός μηχανισμός που αλλάζει τη δομή του ανάλογα με τις πληροφορίες που έχει να επεξεργαστεί. Γενικά προσομοιάζει στον ανθρώπινο εγκέφαλο σε δύο σημεία: (1) η γνώση του Τεχνητού Νευρωνικού Δικτύου αποκτάται μέσω μιας διαδικασίας μάθησης και (2) οι σύνδεσμοι μεταξύ των νευρώνων που ονομάζονται βάρη χρησιμοποιούνται για να αποθηκευτεί αυτή η γνώση.

5.2 Αρχιτεκτονική Νευρωνικών Δικτύων

5.2.1 Βιολογικά Νευρωνικά Δίκτυα

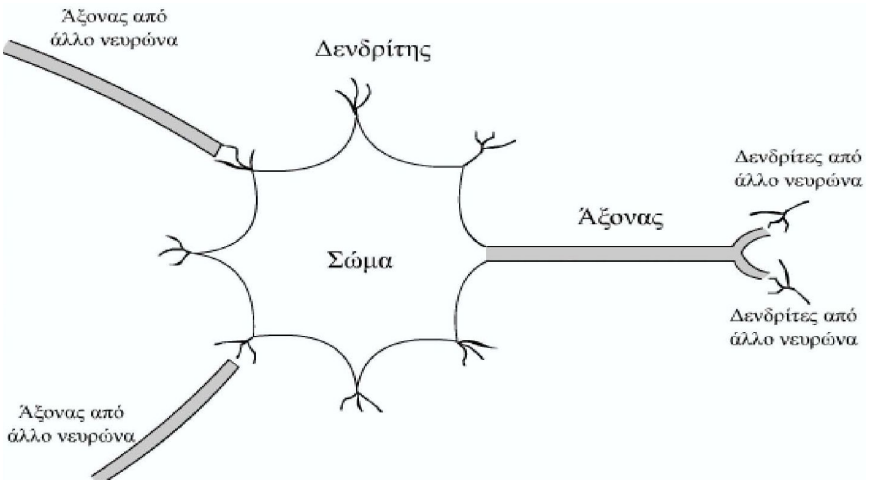
Ένα Βιολογικό Νευρωνικό Δίκτυο αποτελείται από πραγματικούς βιολογικούς νευρώνες οι οποίοι επικοινωνούν μεταξύ τους και βρίσκονται στον ανθρώπινο εγκέφαλο. Κάθε νευρώνας συνδέεται με πολλούς άλλους νευρώνες και ο συνολικός αριθμός νευρώνων και συνδέσεων σε ένα δίκτυο μπορεί να είναι εκτενής. Η επικοινωνία μεταξύ των νευρώνων περιλαμβάνει συχνά μια ηλεκτροχημική διαδικασία. Ο νευρώνας είναι το μικρότερο τμήμα του εγκεφάλου που είναι ικανό να επεξεργαστεί πληροφορίες. Έχει υπολογιστεί ότι η επεξεργασία στους βιολογικούς νευρώνες γίνεται 5 με 6 τάξεις μεγέθους πιο αργά από ότι στις σύχρονες ψηφιακές λογικές πύλες. Ο χρόνος για τις ψηφιακές λογικές πύλες μετριέται σε δισεκατομμυριοστά του δευτερολέπτου, ενώ στούς βιολογικούς νευρώνες σε χιλιοστά του δευτερολέπτου. Όμως ο εγκέφαλος αντισταθμίζει τη σχετικά αργή ταχύτητα λειτουργίας των νευρώνων με τον τεράστιο αριθμό τους και με τον τεράστιο αριθμό των μεταξύ τους συνδέσεων. Υπολογίζεται ότι υπάρχουν 10 δισεκατομμύρια νευρώνες και 60 τρισεκατομμύρια συνδέσεις στον φλοιό του ανθρώπινου εγκεφάλου. Η βασική δομή ενός νευρώνα αποτελείται από τρία βασικά μέρη. Τους δενδρίτες το σώμα και τον άξονα (Σχήμα 68). Όπως φαίνεται από το Σχήμα 68 ένας νευρώνας αποτελείται από το σώμα στο οποίο γίνεται η επεξεργασία



Σχήμα 68: Απλουστευμένο μοντέλο βιολογικού νευρώνα

των πληροφοριών, απο τους δενδρίτες εισόδου από τους οποίους εισέρχονται οι νέες πληροφορίες, και από τον άξονα με τους δενδρίτες εξόδου με τους οποίους γίνεται η εξαγωγή των πληροφοριών και η μεταδοσή τους σε άλλους νευρώνες.Οι πληροφορίες μεταφέρονται μεταξύ των νευρώνων με μορφή ηλεκτρικών ερεθισμάτων κατά μήκος των δενδριτών. Οι εισερχόμενες πληροφορίες που φθάνουν στους δενδρίτες εισόδου του νευρώνα

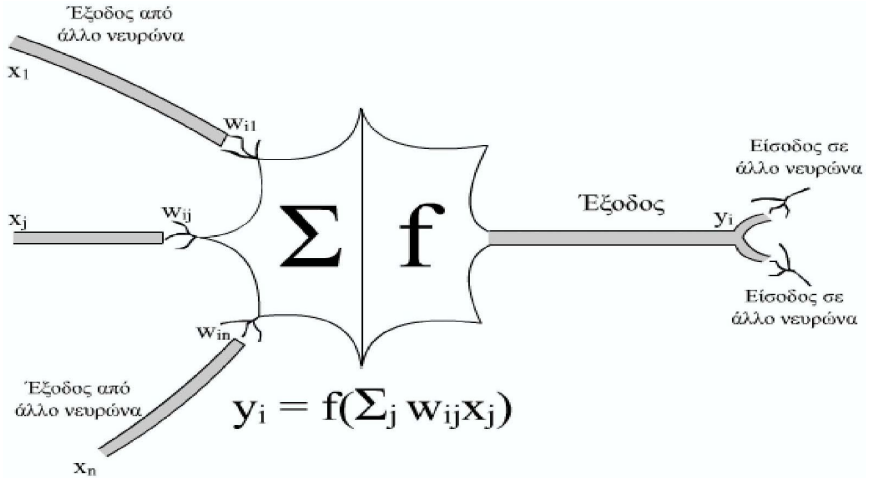
προστίθενται και ύστερα μεταφέρονται κατα μήκος του άξονα του νευρώνα στους δενδρίτες εξόδου, όπου μέσω αυτών οι πληροφορίες θα μεταδοθούν σε άλλους νευρώνες αν το ερέθισμα έχει ξεπεράσει ένα κατώ όριο. Σε αυτή τη περίπτωση ο νευρώνας ενεργοποιείται. Στο Σχήμα 69 φαίνεται πιο καθαρά η πιο πάνω διαδικασία.



Σχήμα 69: Μοντέλο βιολογικού νευρώνα

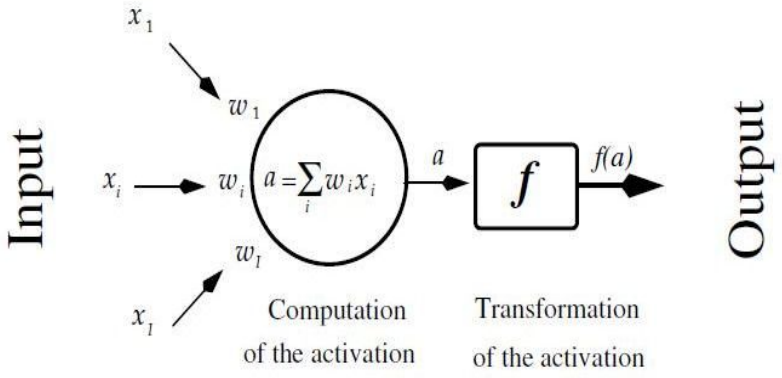
5.2.2 Τεχνητά Νευρωνικά Δίκτυα

Το πραγματικό βιολογικό νευρικό σύστημα είναι ιδιαίτερα σύνθετο και είναι πολύ δύσκολο να κατανοήσουμε με ακρίβεια τον τρόπο λειτουργείας του. Ένα Τεχνητό Νευρωνικό Δίκτυο αποτελείται από διασυνδεδεμένους τεχνητούς νευρώνες οι οποίοι είναι κατασκευάσματα προγραμματισμού που μιμούνται τις ιδιότητες των βιολογικών νευρώνων. Τα Τεχνητά Νευρωνικά Δίκτυα μπορούν να χρησιμοποιηθούν για την επίλυση προβλημάτων τεχνητής νοημοσύνης χωρίς να χρειαστεί να κατασκευάσουμε απαραίτητα ένα πραγματικό μοντέλο βιολογικού συστήματος. Επίσης τα Τεχνητά Νευρωνικά Δίκτυα μας βοηθάνε να κατανοήσουμε καλύτερα τα Βιολογικά Νευρωνικά Δίκτυα. Στο Σχήμα 70 φαίνεται η μορφή ενός τεχνητού νευρώνα.



