

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ

ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ: Τηλεπικοινωνιών και Τεχνολογίας Πληροφορίας

ΕΡΓΑΣΤΗΡΙΟ: Ενσύρματης Τηλεπικοινωνίας

Διπλωματική Εργασία

του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και
Τεχνολογίας Υπολογιστών της Πολυτεχνικής Σχολής του
Πανεπιστημίου Πατρών

Ανδρέα Παπαθανασίου του Σωτηρίου

Αριθμός Μητρώου: 5238

Θέμα

**«Μελέτη, σχεδιασμός και υλοποίηση αλγορίθμων
εντοπισμού θέσης και αναγνώρισης χαρακτήρων σε
τυπωμένες εικόνες»**

Επιβλέπων

Ευάγγελος Δερματάς

Αριθμός Διπλωματικής Εργασίας:

Πάτρα, Ιούλιος 2010

ΠΙΣΤΟΠΟΙΗΣΗ

Πιστοποιείται ότι η Διπλωματική Εργασία με θέμα

«Μελέτη σχεδιασμός και υλοποίηση αλγορίθμων εντοπισμού θέσης και αναγνώρισης χαρακτήρων σε τυπωμένες εικόνες»

Του φοιτητή του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών

Ανδρέα Παπαθανασίου του Σωτηρίου

Αριθμός Μητρώου: 5238

Παρουσιάστηκε δημόσια και εξετάστηκε στο Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών στις
...../...../.....

Ο Επιβλέπων

Ευάγγελος Δερματάς
Επίκουρος Καθηγητής

Ο Διευθυντής του Τομέα

Νίκος Φακωτάκης
Καθηγητής

Αριθμός Διπλωματικής Εργασίας:

Θέμα: «Μελέτη σχεδιασμός και υλοποίηση αλγορίθμων εντοπισμού θέσης και αναγνώρισης χαρακτήρων σε τυπωμένες εικόνες»

Φοιτητής: Παπαθανασίου Ανδρέας
Επιβλέπων: Δερματάς Ευάγγελος

Περίληψη

Για τη παρούσα εργασία πραγματοποιήσαμε μια αναλυτική περιγραφή της δομής ενός συστήματος οπτικής αναγνώρισης χαρακτήρα και των μεθόδων που έχουν χρησιμοποιηθεί από τους διάφορους ερευνητές. Σταθήκαμε περισσότερο στην θεωρία των Κυματιδίων (Wavelets) και των Τεχνητών Νευρωνικών Δικτύων. Στη συνέχεια υλοποιήσαμε ένα σύστημα Οπτικής Αναγνώρισης Χαρακτήρα σε περιβάλλον Matlab χρησιμοποιώντας wavelets για την εξαγωγή παραμέτρων και Radial Basis Function (RBF) νευρωνικό δίκτυο. Στο πείραμα που πραγματοποιηθήσαμε μετρήσαμε την αποδοτικότητα του συστήματος μας για εξαγωγή παραμέτρων με δύο διαφορετικά wavelets(sym4 και Meyer) και αποδείξαμε πως το δεύτερο έχει πολύ καλύτερη επίδοση.

ΠΕΡΙΕΧΟΜΕΝΑ

Κεφάλαιο 1: Εισαγωγή	1
1.1 Χρήσεις OCR Συστημάτων	1
1.2 Μεθοδολογίες OCR Συστημάτων	1
1.2.1 Τρόπος προσέγγισης προβλήματος	2
1.2.2 Η φύση των εφαρμογών	3
1.2.3 Χαρακτηριστικά που χρησιμοποιούνται	3
Κεφάλαιο 2: Δομή OCR Συστήματος	5
2.1 Προεπεξεργασία	5
2.1.1 Μείωση θορύβου	5
2.1.2 Κανονικοποίηση των Δεδομένων	7
2.1.3 Συμπύεση πληροφοριών	8
2.2 Κατάτμηση	10
2.2.1 Εξωτερική κατάτμηση	10
2.2.2 Εσωτερική κατάτμηση	11
2.3 Εξαγωγή και Ανάλυση χαρακτηριστικών	13
2.3.1 Γενικοί μετασχηματισμοί και ανάπτυξη σειρών	13
2.3.2 Χαρακτηριστικά προερχόμενα από στατιστική κατανομή σημείων	16
2.3.3 Γεωμετρικά και τοπολογικά χαρακτηριστικά	17
2.4 Ταξινόμηση	17
2.4.1 Template matching	18
2.4.2 Στατιστικές Τεχνικές	18
2.4.3 Δομικές Τεχνικές	19
2.4.4 Νευρωνικά Δίκτυα	20
2.5 Μετεπεξεργασία	23
Κεφάλαιο 3: Πείραμα	25
3.1 Διαδικασία	25
3.2 Εξαγωγή παραμέτρων	26
3.3 Αναγνώριση	26
3.4 Πείραμα 1	27
3.4.1 Πείραμα 1.α	27
3.4.2 Πείραμα 1.β	29
3.4.3 Συμπεράσματα	30
3.5 Πείραμα 2	31
3.5.1 πείραμα 2.α	31
3.5.2 πείραμα 2.β	32
3.5.3 Συμπεράσματα	33
3.6 Βελτιώσεις	33
Βιβλιογραφία	34