Σχήμα 70: Μοντέλο τεχνητού νευρώνα

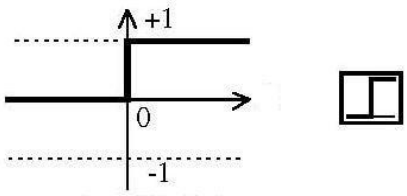
Μπορούμε να παρατηρήσουμε από τα Σχήματα 69 και 70 ότι ένας τεχνητός νευρώνας έχει πολλές ομοιότητες με ένα βιολογικό νευρώνα. Πιο αναλυτικά ένας τεχνητός νευρώνας δέχεται πληροφορίες από άλλους τεχνητούς νευρώνες ή εξωτερικούς παράγοντες με τους οποίους είναι συνδεδεμένος. Σε κάθε σύνδεση αντιστοιχεί ένα βάρος. Αυτά τα βάρη πολλαπλασιάζονται με τις εισερχόμενες πληροφορίες και αθροίζονται με τη βοήθεια μιας αθροιστικής συνάρτησης. Έτσι αν συμβολίσουμε με x_i τις τιμές των εισερχόμενων πληροφοριών και με w_i τις τιμές των αντίστοιχων βαρών τότε έχουμε το άθροισμα $a = \sum x_i w_i$. Στη συνέχεια αυτό το άθροισμα μετασχηματίζεται σε μια απάντηση μέσω μιας συνάρτησης ενεργοποίησης $f(a)$ και διαδίδεται σε άλλους νευρώνες. Στο Σχήμα 71 παρουσιάζεται αυτή η διαδικασία.



Σχήμα 71: Τεχνητός νευρώνας

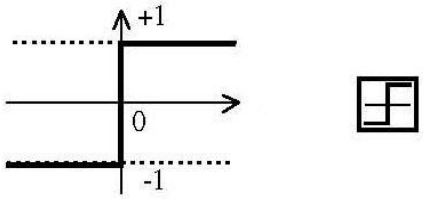
Σε αυτό το σημείο θα παρουσιάσουμε μερικές από τις πιο συχνά χρησιμοποιημένες συναρτήσεις ενεργοποίησης.

• $Hard - Limit\, Transfer\, Function$



Η πιο πάνω συνάρτηση ορίζει την έξοδο του νευρώνα με τιμή 0 αν το άθροισμα a είναι μικρότερο από 0 και με τιμή 1 αν το άθροισμα a είναι μεγαλύτερο ή ίσο από 0.

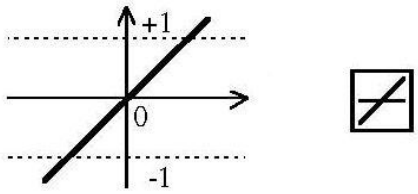
• $Symmetric\, Hard - Limit\, Transfer\, Function$



Η πιο πάνω συνάρτηση ορίζει την έξοδο του νευρώνα με τιμή -1 αν το άθροισμα a είναι μικρότερο από 0 και με

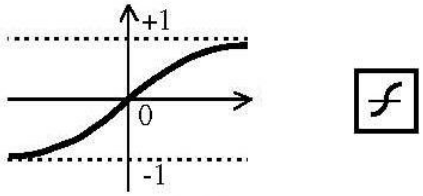
τιμή 1 αν το άθροισμα a είναι μεγαλύτερο ή ίσο από 0.

• $Linear\,Transfer\,Function$



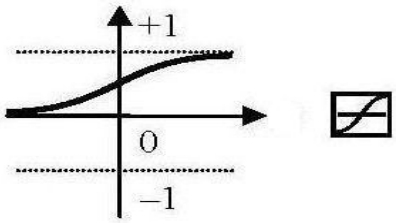
Η γραμμική συνάρτηση δίνει στην έξοδο του νευρώνα την τιμή που έχει το άθροισμα a.

• $Tan-Sigmoid\,Transfer\,Function$



Η υπερβολική εφαπτομένη περιορίζει την έξοδο του νευρώνα στο διάστημα (-1,1).

• $Log-Sigmoid\,Transfer\,Function$

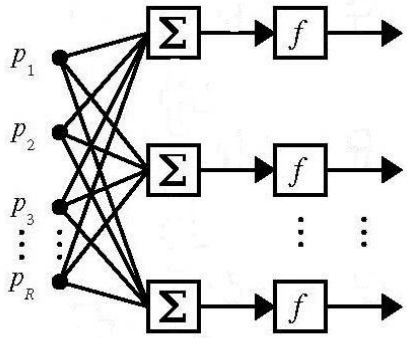


Η λογιστική συνάρτηση περιορίζει την έξοδο του νευρώνα στο διάστημα (0,1).

5.2.3 Αρχιτεκτονική Τεχνητών Νευρωνικών Δικτύων

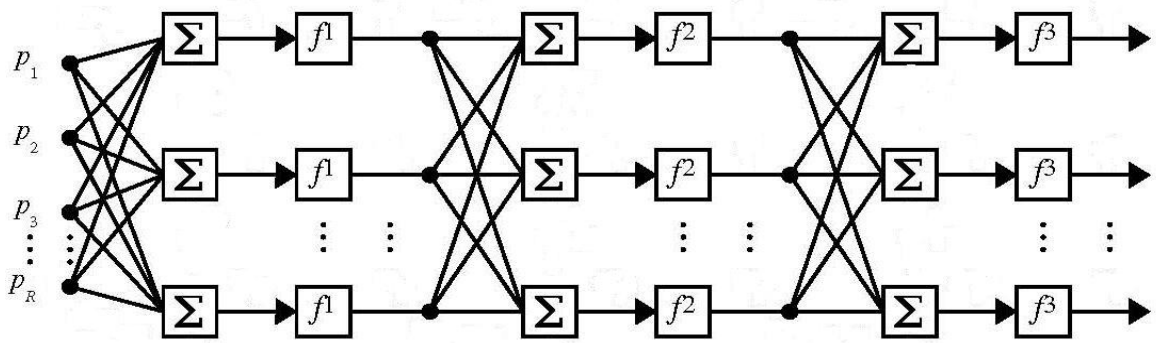
Ο τρόπος με τον οποίο είναι διασυνδεδεμένει οι νευρώνες ενός δικτύου ονομάζεται αριτεκτονική του Τεχνητού Νευρωνικού Δικτύου. Για λόγους εξυπηρέτησης θεωρούμε ότι οι νευρώνες είναι τοποθετημένει σε στρώματα (layers). Σε κάθε στρώμα οι νευρώνες του δικτύου έχουν σχεδόν πάντα την ίδια συνάρτηση ενεργοποίησης. Κάθε Νευρωνικό Δίκτυο έχει τουλάχιστον δύο στρώματα. Το στρώμα εισόδου από το οποίο εισέρχονται τα δεδομένα εκπαίδευσης και το στρώμα εξόδου από το οποίο παίρνουμε την απάντηση του δικτύου. Αν ένα Νευρωνικό Δίκτυο δεν αποτελείται μόνο από το στρώμα εισόδου και το στρώμα εξόδου τότε το δίκτυο ονομάζεται πολυστρωματικό (multilayer) και τα ενδιάμεσα στρώματα ονομάζονται κρυφά στρώματα (hidden layers). Ένας άλλος τρόπος κατηγοριοποίησης των Τεχνητών Νευρωνικών Δικτύων είναι ο τρόπος με τον οποίο διαδίδονται οι πληροφορίες μεταξύ των νευρώνων. Αν οι πληροφορίες διαδίδονται έτσι ώστε να μην υπάρχει νευρώνας που η εξοδός του να είναι είσοδος κάποιου άλλου νευρώνα, του ιδίου ή προηγούμενου στρώματος τότε το δίκτυο ονομάζεται πρόσθιας τροφοδότησης (feedforward). Επίσης αν κάθε νευρώνας ενός στρώματος συνδέεται με

όλους του νευρώνες του επόμενου στρώματος τότε το δίκτυο είναι πλήρως διασυνδεδεμένο (fullyinterconnected). Τέλος αν η έξοδος κάποιου νευρώνα κινείται προς νευρώνες του ιδίου ή προηγούμενων στρωμάτων τότε το δίκτυο παρουσιάζει ανάδραση (feedback). Στο Σχήμα 72 παρουσιάζεται ένα δίκτυο που αποτελείται μόνο από το στρώμα εισόδου και το στρώμα εξόδου.



Σχήμα 72: Μοντέλο Τεχνητού Νευρωνικού Δικτύου με δύο στρώματα

Το δίκτυο στο Σχήμα 72 είναι ένα δίκτυο με δύο στρώματα, το στρώμα εισόδου και το στρώμα εξόδου. Επίσης είναι πλήρως δισυνδεδεμένο αφού κάθε νευρώνας του πρώτου στρώματος (στρώμα εισόδου) συνδέεται με όλους τους νευρώνες του δεύτερου στρώματος (στρώμα εξόδου). Αυτό το δίκτυο δέχεται σαν είσοδο ένα διάνυσμα με R στοιχεία και μπορεί να έχει μια ή περισσότερες εξόδους. Στο Σχήμα 73 φαίνεται ένα δίκτυο με τρία στρώματα.