Κεφάλαιο 1

Εισαγωγή

Το πρόβλημα της οπτικής αναγνώρισης χαρακτηρίζεται ως τομέας μιας προσπάθειας προσομοίωσης της ανθρώπινης όρασης από τον υπολογιστή. Σκοπός είναι ο εντοπισμός, μέσα σε μία εικόνα, των αλφαριθμητικών συμβόλων και η αναγνώριση αυτών. Η προσομοίωση της ανθρώπινης όρασης από μηχανή αποτελεί τομέα έρευνας για πάνω από τρεις δεκαετίες. Πολλές εργασίες και προγράμματα οπτικής αναγνώρισης χαρακτηρίζονται ως πραγματοποιηθεί όλα αυτά τα χρόνια χωρίς όμως να έχει επιτευχθεί ο στόχος, δηλαδή να αποκτήσει η μηχανή τις ίδιες ικανότητες ανάγνωσης με τον άνθρωπο. Έτσι το πεδίο έρευνας είναι ακόμα ανοιχτό.

1.1 Χρήσεις OCR συστημάτων

- Βοηθήματα ανάγνωσης για τυφλούς με χρήση φωτοαισθητήρων και απτικών προσομοιωτών
- Τηλεπικοινωνιακά βοηθήματα για κωφούς
- Χρήση σε ταχυδρομεία για την αναγνώριση διευθύνσεων αυτόματα
- Αναγνώριση πινακίδων αυτοκινήτου από τροχαία ή σε χώρους στάθμευσης
- Άμεση επεξεργασία χειρόγραφων κειμένων
- Μετατροπή μεγάλου όγκου δεδομένων (μηχανογράφηση)
- Ηλεκτρονικές Βιβλιοθήκες
- Αναγνώριση υπογραφής
- Χρήση στην εκδοτική βιομηχανία

1.2 Μεθοδολογίες OCR

Στη βιβλιογραφία έχουν προταθεί πολλές μέθοδοι αναγνώρισης χαρακτηρίζονται και συνδυασμοί αυτών. Διαφορετική είναι η προσέγγιση και άλλες οι μέθοδοι που χρησιμοποιούμε όταν θέλουμε να αναγνωρίσουμε τυπωμένο από μηχανή κείμενο και διαφορετική για χειρόγραφο. Μάλιστα σε γλώσσες με διαφορετικά σύμβολα προτείνονται διαφορετικοί τρόποι αναγνώρισης.

Οι μεθοδολογίες στην οπτική αναγνώριση χαρακτηρίζονται μπορούν να κατηγοριοποιηθούν με διάφορους τρόπους όπως:

- 1 Τρόπος προσέγγισης προβλήματος

- 2 Η φύση των εφαρμογών
- 3 Χαρακτηριστικά που χρησιμοποιούνται

1.2.1 Προσεγγίσεις αναγνώρισης χαρακτήρα

Δύο είναι οι κύριες προσεγγίσεις στην αναγνώριση προτύπων:

Η **στατιστική/θεωρία αποφάσεων** και η **συντακτική/γλωσσική/γραμματική/δομική** προσέγγιση. Κάθε μια έχει τα πλεονεκτήματα και τα μειονεκτήματά της. π.χ. Στατιστικές μέθοδοι δεν μπορούν να χειριστούν καλά δομικές πληροφορίες που αφορούν τις διασυνδέσεις ενός περίπλοκου προτύπου, αλλά από την άλλη η χρήση αυστηρών συντακτικών κανόνων πολλές φορές εμποδίζει την διαδικασία. Επίσης η γλωσσική προσέγγιση δεν λαμβάνει υπόψη τους περιορισμούς που θέτει η εξαγωγή χαρακτηριστικών.

Συνεπώς ένα υβριδικό μοντέλο θα ήταν η καλύτερη λύση. Για την αναγνώριση χαρακτήρα χρειαζόμαστε τις τεχνικές για να περιγράψουμε έναν μεγάλο αριθμό παρόμοιων δομών της ίδιας κατηγορίας επιτρέποντας τις ευδιάκριτες περιγραφές μεταξύ των διαφορετικών προτύπων. Ο τελικός στόχος της έρευνας αναγνώρισης χαρακτήρα είναι να αναπτυχθούν μηχανές που μπορούν να διαβάσουν οποιοδήποτε κείμενο με την ίδια ικανότητα αναγνώρισης του ανθρώπου. Η προσδοκία είναι ότι, εάν η λειτουργία της ανθρώπινης όρασης περιγραφεί κατάλληλα και χρησιμοποιηθεί σε έναν αλγόριθμο αναγνώρισης χαρακτήρα, ο αλγόριθμος πρέπει να λειτουργήσει όπως ένας άνθρωπος. Αυτό είναι το κίνητρο για την αποκαλούμενη περιγραφική προσέγγιση, η οποία είναι τώρα δημοφιλής στην έρευνα αναγνώρισης χαρακτήρα.

Σε μια περιγραφική προσέγγιση μπορούν να παρασχεθούν εύκολα οι ευελιξίες που απαιτούνται για να φροντίσουν την άπειρη παραλλαγή των μορφών χαρακτήρα σε μια κατηγορία-περιγραφή. Η περιγραφή αντιπροσωπεύει ένα πιο υψηλό επίπεδο νοημοσύνης. Η περιγραφή ενός χαρακτήρα περιλαμβάνει τα χαρακτηριστικά γνωρίσματα (δομικές λεπτομέρειες) και τους κανόνες κάτω από τους οποίους συντίθεται ένας χαρακτήρας. Για να πετύχουμε μια αποδοτική περιγραφή, τα χρησιμοποιούμενα χαρακτηριστικά γνωρίσματα πρέπει να είναι ανεξάρτητα. Δηλαδή η παρουσία νέου χαρακτηριστικού γνωρίσματος ή η απουσία παλαιού χαρακτηριστικού γνωρίσματος, δεν πρέπει να έχει επιπτώσεις στην περιγραφή των υπόλοιπων χαρακτηριστικών γνωρισμάτων. Αυτό θα παράσχει κάποια ασυλία στους περιορισμούς εκ μέρους των εξαγωγέων χαρακτηριστικών γνωρισμάτων.

1.2.2 Φύση των εφαρμογών

Όσον αφορά αυτή τη κατηγοριοποίηση, μπορούμε να διακρίνουμε δύο μεγάλες κατηγορίες: Την **off-line** και την **on-line** αναγνώριση χαρακτήρα ανάλογα με το αν η αναγνώριση γίνεται σε πραγματικό χρόνο ή όχι.

Υποκατηγορίες τη off-line αναγνώρισης είναι οι:

1. Fixed-font αναγνώριση (μίας γραμματοσειράς)
2. Multi-font αναγνώριση (ποικίλες τυπωμένες γραμματοσειρές)
3. Omni-font αναγνώριση (οποιαδήποτε τυπωμένη γραμματοσειρά)
4. Handwritten αναγνώριση, η οποία ασχολείται με την αναγνώριση χειρόγραφου απομονωμένου χαρακτήρα
5. Script αναγνώριση η οποία ασχολείται με τη αναγνώριση χειρόγραφου κειμένου με ενωμένους ή μη χαρακτήρες

1.2.3 Χαρακτηριστικά που χρησιμοποιούνται

Όσον αφορά τα χαρακτηριστικά, οι τεχνικές που χρησιμοποιούνται μπορούν να χωριστούν σε δυο κατηγορίες:

1. Εύρεση και ταύτιση με υπόδειγμα
2. Εύρεση και ταύτιση χαρακτηριστικού

- **Εύρεση και ταύτιση με υπόδειγμα**

Αυτή η μέθοδος κάνει σύγκριση του χαρακτήρα που εισάγουμε με μία σειρά από αποθηκευμένα στο σύστημα πρωτότυπα. Το πρωτότυπο με το μεγαλύτερο δείκτη ομοιότητας αναγνωρίζεται ως ο εισηγμένος χαρακτήρας. Οι μέθοδοι σύγκρισης διαφέρουν. Μπορούν να είναι απλοί(έναν προς ένα χαρακτήρας) ή σύνθετοι (δένδρο απόφασης).Αυτή η τεχνική παρουσιάζει προβλήματα όταν υπάρχει θόρυβος και δεν προσαρμόζεται εύκολα σε αλλαγές του γραφικού χαρακτήρα. Επίσης έχει απορριφθεί ως αλγόριθμος προσομοίωσης στην ανθρώπινη όραση.