Σχήμα 73: Μοντέλο πολυστρωματικού Τεχνητού Νευρωνικού Δικτύου

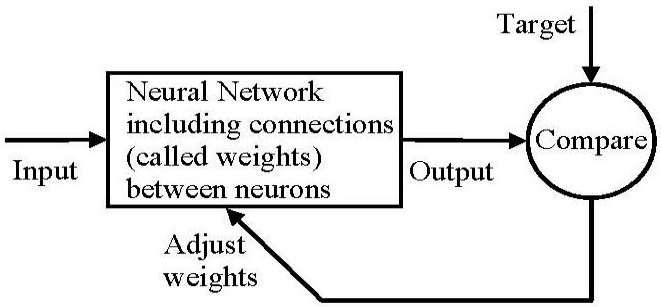
Το δίκτυο στο Σχήμα 73 αποτελείται από τρία στρώματα, το στρώμα εισόδου, ένα κρυφό στρώμα και το στρώμα εξόδου δηλαδή είναι πολυστρωματικό (multilayer). Είναι πρόσθιας τροφοδότησης feedforward) αφού οι πληροφορίες που φεύγουν από κάθε νευρώνα διαδίδονται μόνο σε νευρώνες επόμενων στρωμάτων. Επίσης είναι πλήρως διασυνδεδεμένο (fully interconnected) αφού κάθε νευρώνας ενός στρώματος συνδέεται με όλους του νευρώνες του επόμενου στρώματος. Μπορούμε να παρατηρήσουμε ότι οι νευρώνες που βρίσκονται στο ίδιο στρώμα έχουν την ίδια συνάρτηση ενεργοποίησης. Το κάθε στρώμα δεν έχει απαραίτητα το ίδιο πλήθος νευρώνων με τα άλλα στρώματα. Όπως πρίν αυτό το δίκτυο δέχεται σαν είσοδο ένα διάνυσμα με R στοιχεία και μπορεί να έχει μια ή περισσότερες εξόδους. Στη συνέχεια θα ασχοληθούμε με δίκτυα πολυστρωματικά, πλήρως δισυνδεδεμένα και πρόσθιας τροφοδότησης.

5.3 Εκπαίδευση Τεχνητών Νευρωνικών Δικτύων

Με τον όρο εκπαίδευση του νευρωνικού δικτύου ενοούμε τη διαδικασία με την οποία μεταβάλουμε τα βάρη του δικτύου έτσι ώστε αυτό να μάθει τη σχέση μεταξύ των προτύπων εκπαίδευσης και της επιθυμητής εξόδου. Η προσαρμογή των βαρών γίνεται με κάποιον αλγόριθμο εκπαίδευσης που συνήθως είναι κάποιος αλγόριθμος βελτιστοποίησης. Οι μέθοδοι εκπαίδευσης μπορούν να χωριστούν σε δύο κατηγορίες. Στις μεθόδους εκπαίδευσης με επίβλεψη (supervised learning) και στις μεθόδους εκπαίδευσης χωρίς επίβλεψη (unsupervised learning).

• Εκπαίδευση με επίβλεψη.

Η εκπαίδευση με επίβλεψη είναι μια διαδικασία με την οποία το δίκτυο προσαρμόζεται έτσι ώστε να δίνει συγκεκριμένες απαντήσεις σε συγκεκριμένες εισόδους. Τα δεδομένα εκπαίδευσης αποτελούνται απο ζευγάρια με τα δεδομένα εισόδου και την επιθυμητή έξοδο. Τα δεδομένα εισόδου εισέρχονται στο δίκτυο και αυτό παράγει μια έξοδο. Η έξοδος του δικτύου συγκρίνεται με την επιθυμητή έξοδο και υπολογίζεται η διαφορά τους. Αυτή η διαφορά αποτελεί το σφάλμα εκπαίδευσης του δικτύου. Στη συνέχεια τα βάρη του δικτύου μεταβάλλονται έτσι ώστε η νέα έξοδος του δικτύου να προσεγγίζει με μεγαλύτερη ακρίβεια την επιθυμητή έξοδο και αυτό θα έχει σαν συνέπεια την μείωση του σφάλματος του δικτύου. Ο σκοπός της εκπαίδευσης με επίβλεψη είναι το δίκτυο να μπορεί να προβλέψει με επιτυχία την επιθυμητή έξοδο για δεδομένα που δεν είχε δεί κατά την διάρκεια της εκπαιδευσής του. Βέβαια για να συμβεί κάτι τέτοιο θα πρέπει τα άγνωστα δεδομένα να έχουν ομοιότητες και να αναφέρονται στο ίδιο πρόβλημα με τα δεδομένα εκπαιδευσής. Στο Σχήμα 74 παρουσιάζεται συνοπτικά η πιο πάνω διαδικασία.



Σχήμα 74: Εκπαίδευση με επίβλεψη

• Εκπαίδευση χωρίς επίβλεψη.

Σε πολλές περιπτώσεις είναι δύσκολο να χαρακτηριστούν τα δεδομένα εισόδου όταν αυτά είναι πολλά ή όταν προέρχονται από κάποια άγνωστη διεργασία. Σε αυτές τις περιπτώσεις είναι δυνατό να χρησιμοποιηθεί η εκπαίδευση χωρίς επίβλεψη. Στην εκπαίδευση χωρίς επίβλεψη τα βάρη τροποποιούνται έτσι ώστε να ανταποκρίνονται μόνο στα δεδομένα εισόδου. Δεδομένα που καθορίζουν κάποια επιθυμητή έξοδο δεν υπάρχουν. Οι περισσότεροι τέτοιοι αλγόριθμοι πετυχαίνουν ομαδοποίηση των δεδομένων σε ένα πεπερασμένο σύνολο κατηγοριών. [16,18,19]

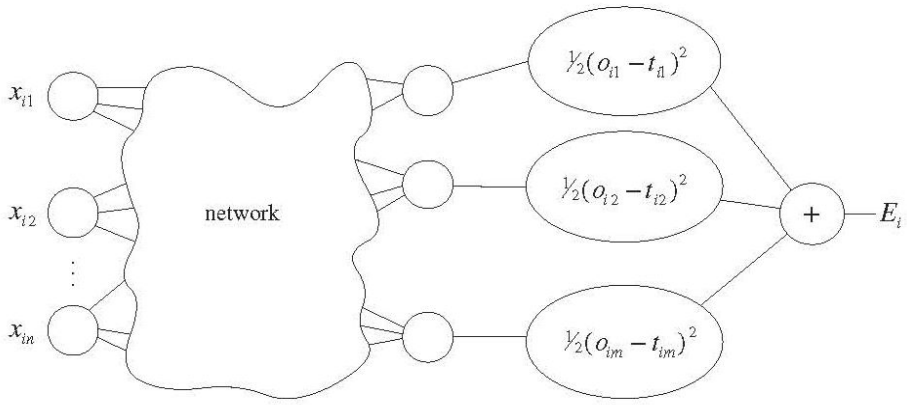
5.4 Αλγόριθμος Backpropagation

Όπως αναφέραμε πρίν ένα δίκτυο πρόσθιας τροφοδότησης είναι ένα γράφος του οποίου οι κόμβοι είναι ανεξάρτητες μονάδες επεξεργασίας και οι ακμές του έχουν την ικανότητα να μεταφέρουν αριθμητικές πληροφορίες από κόμβο σε κόμβο. Το πρόβλημα της εκπαίδευσης αποτελείται από την εύρεση του βέλτιστου συνδυασμού των βαρών του δικτύου έτσι ώστε αυτό να μπορεί να προσεγγίσει όσο γίνεται καλύτερα τις τιμές μιας επιθυμητής συνάρτησης. Έστω ένα δίκτυο πρόσθιας τροφοδότησης το οποίο αποτελείται από n εισόδους και m εξόδους. Αυτό μπορεί να έχει οποιοδήποτε αριθμό από κρυφούς νευρώνες και να είναι πλήρως ή οχι δισυνδεδεμένο. Επίσης

δίνεται ένα σύνολο εκπαίδευσης $(\mathbf{x}_1, \mathbf{t}_1), ..., (\mathbf{x}_p, \mathbf{t}_p)$ το οποίο αποτελείται απο p ζευγάρια εκπαίδευσης. Το κάθε ζευγάρι $(\mathbf{x}_i, \mathbf{t}_i)$ αποτελείται από τα δεδομένα εισόδου $\mathbf{x}_i$ που έχουν διάσταση n και τα δεδομένα εξόδου που έχουν διάσταση m. Τα βάρη του δικτύου είναι πραγματικοί αριθμοί οι οποίοι αρχικοποιούνται τυχαία. Όταν τα δεδομένα εισόδου $\mathbf{x}_i$ απο το σύνολο εκπαίδευσης εισέρχονται στο δίκτυο τότε αυτό παράγει μια έξοδο $\mathbf{o}_i$ που γενικά θα είναι διαφορετική από την επιθυμητή έξοδο $\mathbf{t}_i$. Αυτό που θέλουμε να πετύχουμε είναι να κάνουμε τα $\mathbf{o}_i$ και $\mathbf{t}_i$ πανομοιότυπα για όλα τα $i = 1, 2, ..., p$ χρησιμοποιώντας έναν αλγόριθμο εκπαίδευσης. Πιο συγκεκριμένα θέλουμε να ελαχιστοποιήσουμε τη συνάρτηση σφάλματος του δικτύου:

$$E = \frac{1}{2} \sum_{i=1}^{p} ||\mathbf{o}_i - \mathbf{t}_i||^2 \qquad (5.1)$$

Αφού ελαχιστοποιήσουμε την πιο πάνω συνάρτηση, νέα άγνωστα δεδομένα θα εισέλθουν στο δίκτυο και αυτό που περιμένουμε να κάνει είναι να μπορέσει να αναγνωρίσει αν αυτά τα δεδομένα έχουν κάποια σχέση με κάποια από τα δεδομένα εκπαίδευσης και να δώσει μια παρόμοια απάντηση. Ο αλγόριθμος Backpropagation χρησιμοποιείται για να βρεί ένα τοπικό ελάχιστο της συνάρτησης σφάλματος του δικτύου. Η κλίση της συνάρτησης σφάλματος υπολογίζεται και χρησιμοποιείται για να διορθώσει τα αρχικά βάρη. Έτσι ο σκοπός μας είναι να υπολογίσουμε αυτή την κλίση με επαναληπτικό τρόπο. Το πρώτο βήμα της διαδικασίας ελαχιστοποίησης είναι να επεκτείνουμε το δίκτυο έτσι ώστε η συνάρτηση σφάλματος να υπολογίζεται αυτόματα Σχήμα 75.



Σχήμα 75: Επεκτεταμένο δίκτυο για τον υπολογίσμό της συνάρτησης σφάλματος

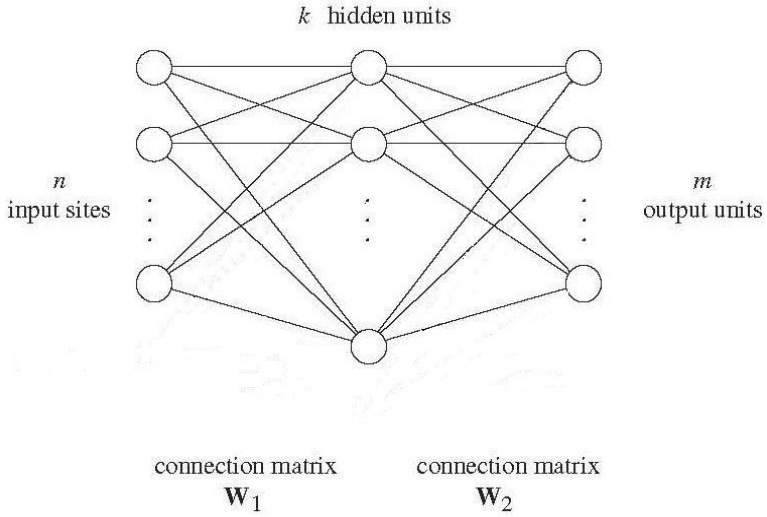
Κάθε ένας από τους κόμβους εξόδου του δικτύου είναι συνδεδεμένος με ένα κόμβο ο οποίος υπολογίζει τη συνάρτηση $\frac{1}{2}(o_{ij} - t_{ij})^2 \;\; j = 1, 2, ..., m$ όπου τα o_{ij} και t_{ij} είναι οι $j-$ωστές συνιστώσες των διανυσμάτων $\mathbf{o}_i$ και $\mathbf{t}_i$ αντίστοιχα. Στο τέλος αυτές οι συναρτήσεις αθροίζονται και δίνουν το άθροισμα E_i. Για κάθε ζευγάρι $(\mathbf{x}_i, \mathbf{t}_i)$ υπολογίζεται μία συνάρτηση E_i.Μόλις εισέλθουν και τα p ζευγάρια εκπαίδευσης στο δίκτυο θα έχουμε τις συναρτήσεις $E_1, E_2, ..., E_p$. Το άθροισμα αυτών των συναρτήσεων δίνουν την συνάρτηση σφάλματος του δικτύου $E = E_1 + E_2 + ... + E_p$. Τα βάρη του δικτύου είναι οι μόνοι παράμετροι που μπορούν να τροποποιηθούν για να μειωθεί το τετραγωνικό σφάλμα E. Για την μείωση του τετραγωνικού σφάλματος θα χρησιμοποιήσουμε την επαναληπτική μέθοδο Gradient Descent για την οποία πρέπει να υπολογίσουμε την κλίση:

$$\nabla E = (\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, ..., \frac{\partial E}{\partial w_l}) \qquad (5.2)$$

Κάθε βάρος αλλάζει σύμφωνα με την πιο κάτω σχέση:

$$w_{i+1} = w_i - \gamma \frac{\partial E}{\partial w_i}, \; i = 1, 2, ..., l-1 \qquad (5.3)$$

όπου το γ είναι μια παράμετρος που ορίζει το βήμα σε κάθε επανάληψη στην κατεύθυνση της αρνητικής κλίσης. Τώρα το πρόβλημα της εκπαίδευσης περιορίζεται στο να βρούμε ένα τρόπο για να υπολογίσουμε την κλίση της συνάρτησης σφάλματος ως προς τα βάρη του δικτύου. Αν έχουμε μια μέθοδο που θα υπολογίζει αυτή την κλίση τότε μπορούμε να διορθώσουμε τα βάρη του δικτύου επαναληπτικά. Στη συνέχεια θα παρουσιάσουμε τα βήματα που ακολουθεί ο αλγόριθμος Backpropagation . Υποθέτουμε ότι έχουμε το δίκτυο που φαίνεται στο Σχήμα 76.



Σχήμα 76:

Αυτό το δίκτυο αποτελείται από το στρώμα εισόδου με n νευρώνες, από ένα κρυφό στρώμα με k κρυφούς νευρώνες και από το στρώμα εξόδου με m νευρώνες. Το βάρος της ακμής που συνδέει το νευρώνα i από το στρώμα εισόδου και τον νευρώνα j από το κρυφό στρώμα το συμβολίζουμε με $w_{ij}^{(1)}$. Το βάρος της ακμής που συνδέει το νευρώνα i από το κρυφό στρώμα και τον νευρώνα j από το στρώμα εξόδου το συμβολίζουμε με $w_{ij}^{(2)}$. Υπάρχουν $n \times k$ βάρη μεταξύ των νευρώνων του στρώματος εισόδου και του κρυφού στρώματος και $k \times m$ βάρη μεταξύ των νευρώνων του κρυφού στρώματος και του στρώματος εξόδου. Συμβολίζουμε με W_1 τον πίνακα $n \times k$ με συντελεστή το $w_{ij}^{(1)}$ στην i-γραμμή και j-στήλη. Όμοια συμβολίζουμε με W_2 τον πίνακα $k \times m$ με συντελεστή το $w_{ij}^{(2)}$ στην i-γραμμή και j-στήλη. Επίσης θεωρούμε το διάνυσμα με δεδομένα εισόδου $O = (o_1, o_2, ...o_n)$. Το άθροισμα στον j-νευρώνα του κρυφού στρώματος συμβολίζεται με net_j και περιγράφεται από την πιο κάτω σχέση:

$$net_j = \sum_{i=1}^{n} w_{ij}^{(1)} o_i \qquad (5.4)$$

Η συνάρτηση ενεργοποίησης είναι σιγμοειδής και η έξοδος από τον j-νευρώνα του κρυφού στρώματος συμβολίζεται με $o_j^{(1)}$ και δίνεται από την πιο κάτω σχέση:

$$o_j^{(1)} = s(\sum_{i=1}^{n} w_{ij}^{(1)} o_i) \qquad (5.5)$$

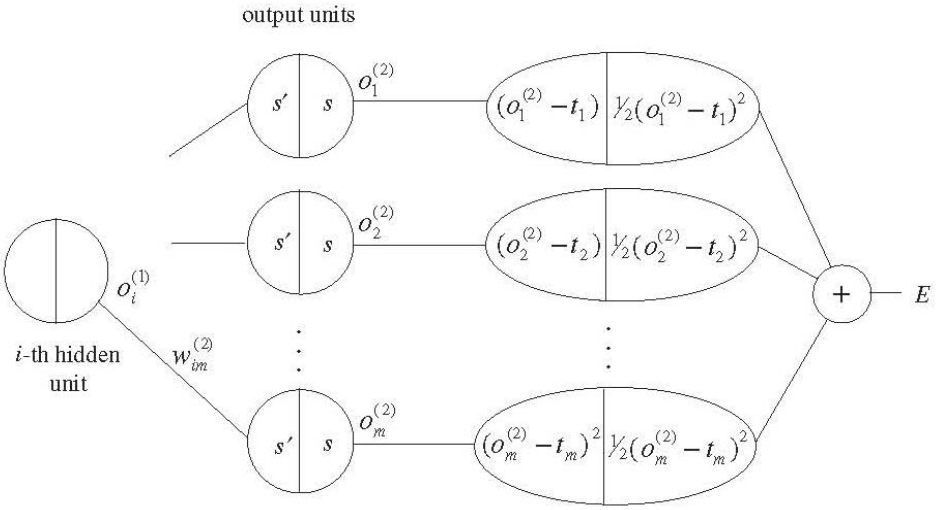
Το άθροισμα που προκύπτει σε κάθε κρυφό νευρώνα μπορεί να υπολογιστεί με τον πολλαπλασισμό του διανύσματος O και του πίνακα W_1. Το διάνυσμα του οποίου τα στοιχεία είναι οι έξοδοι των κρυφών νευρώνων δίνεται απο την σχέση:

$$O^{(1)} = s(OW_1) \qquad (5.6)$$

Έτσι προκύπτει το διάνυσμα $O^{(1)} = (o_1^{(1)}, o_2^{(1)}, ..., o_k^{(1)})$. Στη συνέχεια με πολλαπλασιασμό του διανύσματος $O^{(1)}$ με τον πίνακα W_2 προκύπτουν τα αθροίσματα στους νευρώνες του στρώματος εξόδου. Οι έξοδοι αυτών των νευρώνων δίνονται από την πιο κάτω σχέση:

$$O^{(2)} = s(O^{(1)} W_2) \qquad (5.7)$$

Στο Σχήμα 77 έχουμε επεκτείνει το δίκτυο έτσι ώστε να υπολογίζει τη συνάρτηση σφάλματος. Θεωρούμε οτι σε κάθε κόμβο του δικτύου αποθηκεύονται δύο τιμές. Στους κόμβους του νέου στρώματος αποθηκεύονται στην δεξιά πλευρά οι τετραγωνικές διαφορές $\frac{1}{2}(o_i^{(2)} - t_i)^2$ και στην αριστερή πλευρά οι παράγωγοι αυτών των τετραγωνικών διαφορών $(o_i^{(2)} - t_i)$ με $i = 1, 2, ..., m$. Στους υπόλοιπους κόμβους του δικτύου αποθηκεύεται στην δεξιά πλευρά η σιγμοειδής συνάρτηση s και στην αριστερή πλευρά η παράγωγος αυτής s'. Για κάθε νευρώνα εξόδου i υπολογίζεται η σιγμοειδή συνάρτηση s και παράγεται η έξοδος $o_i^{(2)}$. Το άθροισμα των τετραγωνικών διαφορών δίνουν το σφάλμα E για το ένα ζευγάρι από το σύνολο εκπαίδευσης. Αφού υπολογίσουμε το σφάλμα και για τα p ζευγάρια από το σύνολο εκπαίδευσης τότε το άθροισμα αυτών των σφαλμάτων δίνει τη συνάρτηση σφάλματος του δικτύου.



Σχήμα 77: Επεκτεταμένο δίκτυο

Ο αλγόριθμος Backpropagation μπορεί να περιγραφθεί από τα επόμενα τέσσερα βήματα:

• Feed-forward computation .

• Backpropagation to the output layer .

• Backpropagation to the hidden layer .

• Weight updates .

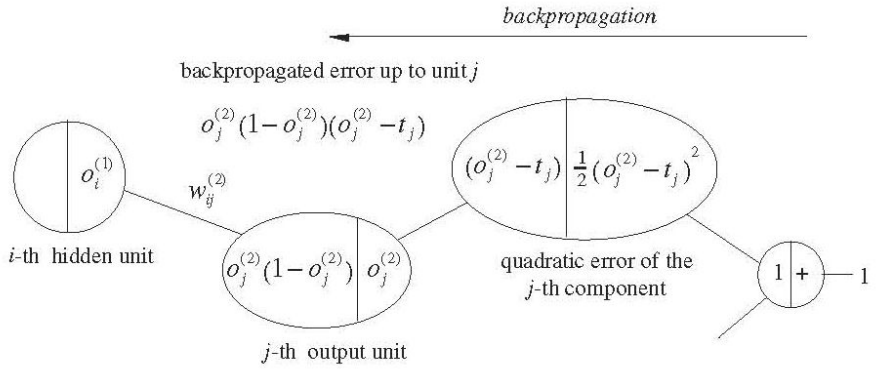
Ο αγόριθμος σταματάει οταν η τιμή της συνάρτησης σφάλματος γίνει ικανοποιητικά μικρή.

Πρώτο βήμα: Feed-forward computation

Το διάνυσμα O εισέρχεται στο δίκτυο. Τα διανύσματα $O^{(1)}$ και $O^{(2)}$ υπολογίζονται και αποθηκεύονται. Επίσης υπολογίζεται το σφάλμα E και οι παράγωγοι των συναρτήσεων ενεργοποίησης σε κάθε κόμβο.

Δεύτερο βήμα: Backpropagation to the output layer

Σε αυτό το βήμα θέλουμε να υπολογίσουμε τις μερικές παραγώγους $\partial E/\partial w_{ij}^{(2)}$ Το μονοπάτι που ακολουθεί ο αλγόριθμος από την έξοδο του επεκτεταμένου δικτύου μέχρι τον κόμβο εξαγωγής j φαίνεται στο Σχήμα 78.



Σχήμα 78: Μονοπάτι μέχρι τον κόμβο εξαγωγής j

Ο αλγόριθμος κινείται από το τέλος προς την αρχή του δικτύου. Σαν είσοδο του δικτύου χρησιμοποιούμε το 1. Αφού πολλαπλασιάσουμε όλα τα αριστερά μέλη των κόμβων του μονοπατιού μέχρι και τον κόμβο j προκύπτει το σφάλμα μέχρι τον κόμβο j. Η σχέση $(o_j^{(2)} - t_j)$ είναι η παράγωγος της τετραγωνικής διαφοράς $\frac{1}{2}(o_j^{(2)} - t_j)^2$ και η σχέση $o_j^{(2)}(1 - o_j^{(2)})$ είναι η παράγωγος της σιγμοειδούς συνάρτησης του κόμβου j. Μετά τον πολλαπλασιασμό το σφάλμα στον κόμβο j δίνεται από την σχέση:

$$\delta_j^{(2)} = o_j^{(2)}(1 - o_j^{(2)})(o_j^{(2)} - t_j) \qquad (5.8)$$

Αποδεικνύεται με τον κανόνα της αλυσίδας ότι :

$$\frac{\partial E}{\partial w_{ij}^{(2)}} = [o_j^{(2)}(1 - o_j^{(2)})(o_j^{(2)} - t_j)]o_i^{(1)} = \delta_j^{(2)} o_i^{(1)} \qquad (5.9)$$

Τρίτο βήμα: Backpropagation to the hidden layer

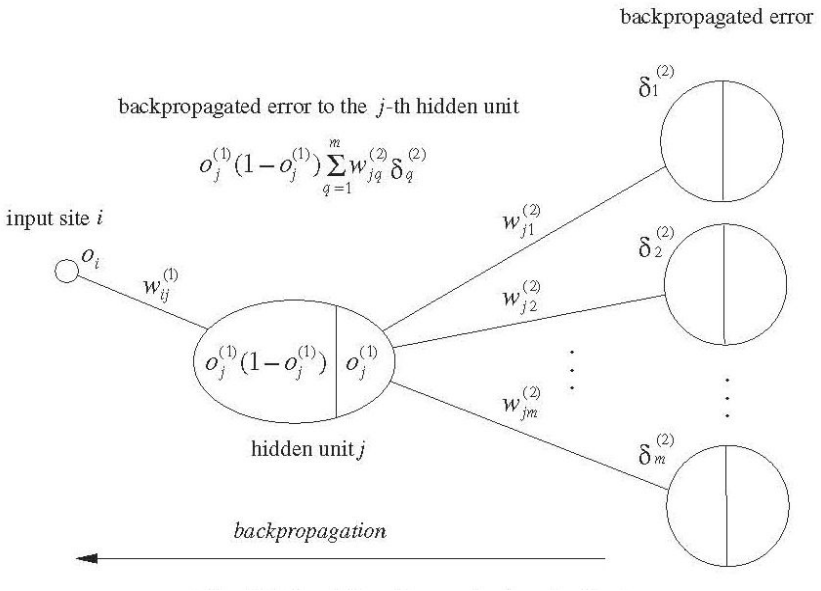
Στο τρίτο βήμα θέλουμε να υπολογίσουμε τις μερικές παραγώγους $\partial E/\partial w_{ij}^{(1)}$. Κάθε κόμβος j του κρυφού στρώματος είναι συνδεδεμένος με κάθε κόμβο q του στρώματος εξαγωγής με μια ακμή με βάρος $w_{jq}^{(2)}$ για $q = 1, 2, ..., m$. Το σφάλμα μέχρι και τον κόμβο j του κρυφού στρώματος υπολογίζεται αφού χρησιμοποιήσουμε όλα τα μονοπάτια που καταλήγουν σε αυτόν τον κόμβο από τους κόμβους του στρώματος εξαγωγής Σχήμα 79. Το σφάλμα στον κόμβο j του κρυφού στρώματος υπολογίζεται από την σχέση:

$$\delta_j^{(1)} = o_j^{(1)}(1 - o_j^{(1)}) \sum_{q=1}^{m} w_{jq}^{(2)} \delta_q^{(2)} \qquad (5.10)$$

Με τον κανόνα της αλυσίδας προκύπτουν οι μερικοί παράγωγοι:

$$\partial E/\partial w_{ij}^{(1)} = \delta_j^{(1)} o_i \qquad (5.11)$$

Με τον ίδιο τρόπο μπορούμε να υπολογίσουμε το σφάλμα για οποιοδήποτε αριθμό κρυφών στρωμάτων που έχουν τη δομή του δικτύου στο Σχήμα 76.