- **Εύρεση και ταύτιση χαρακτηριστικού**

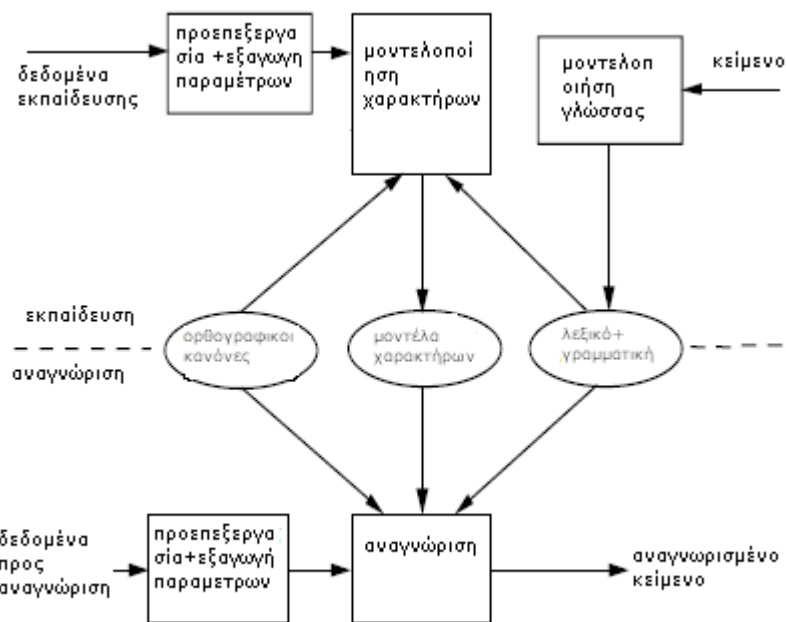
Οι τεχνικές αυτές βασίζονται στο ταίριασμα επιπέδων ή χώρων χαρακτηριστικών οι οποίοι διανέμονται σε ένα δισδιάστατο επίπεδο. Σε αυτές τις μεθόδους, χαρακτηριστικά γνωρίσματα εξάγονται από κάθε χαρακτήρα και συγκρίνονται με τις περιγραφές των ιδανικών χαρακτήρων(προτύπων). Έτσι η περιγραφή που πλησιάζει περισσότερο σε κάποιο πρότυπο μας δίνει αναγνώριση αυτού. Η λειτουργία της ανθρώπινης όρασης έτσι προσομοιάζεται καλύτερα από ότι με την ταύτιση με υπόδειγμα. Πολλές τεχνικές ανάλυσης χαρακτηριστικών έχουν αναπτυχθεί και έχουν εφαρμοστεί στην αναγνώριση χαρακτήρα. Οι περισσότερες από τις οποίες είναι παραδείγματα των παραδοσιακών μεθόδων αναγνώρισης

προτύπων, και είναι συνήθως κατάλληλες για εφαρμογή σε περιορισμένους τομείς.

Κεφάλαιο 2

Δομή συστήματος αναγνώρισης χαρακτήρα

Η δομή ενός προγράμματος αναγνώρισης χαρακτήρα(OCR) χωρίζεται στα εξής βήματα : ανάλυση δομής κειμένου, προεπεξεργασία, κατάτμηση, ταξινόμηση, μετεπεξεργασία.



2.1 Προεπεξεργασία

Προεπεξεργασία είναι το στάδιο κατά το οποίο πραγματοποιούμε διάφορες διεργασίες στη πρωτότυπη εικόνα κειμένου ώστε αυτή να είναι πιο καθαρή και πιο εύκολα επεξεργάσιμη στα επόμενα στάδια αναγνώρισης. Οι κύριοι στόχοι της προεπεξεργασίας είναι :

1. Μείωση θορύβου
2. Κανονικοποίηση των δεδομένων
3. Συμπίεση του πλήθους πληροφοριών που διατηρούνται

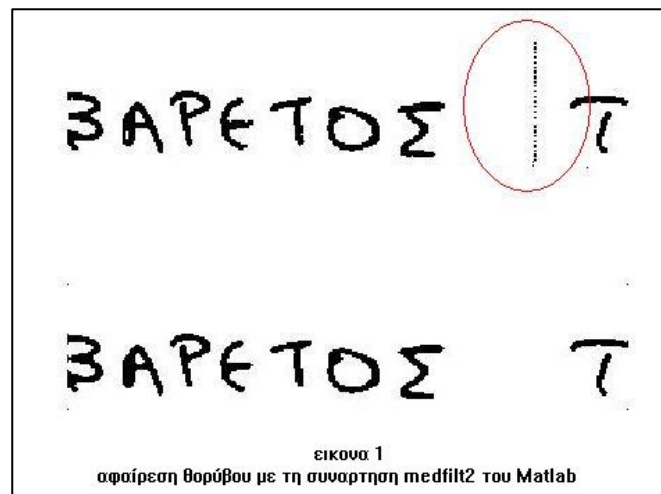
2.1.1 Μείωση θορύβου

Ο θόρυβος μπορεί να προέρχεται από τη μηχανή σάρωσης ή από το μέσο που χρησιμοποιούμε για να γράψουμε και έχει ως αποτέλεσμα οι χαρακτήρες να είναι ελαττωματικοί π.χ. διακεκομμένες γραμμές. Οπότε είναι απαραίτητο να ελαχιστοποιήσουμε τον θόρυβο πριν περάσουμε στην αναγνώριση χαρακτήρα. Οι

τεχνικές μείωσης θορύβου είναι εκατοντάδες αλλά μπορούν να κατηγοριοποιηθούν σε τρεις μεγάλες κατηγορίες:

- **Φιλτράρισμα**

Σκοπός του φιλτραρίσματος είναι η απομάκρυνση θορύβου και η μείωση των πλαστών σημείων που συνήθως εισάγονται από το μέσο γραφής ή από κακή συχνότητα δειγματοληψίας κατά τη σάρωση. Διάφορα φίλτρα χώρου και συχνότητας μπορούν να σχεδιαστούν για αυτό το σκοπό. Η βασική ιδέα είναι η εφαρμογή μίας προσχεδιασμένης μάσκας στην εικόνα. Η μάσκα θα αναθέτει τιμή στα πίξελ σύμφωνα με κάποια συνάρτηση που λαμβάνει υπόψη τις τιμές των γειτονικών πιξελ. Τα φίλτρα μπορούν να χρησιμοποιηθούν ακόμα για εξομάλυνση εικόνας, αφαίρεση φόντο, ρύθμιση αντίθεσης κτλ.