Σχήμα 79: Μονοπάτια μέχρι τον κρυφό κόμβο j

Τέταρτο βήμα: Weight updates

Αφού υπολογίσουμε τις μερικές παραγώγους τότε μπορούμε να υπολογίσουμε τα νέα βάρη του δικτύου. Χρησιμοποιώντας την επαναληπτική μέθοδο Gradient Descent έχουμε:

$$\Delta w_{ij}^{(2)} = -\gamma o_i^{(1)} \delta_j^{(2)} \ , \ i = 1, 2, ..., k+1 \ , \ j = 1, 2, ..., m \qquad (5.12)$$

και

$$\Delta w_{ij}^{(1)} = -\gamma o_i \delta_j^{(1)} \ , \ i = 1, 2, ..., n+1 \ , \ j = 1, 2, ..., k \qquad (5.13)$$

όπου $o_{n+1} = o_{k+1}^{(1)} = 1$.

Με την διόρθωση των βαρών του δικτύου ολοκληρώνεται μια επανάληψη του αλγορίθμου Backpropagation . Τα πιο πάνω βήματα θα εκτελεστούν πολλές φορές μέχρι το σφάλμα του δικτύου να μειωθεί ικανοποιητικά. Ο αριθμός των επαναλήψεων δεν είναι γνωστός από την αρχή και εξαρτάται από το πλήθος και το είδος των δεδομένων του προβλήματος που έχουμε να αντιμετωπίσουμε [17,20,21].

6 Αναγνώριση αριθμών

Ενα πρόβλημα που θα προσπαθήσουμε να προσεγγίσουμε είναι η αναγνώριση αριθμών από φωτογραφία. Αρχικά θα περιγράψουμε συνοπτικά το πρόβλημα και στη συνέχεια θα δώσουμε πιο αναλυτικά τη διαδικασία που ακολουθήσαμε προκυμένου να το προσεγγίσουμε. Οι φωτογραφίες με τους αριθμούς έχουν παραχθεί από κείμενο Microsoft Office Word 2003 , είναι τύπου .jpg και οι αριθμοί έχουν μέγεθος γραμματοσειράς 26 (Σχήμα 80).

Σχήμα 80:

Σκοπός είναι να κατασκευαστεί μια διαδικασία που θα μπορεί να αναγνωρίζει το κάθε ψηφίο (0,1,2,...,9) χωριστά και στη συνέχεια να αναγνωρίζει ολόκληρο τον αριθμό. Για να το πετύχουμε αυτό θα χρειαστεί να χρησιμοποιήσουμε ορισμένες τεχνικές επεξεργασίας εικόνας που έχουμε ήδη αναφέρει σε συνδυασμό με τα Τεχνητά Νευρωνικά Δίκτυα. Σχετικά με το κομμάτι του προγραμματισμού έχει χρησιμοποιηθεί η γλώσσα προγραμματισμού Matlab καθώς διαθέτει μεγάλο πλήθος συναρτήσεων που έχουν σχέση με την επεξεργασία εικόνας και τα Τεχνητά Νευρωνικά Δίκτυα. Στη συνέχεια θα περιγράψουμε την διαδικασία αναγνώρισης των αριθμών σε βήματα.

Πρώτο βήμα: Εντοπισμός συνόρου

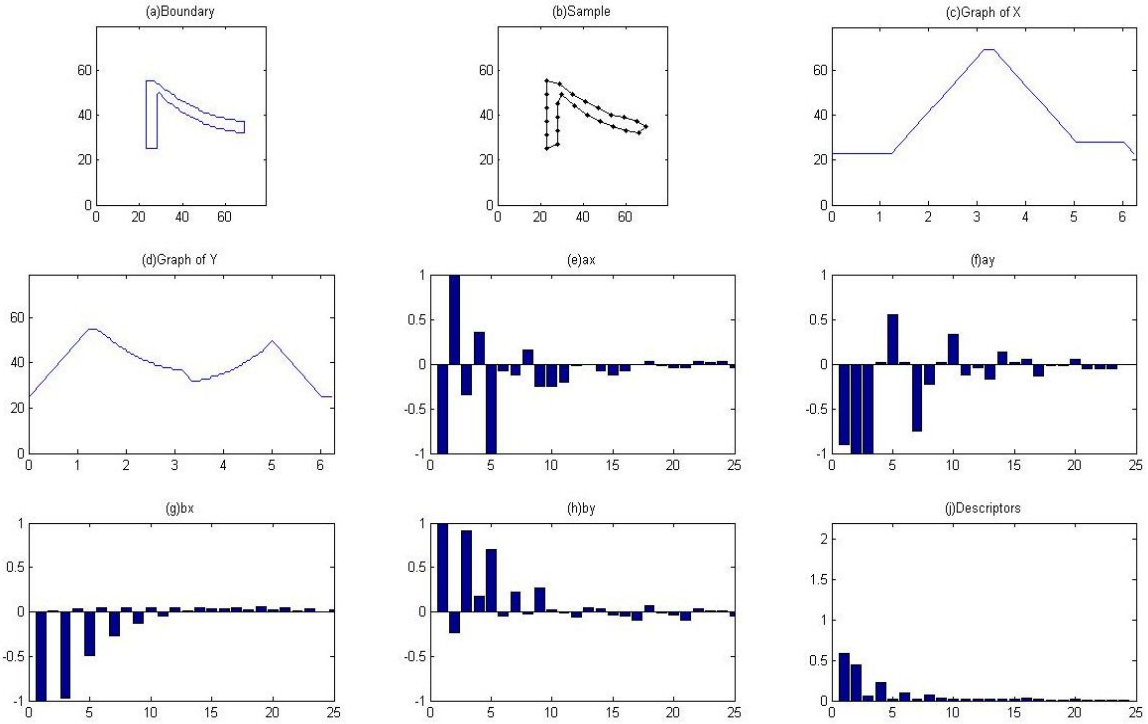
Αρχικά αυτό που θέλουμε να βρούμε είναι το σύνορο του κάθε ψηφίου. Χρησιμοποιώντας τη διαδικασία που έχουμε αναφέρει στην παράγραφο 4 μπορούμε να εντοπίσουμε το σύνορο του κάθε ψηφίου (Σχήμα 81).

Σχήμα 81: Σύνορο αριθμού

Στο Σχήμα 81 έχουμε χρησιμοποιήσει τον αριθμό 7. Χρησιμοποιώντας τη διαδικασία της συνδετικότητας με οκτώ γείτονες που έχουμε περιγράψει στην παράγραφο 4.2 εντοπίζουμε το περίγραμμα του αριθμού επτά.

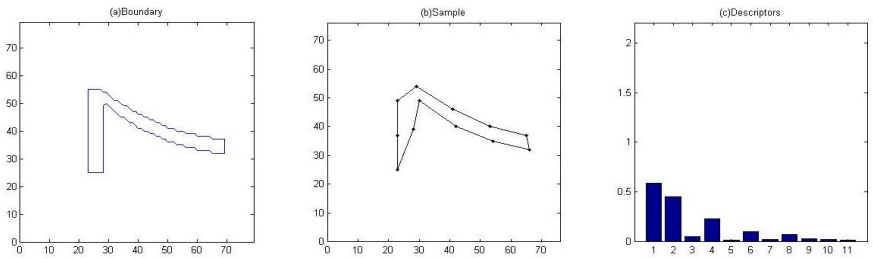
Δεύτερο βήμα: Περιγραφή συνόρου με θεωρία Fourier

Αφού έχουμε βρεί το σύνορο του ψηφίου τότε μπορούμε να το περιγράψουμε χρησιμοποιώντας τη θεωρία Fourier που αναφέραμε στην Παράγράφο 4. Στο Σχήμα 82 έχουμε σχηματίσει τον αριθμό 7 με 25 σημεία από το σύνορό του τα οποία ισαπέχουν μεταξύ τους.



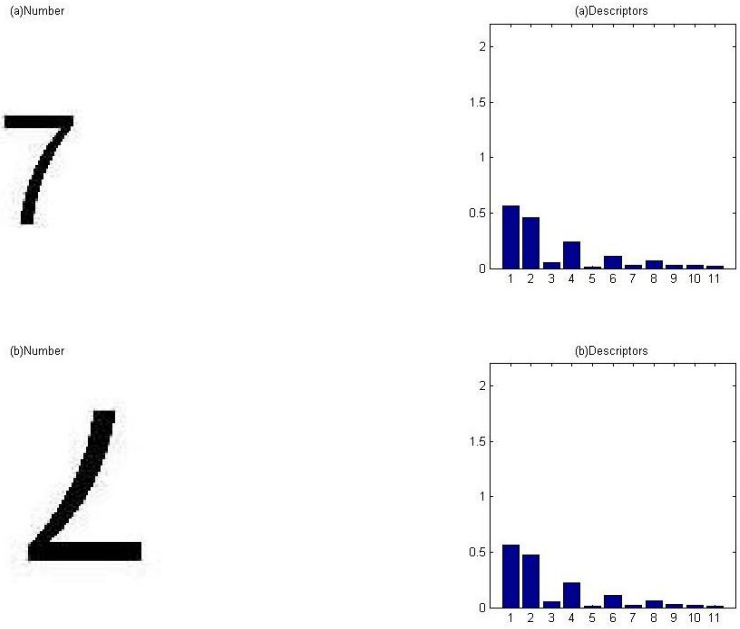
Σχήμα 82:

Στο Σχήμα 82a και στο Σχήμα 82b έχουμε το αρχικό περίγραμμα και το περίγραμμα που σχηματίζεται από 25 ισαπέχοντα σημεία αντίστοιχα. Στα Σχήματα 82c και 82d έχουμε την ανάλυση του αρχικού περιγράμματος ως προς την x και y συντεταγμένη. Τα Σχήματα 82e , 82f , 82g και 82h παρουσιάζουν τους συντελεστές Fourier της σχέσης 4.29 αντίστοιχα. Τέλος στο Σχήμα 82j φαίνονται οι αναλλοίωτοι συντελεστές Fourier της Σχέσης 4.45. Δεν υπάρχει κάποιος κανόνας που να καθορίζει τον αριθμό των συντελεστών. Αυτός προκύπτει ύστερα από πειραματισμό και από τις ανάγκες του προβλήματος. Ο αριθμός των συντελεστών που έχουμε επιλέξει για να περιγράψουμε το σύνορο του κάθε ψηφίου είναι 12 (Σχήμα 83).



Σχήμα 83:

Σε αυτό το σημείο ιδιαίτερα σημαντικό είναι να δείξουμε τις ιδιότητες των συντελεστών Fourier (Μετατόπιση - Αλλαγή Κλίμακας - Περιστροφή) που έχουμε αναφέρει στην παράγραφο 4.2. Στο Σχήμα 84 μπορούμε να δούμε ότι οι συντελεστές Fourier δεν μεταβάλονται οταν ο αριθμός 7 μετατοπιστεί, αλλάξει στην κλίμακα και περιστραφεί.

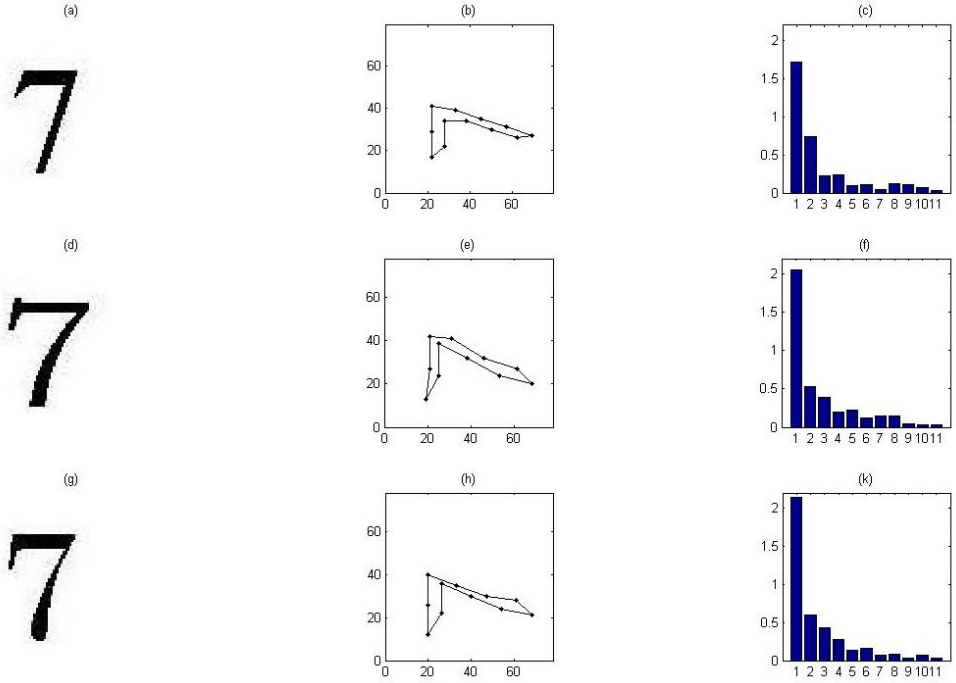


Σχήμα 84: Αναλλοίωτοι συντελεστές Fourier

Στο Σχήμα 84 ο αριθμός 7 μεγευθύνεται, μετατοπίζεται δεξιά σε σχέση με την αρχική του θέση και περιστρέφεται κατά 180 μοίρες. Οι συντελεστές Fourier του αριθμού 7 είναι σχεδόν ίδιοι πριν και μετά την τροποποιησή του. Υπάρχουν βέβαια μικρές αποκλίσεις οι οποίες οφείλονται σε διάφορους παράγοντες όπως είναι η ανάλυση της εικόνας, η προεργασία εύρεσης και κατασκευής της καμπύλης κ.α. Οι αποκλίσεις όμως είναι αρκετά μικρές που όπως θα δούμε στην συνέχεια δεν επηρεάζουν σημαντικά τη διαδικασία αναγνώρησης των ψηφίων αν αυτά έχουν τροποποιηθεί με παρόμοιο τρόπο.

Τρίτο βήμα: Κατασκευή βάσης δεδομένων

Αφού μπορούμε να χαρακτηρίσουμε το σύνορο του κάθε ψηφίου με έναν πεπερασμένο αριθμό συντελεστών Fourier , τότε αυτό που θα κάνουμε στη συνέχεια είναι να κατασκευάσουμε μια βάση δεδομένων η οποία θα περιέχει αυτούς τους συντελεστές για κάθε ψηφίο. Όπως έχουμε αναφέρει και πρίν 12 θα είναι οι συντελεστές που θα αντιπροσωπέυουν το κάθε ψηφίο. Πιο συγκεκριμένα για κάθε ψηφίο έχουν χρησιμοποιηθεί 40 διαφορετικοί γραφικοί χαρακτήρες από τη λίστα γραμματοσειράς του Microsoft Office Word 2003 με μέγεθος γραμματοσειράς 26. Στο Σχήμα 85 έχουμε τον αριθμό 7 σε τρείς διαφορετικούς γραφικούς τύπους. Η κατασκευή του κάθε αριθμού με 12 σημεία μας βοηθάει να διατηρήσουμε την αρχική του μορφή χωρίς να κρατήσουμε κάποιες περιττές λεπτομέρειες του συνόρου. Τα νέα σύνορα που έχουν κατασκευαστεί απο 12 σημεία έχουν πολλές ομοιότητες μεταξύ τους ακόμα και αν τα αρχικά σύνορα είχαν σημαντικές διαφορές. Αποτέλεσμα είναι οι συντελεστές Fourier που προκύπτουν για τα σύνορα με 12 σημεία να είναι όμοιοι μεταξύ τους.



Σχήμα 85: Συντελεστές για τρείς διαφορετικούς γραφικούς τύπους

Τέταρτο βήμα: Κατασκευή και εκπαίδευση Τεχνητού Νευρωνικού Δικτύου

Η βάση δεδομένων που έχουμε κατασκευάσει στο τρίτο βήμα τώρα μπορεί να χρησιμοποιηθεί για την εκπαίδευση ενός νευρωνικού δικτύου το οποίο θα αναγνωρίζει αριθμούς. Για την κατασκευή του δικτύου έχουμε χρησιμοποιήσει έτοιμες συναρτήσεις που διαθέτει η γλώσσα προγραμματισμού Matlab . Στη συνέχεια δίνεται τμήμα κώδικα που ορίζει τη δομή του δικτύου που χρησιμοποιήθηκε για τον πιο πάνω σκοπό.

(Κώδικας)

```
⋮
net = newff(minmax(inputs) , [15 15 1] , 'logsig' , 'logsig' , 'purelin' , 'trainlm');
net = init(net);

net.trainParam.epochs = 2000;
net.trainParam.show = 50;
net.trainParam.lr = 0.3;
net.trainParam.goal = 1.e-2;
⋮
```

Το πιο πάνω δίκτυο είναι ένα πολυστρωματικό, πλήρως δισυνδεδεμένο και πρόσθιας τροφοδότησης δίκτυο. Αποτελείται από το στρώμα ειδόδου με 12 νευρώνες, δύο κρυφά στρώματα με 15 κρυφούς νευρώνες το καθένα

και το στρώμα εξόδου με ένα νευρώνα. Στα κρυφά στρώματα έχει χρησιμοποιηθεί η λογιστκή συνάρτηση ενεργοποίησης και στο στρώμα εξόδου γραμμική συνάρτηση ενεργοποίησης. Η συνάρτηση εκπαίδευσης που χρησιμοποιήθηκε είναι η trainlm (Levenberg-Marquardt Backpropagation) . Ο μέγιστος αριθμός επαναλήψεων που θα εκτελέσει το δίκτυο είναι 2000, το βήμα του αλγορίθμου είναι 0.3 και θέλουμε το δίκτυο να έχει ακρίβεια δύο δεκαδικών ψηφίων. Για την εκπαίδευση του δικτύου χρησιμοποιήθηκε η επόμενη συνάρτηση.

```
⋮
net1 = train(net,inputs,targets);
⋮
```

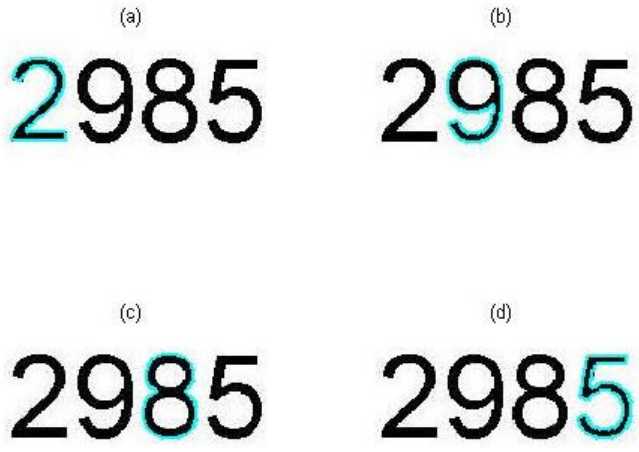
Μετά την εκπαίδευση του δικτύου τα βάρη που έχουν προκύψει αποθηκέυονται στον πίνακα net1 και με τη συνάρτηση:

```
⋮
sim(net1,newinputs)
⋮
```

μπορούμε να ελέγξουμε την απόδοση του δικτύου σε νέα δεδομένα.