- **Μορφολογική επεξεργασία**

Στη μορφολογική επεξεργασία προσπαθούμε να φιλτράρουμε το κείμενο με logical operations. Διάφοροι τρόποι μορφολογικής επεξεργασίας μπορούν να σχεδιαστούν ώστε να ενώνουν σπασμένες γραμμές, να διαχωρίζουν ενωμένους χαρακτήρες, να εξομαλύνουν το κείμενο. Γι' αυτό το λόγο η χρήση μορφολογικής επεξεργασίας για την αφαίρεση θορύβου που παράγεται λόγω κακής ποιότητας χαρτιού ή ασταθούς κίνησης χεριού είναι αρκετά διαδεδομένη.

- **Μοντελοποίηση θορύβου**

Ο θόρυβος μπορεί να αφαιρεθεί και με κάποιες τεχνικές βαθμονόμησης εφόσον υπάρχει κάποιο μοντέλο κατάλληλο. Συνήθως όμως η μοντελοποίηση του θορύβου είναι αδύνατη. Γενικώς δεν υπάρχει μεγάλη πρόοδος στη μοντελοποίηση θορύβου που εισάγεται από οπτική παραμόρφωση όπως σκιές και θάμπωση. Παρ' όλα αυτά μπορούμε να αξιολογήσουμε την ποιότητα ενός εγγράφου και

μέχρι ενός σημείου να αφαιρέσουμε τον θόρυβο χρησιμοποιώντας τη μοντελοποίηση.

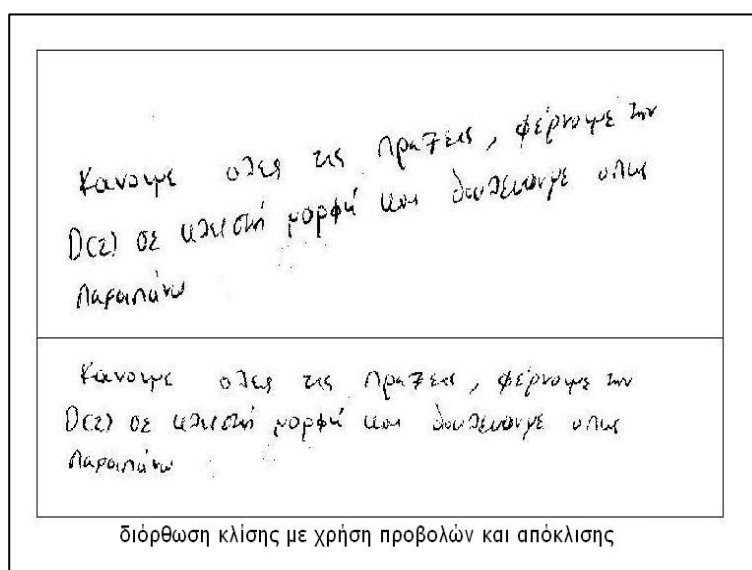
2.1.2 Κανονικοποίηση

Οι μέθοδοι κανονικοποίησης επιχειρούν να εξαλείψουν τις διαφοροποιήσεις που προκύπτουν από διαφορετικούς τρόπους γραφής ώστε να είναι εύκολη η αποκόμιση δεδομένων. Στη συνέχεια παρατίθενται βασικές μέθοδοι κανονικοποίησης:

- **Κανονικοποίηση κλίσης και εξαγωγή γραμμής στοίχισης**

Λόγο σφαλμάτων κατά τη σάρωση ή κατά τη συγγραφή ενός κειμένου, υπάρχει περίπτωση αυτό να πάρει κάποια κλίση ή να είναι καμπύλο. Αυτό έχει αρνητική επίδραση στην αποτελεσματικότητα των αλγορίθμων αναγνώρισης χαρακτήρα που θα επακολουθήσουν. Για αυτό το λόγο αυτές οι ατέλειες θα πρέπει να διορθωθούν. Επίσης κάποιοι χαρακτήρες (όπως το 9 και το g) διαφοροποιούνται και αναγνωρίζονται λαμβάνοντας υπόψη τη θέση τους σε σχέση με τη στοίχιση.

Οι μέθοδοι εξαγωγής της γραμμής στοίχισης περιλαμβάνουν τεχνικές όπως η συστοίχιση πλησιέστερων γειτόνων, η αλληλοσυσχέτιση μεταξύ γραμμών και η χρήση του μετασχηματισμού Hough. Ο μετασχηματισμός Hough εντοπίζει καμπύλες εκμεταλλευόμενος τη διπλότητα μεταξύ σημείων σε μία καμπύλη και τις παραμέτρους της καμπύλης αυτής και χρησιμοποιείται για να χαρακτηρίσει τις παραμετρικές καμπύλες της γραφής.



- **Κανονικοποίηση Λοξότητας**

Ένας ακόμα τρόπος να κανονικοποιήσουμε σε τυπική μορφή τα διάφορα στυλ γραφής είναι ο υπολογισμός της λοξότητας των χαρακτήρων, δηλαδή της γωνίας μεταξύ της πιο μακριάς γραμμής κάποιου χαρακτήρα και της νοητής καθέτου. Η πιο διαδεδομένη μέθοδος εκτίμησης της λοξότητας είναι ο υπολογισμός της μέσης τιμής της γωνίας των σχεδόν κάθετων στοιχείων.

- **Κανονικοποίηση μεγέθους**

Σκοπός της κανονικοποίησης μεγέθους είναι η προσαρμογή των χαρακτήρων σε ένα καθορισμένο μέγεθος. Αυτό μπορεί να πραγματοποιείται κάθετα, οριζόντια ή και τα δύο. Πέρα από το στάδιο της προεπεξεργασίας, η κανονικοποίηση μεγέθους μπορεί να πραγματοποιηθεί και κατά το στάδιο της εκπαίδευσης όπου οι παράμετροι μεγέθους υπολογίζονται ξεχωριστά για κάθε περίπτωση.

- **Εξομάλυνση περιγράμματος**

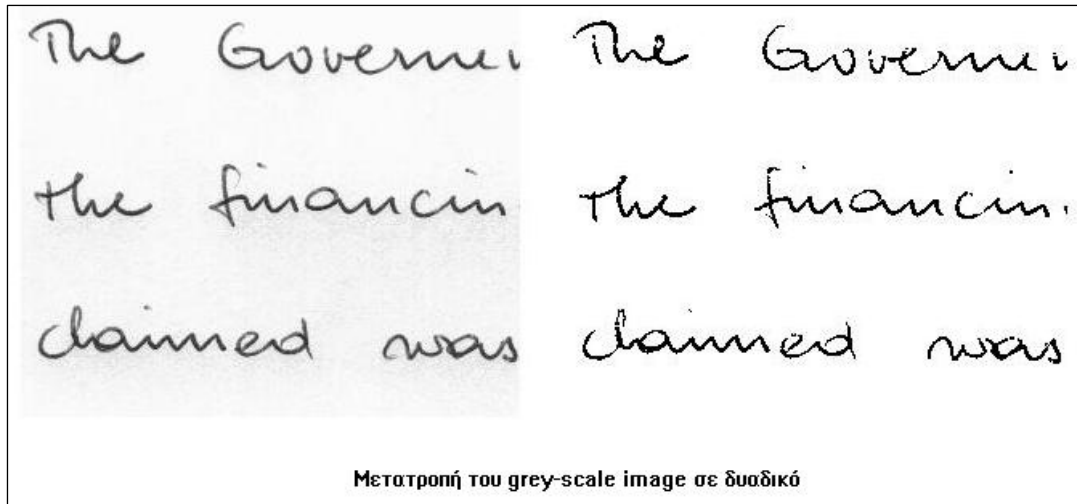
Απομακρύνει ατέλειες από το περίγραμμα των χαρακτήρων οι οποίες προκύπτουν από την αστάθεια του ανθρώπινου χεριού όταν γράφει. Γενικά, μειώνει τον αριθμό των δειγμάτων που είναι απαραίτητα για την αναπαράσταση του κειμένου, οπότε βελτιώνεται η αποτελεσματικότητα στα υπόλοιπα στάδια προεπεξεργασίας.

2.1.3 Συμπίεση

Όπως γνωρίζουμε στη οι κλασσικές μέθοδοι συμπίεση εικόνας μετασχηματίζουν την εικόνα από το πεδίο του χώρου σε άλλα πεδία τα οποία δεν είναι κατάλληλα για αναγνώριση. Οι μέθοδοι συμπίεσης για αναγνώριση χαρακτήρα απαιτούν τεχνικές οι οποίες διατηρούν το πεδίο του χώρου ώστε να διατηρούνται οι πληροφορίες που αφορούν το σχήμα του χαρακτήρα. Οι πιο διαδεδομένες μέθοδοι συμπίεσης είναι η χρήση κατωφλίου και η λέπτυνση.

- **Χρήση κατωφλίου**

Πολλές φορές για να εξοικονομήσουμε χώρο αποθήκευσης δεδομένων και να επιταχύνουμε την επεξεργασία των δεδομένων επιλέγουμε να μετασχηματίσουμε έγχρωμες (ή σε κλίμακα του γκρι) εικόνες σε δυαδικές εικόνες διαλέγοντας μία τιμή κατωφλίου. Οι επιλογές μας να το κάνουμε αυτό είναι δύο: Γενικό κατώφλι και τοπικό. Χρησιμοποιώντας γενικό κατώφλι επιλέγουμε μια τιμή κατωφλίου για ολόκληρη την εικόνα κειμένου, συνήθως υπολογίζοντας την τιμή αυτή από το ιστόγραμμα έντασης της εικόνας. Στη χρήση τοπικών κατωφλίων χρησιμοποιούμε διαφορετική τιμή κατωφλίου για κάθε πιξελ λαμβάνοντας υπόψη πληροφορίες από το τοπικό περιβάλλον.