Πέμπτο βήμα: Σάρωση της φωτογραφίας από αριστερά προς τα δεξιά

Μέχρι τώρα αυτό που έχουμε πετύχει είναι η αναγνώριση μεμονωνμένων ψηφίων. Ο επόμενος στόχος είναι η κατασκευή μιας διαδικασίας που θα μπορεί να διαβάζει και να αναγνωρίζει αριθμούς με πεπερασμένο πλήθος ψηφίων. Στο πέμπτο και τελευταίο βήμα κατασκευάζουμε εναν αλγόριθμο με τον οποίο σαρώνουμε την φωτογραφία από αριστερά προς τα δεξιά και κάθε φορά που συναντάμε ένα ψηφίο το επεξεργαζόμαστε (βήμα 1,βήμα 2) και στη συνέχεια το νευρωνικό δίκτυο που έχουμε κατασκευάσει θα προσπαθήσει να το αναγνωρίσει (Σχήμα 86).



Σχήμα 86: Σάρωση από αριστερά προς τα δεξιά

Στο Σχήμα 86 έχουμε τον αριθμό 2985. Αρχικά ο αλγόριθμος θα συναντήσει το δύο. Αφού εντοπιστεί το συνορό του (Σχήμα 86a) τότε θα κατασκευαστούν 12 συντελεστές Fourier οι οποίοι θα το περιγράφουν. Αυτοί οι συντελεστές θα χρησιμοποιηθούν από το δίκτυο προκυμένου αυτό να αναγνωρίσει σε ποιο ψηφίο αντιστοιχούν.

Στη συνέχεια θα συναντήσει το εννέα (Σχήμα 86b) και θα ακολουθήσει την ίδια διαδικασία αναγνώρισης. Αυτή η διαδικασία θα συνεχιστεί μέχρι να φτάσουμε στο δεξιότερο σημείο της φωτογραφίας όπου και θα τερματίσει ο αλγόριθμος.

Στη συνέχεια δίνονται μερικά παραδείγματα στα οποία εξετάστηκε το δίκτυο και δεν έχουν χρησιμοποιηθεί στην εκπαίδευση του.

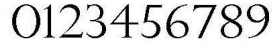
Test:1

Neural network success: 0 1 2 3 4 5 6 7 8 9

Test:2

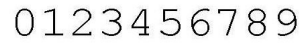
Neural network success: 0 1 2 3 4 5 6 7 8 9

Test:3



Neural network success: 0 1 2 3 4 5 6 7 8 9

Test:4



Neural network success: 0 1 2 3 4 5 6 7 8 9

Test:5

Neural network success: 0 1 2 3 4 5 6 7 8 9

Test:6

Neural network success: 0 0 2 3 4 4 6 7 8 6

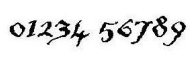
Test:7

Neural network success: 0 1 2 2 2 1 7 7 8 9

Test:8

Neural network success: 1 1 9 -1 5 4 5 5 6 6

Test:9



Neural network success: 0 0 4 6 4 4 8 7 -1 -1

Test:10

Neural network success: 0 5 5 8 7 7 8 7 7 8

Σε αυτό το σημείο θα κάνουμε κάποιες παρατηρήσεις σχετικά με την απόδοση του δικτύου στα πιο πάνω test. Στα πέντε πρώτα tests (0-5) το δίκτυο μπόρεσε να αναγνωρίσει με επιτυχία και τα δέκα ψηφία στις φωτογραφίες. Αυτό οφείλεται στο γεγονός ότι τα πρότυπα εκπαίδευσης και τα νέα δεδομένα είχαν αρκετές ομοιότητες μεταξύ τους τις οποίες μπόρεσε να εντοπίσει το δίκτυο με αποτέλεσμα να δώσει σωστές απαντήσεις. Στα επόμενα πέντε tests (6-10) έχουν χρησιμοποιηθεί πέντε γραφικοί χαρακτήρες οι οποίοι δεν μοιάζουν με κανένα πρότυπο εκπαίδευσης. Εδώ μπορούμε να δούμε ότι η απόδοση του δικτύου είναι πολύ χαμηλή. Στα tests 6 και 7 μπορεί και εντοπίζει περίπου τα μισά από τα δέκα ψηφία ενώ στα tests 8, 9 και 10 η απόδοση είναι ακόμα χειρότερη αφού μπορεί και εντοπίζει ένα δύο ή κανένα ψηφίο. Επίσης στα tests 8 και 9 το δίκτυο δίνει την απάντηση -1. Το -1 δίνεται οταν η έξοδος που παράγεται από το δίκτυο δεν έχει καμία σχέση με αυτή των προτύπων εκπαίδευσης. από τα παραπάνω φαίνεται η δυσκολία κατασκευής ενός δικτύου που θα μπορεί να αναγνωρίζει με επιτυχία ψηφία με οποιοδήποτε γραφικό χαρακτήρα. Το πρόβλημα αναγνώρισης αριθμών θα ηταν πιο απλό αν είχαμε να ασχοληθούμε με λίγους γραφικούς χαρακτήρες όμως η δυσκολία είναι πιο μεγάλη καθώς αυξάνει το πλήθος αυτών. Η προσέγγιση Fourier είναι ένας αποτελεσματικός τρόπος για να χαρακτηρίσουμε μια καμπύλη. Με τη χρήση μόνο θεωρίας Fourier μπορέσαμε να κατασκευάσουμε ένα δίκτυο που μπορεί να αναγνωρίζει αριθμούς με αρκετά καλή επιτυχία. Η χρήση τεχνικών που θα μπορούσαν να δώσουν περισσότερες πληροφορίες για τη μορφή και το σχήμα του κάθε ψηφίου θα ηταν ιδιαίτερα χρήσιμες στην βελτίωση της πιο πάνω προσπάθειας.

Αναφορές

[1] Nixon, Aguado. Feature Extraction and Image Processing.

[2] Baxes, G. A., Digital Image Processing, Principles and Applications, Wiley and Sons, 1994.

[3] Otsu, N., A Threshold Selection Method from Gray-Level Histograms, 1979.

[4] Digital Image Processing, Third Edition, WILLIAM K. PRATT, 2001.

[5] Evans, A. N. and Nixon, M. S., Mode Filtering to Reduce Ultrasound Speckle for Feature Extraction, Image and Signal Processing, 1995.

[6] Campbell, J. D., Edge Structure and the Representation of Pictures, PhD Thesis, University, Missouri, Columbia USA, 1969.

[7] Canny, J., A Computational Approach to Edge Detection, 1986.

[8] Prewitt, J. M. S. and Mendelsohn, M. L., The Analysis of Cell Images, 1966.

[9] Roberts, L. G., Machine Perception of Three-Dimensional Solids, Optical and Electro-Optical Information Processing, MIT Press, 1965.

[10] Bracewell, R. N., The Fourier Transform and its Applications, Revised 2nd Edition, McGraw-Hill Book Co., Singapore, 1986.

[11] Aguado, A. S., Nixon, M. S. and Montiel, E., Parameterising Arbitrary Shapes via Fourier Descriptors for Evidence-Gathering Extraction, CVIU, 1998.

[12] Crimmins, T. R., A Complete Set of Fourier Descriptors for Two-Dimensional Shapes, Man and Cybernetics, 1982.

[13] Kuhl, F. P. and Giardina, C. R., Elliptic Fourier Descriptors of a Closed Contour, 1982.

[14] Persoon, E. and Fu, K.-S., Shape Description Using Fourier Descriptors, 1977.

[15] Chun-Shin Lin ,Chia-Lin Hwang, New Forms of Shape Invariants from Elliptic Fourier Descriptors, 1986.

[16] Howard Demuth and Mark Beale, Neural Network Toolbox, For Use with MATLAB.

[17] R. Rojas:Neural Networks,Springer-Verlag,The Backpropagation Algorithm, Berlin, 1996.

[18] Herve Abdi, Neural Networks, The University of Texas at Dallas.

[19] Βασίλης Πλαγιανάκος, Νέες Μέθοδοι Εκπαίδευσης Τεχνητών Νευρωνικών Δικτύων, Βελτιστοποίησης και Εφαρμογές.

[20] Jorge Nocedal, Stephen J. Wright, Numerical Optimization.

[21] Edwin K.P. Chong and Stanislaw H. Zak, An Introduction to Optimization, Second Edition.