- **Λέπτυνση**

Αυτή η μέθοδος καταφέρνει πολύ σημαντική μείωση δεδομένων. Λειτουργία της είναι να εξάγει τις πληροφορίες που αφορούν το σχήμα του χαρακτήρα.. Οι δύο βασικές προσεγγίσεις λέπτυνσης είναι η pixel wise και η non-pixel wise. Στη πρώτη περίπτωση η μέθοδος επεξεργάζεται την εικόνα τοπικά και επαναληπτικά έως ότου το πλάτος του σκελετού γίνει μόλις ένα pixel.Αυτός ο τρόπος



είναι πολύ ευαίσθητος στο θόρυβο και μπορεί να παραμορφώσει το σχήμα των χαρακτήρων. Στη δεύτερη περίπτωση η μέθοδος χρησιμοποιεί κάποιες γενικές πληροφορίες που αφορούν τον χαρακτήρα κατά τη διάρκεια της λέπτυνσης και δημιουργεί το σκελετό χωρίς να εξετάζει κάθε pixel ξεχωριστά. Π.χ. μέθοδοι βασισμένοι σε συστοιχίες(clusters) αναγνωρίζουν το κέντρο αυτών και δημιουργούν έτσι το σκελετό του χαρακτήρα. Κάποιοι άλλοι αναγνωρίζουν τα μοναδικά χαρακτηριστικά κάθε χαρακτήρα όπως βρόγχους, διασταυρώσεις γραμμών κτλ. Αυτά τα σημεία προκαλούν συνήθως τα μεγαλύτερα προβλήματα και αντιμετωπίζονται με γενικούς κανόνες.

Οι απαραίτητες επαναλήψεις για τη διαδικασία της λέπτυνσης γίνονται είτε με αλληλοδιαδοχικούς, είτε με παράλληλους αλγορίθμους. Οι διαδοχικοί αλγόριθμοι εξετάζουν το περίγραμμα με σάρωση κοκκιδοπλαισίου(raster scanning) ή με παρακολούθηση του περιγράμματος. Οι παράλληλοι αλγόριθμοι είναι προφανώς ανώτεροι των αλληλοδιαδοχικών από τη στιγμή που εξετάζουν όλα τα pixel

ταυτόχρονα, λαμβάνοντας υπόψη για όλα τις ίδιες συνθήκες για την απάλειψη ή συντήρησή τους. Πλεονέκτημά τους είναι ότι μπορούν να εφαρμοστούν αποδοτικά σε παράλληλο hardware. Οι τεχνικές προεπεξεργασίας εξερευνούνται καλά και εφαρμόζονται σε πολλούς άλλους τομείς της επεξεργασίας εικόνας εκτός από την αναγνώριση χαρακτήρα. Πρέπει να σημειώσουμε ότι, οι ανωτέρω τεχνικές έχουν επιπτώσεις στα στοιχεία και μπορούν να εισαγάγουν τις απροσδόκητες διαστρεβλώσεις στην εικόνα εγγράφων. Κατά συνέπεια, αυτές οι τεχνικές μπορούν να προκαλέσουν την απώλεια σημαντικών πληροφοριών και πρέπει να εφαρμοστούν με την προσοχή.

2.2 Κατάτμηση

Αφού λοιπόν έχουμε λάβει αρκετές πληροφορίες , έχουμε συμπιέσει την εικόνα και έχουμε απομακρύνει κάποιο θόρυβο κατά τη διαδικασία της προεπεξεργασίας περνάμε στο στάδιο κατάτμησης του κειμένου. Κατά την κατάτμηση γίνεται ο διαχωρισμός του κειμένου σε γράμματα, λέξεις και γραμμές. Συνεπώς η διαδικασία αυτή είναι σημαντική διότι επηρεάζει άμεσα την ποιότητα και την επιτυχία της αναγνώρισης. Οι δύο τύποι κατάτμησης είναι: Εξωτερική και εσωτερική κατάτμηση

2.2.1 Εξωτερική κατάτμηση είναι η απομόνωση των τμημάτων κειμένου, όπως η παράγραφος, η πρόταση ή η λέξη.

Στόχος της εξωτερικής κατάτμησης είναι η ανάλυση του κειμένου στα λογικά του μέρη, πράξη που αποτελεί το σημαντικότερο κομμάτι της ανάλυσης κειμένου, η οποία είναι απαραίτητη για τη πραγματοποίηση της off line αναγνώρισης χαρακτήρα. Παρότι η ανάλυση κειμένου είναι σχετικά ξεχωριστό πεδίο έρευνας με τις δικές της μεθοδολογίες και τεχνικές, η κατάτμηση του κειμένου σε τομείς κειμένου και τομείς ανυπαρξίας κειμένου είναι ένα ολοκληρωμένο τμήμα ενός προγράμματος αναγνώρισης χαρακτήρα. Συνεπώς κάποιος που ασχολείται με το πεδίο αναγνώρισης χαρακτήρα θα πρέπει να έχει μια γενική εικόνα των τεχνικών ανάλυσης κειμένου.

Το στάδιο της ανάλυσης διάταξης κειμένου παίζει κρίσιμο ρόλο στη διαδικασία αναγνώρισης διότι όσο πιο ακριβής είναι ,τόσο λιγότερα λάθη θα προκύψουν στα μετέπειτα στάδια. Η ανάλυση διάταξης σελίδας πραγματοποιείται σε δύο στάδια:

- **Δομική Ανάλυση**

Η **δομική ανάλυση** η οποία είναι υπεύθυνη για την κατάτμηση της εικόνας σε τμήματα κειμένου(παράγραφος, σειρά, λέξη κτλ). Στο σκέλος της δομικής

ανάλυσης του κειμένου ο εκάστοτε αλγόριθμος αντιστοιχεί μια λογική ιδιότητα σε κάθε λέξη ή γενικότερα κομμάτι του κειμένου. Δημιουργεί δηλαδή ενός είδος συντακτικού. Η αναπαράσταση συνήθως γίνεται με μορφή λογικού δένδρου ντετερμινιστικά χρησιμοποιώντας κάποιου είδους κανόνες.

- **Λειτουργική Ανάλυση**

Η **λειτουργική ανάλυση** χρησιμοποιεί διαφόρους κανόνες που αφορούν το μέγεθος, τη τοποθεσία και τη διάταξη του κειμένου ώστε να αποδώσει σε αυτά τα τμήματα κάποιο ρόλο όπως τίτλος, περίληψη κτλ. Για την λειτουργική ανάλυση έχουν προταθεί τριών ειδών αλγόριθμοι : Από τα πάνω προς τα κάτω προσέγγιση, από τα κάτω προς τα πάνω και μίξη αυτών.

Διάφορες προσεγγίσεις θεωρούν μια ομοιογενή περιοχή σε μια εικόνα εγγράφων ως κατασκευασμένη περιοχή. Η κατάτμηση σελίδων εφαρμόζεται έπειτα με την εύρεση των κατασκευασμένων περιοχών στις εικόνες γκριζας κλίμακας ή χρώματος. Πολλές προσεγγίσεις για την κατάτμηση σελίδων επικεντρώνονται στην επεξεργασία των pixel υποβάθρου ή τη χρησιμοποίηση του άσπρου διαστήματος σε μια σελίδα για να προσδιορίσουν τις ομοιογενείς περιοχές. Αυτές οι τεχνικές περιλαμβάνουν το X-Y δέντρο , pixel based σχεδιάγραμμα προβολής , σχεδιάγραμμα προβολής βασισμένο στο συνδεδεμένο συστατικό, εντοπισμό του λευκού διαστήματος , και την λέπτυνση λευκού διαστήματος . Μπορούν να θεωρηθούν ως από επάνω προς τα κάτω προσεγγίσεις, οι οποίες τέμνουν μια σελίδα, κατ' επανάληψη, ξεκινώντας από τα μεγάλα συστατικά, συνεχίζοντας στα μικρά συστατικά και φθάνοντας τελικά στους μεμονωμένους χαρακτήρες.

Αφ' ετέρου, υπάρχουν μερικές από τα κάτω προς τα επάνω μέθοδοι βασισμένες στην επεξεργασία των pixel και τα συνδεδεμένα συστατικά, που κατ' επανάληψη αναπτύσσουν τις ομοιογενείς περιοχές από τα μικρά συστατικά. Ένα παράδειγμα αυτής της προσέγγισης μπορεί να είναι μέθοδος Docstrum, η οποία χρησιμοποιεί την συγκέντρωση πλησιέστερων γειτόνων . Μερικές τεχνικές συνδυάζουν και τις από επάνω προς τα κάτω και από κάτω προς τα επάνω τεχνικές.

2.2.2 Εσωτερική κατάτμηση είναι η απομόνωση των γραμμάτων, ειδικά στις ρευστά γραπτές λέξεις.

Η εσωτερική κατάτμηση είναι μια λειτουργία που επιδιώκει να αποσυνθέσει μια εικόνα μιας ακολουθίας χαρακτήρων σε μεμονωμένα σύμβολα. Αν και, οι μέθοδοι έχουν αναπτυχθεί εντυπωσιακά στην τελευταία δεκαετία και ποικίλες τεχνικές έχουν

προκύψει, η κατάτμηση του ρέοντος κειμένου σε γράμματα είναι ακόμα ένα άλτο πρόβλημα. Οι στρατηγικές κατάτμησης χαρακτήρα διαιρούνται σε τρεις κατηγορίες .

- **Ρητή κατάτμηση**

Σε αυτήν την στρατηγική, τα τμήματα προσδιορίζονται με βάση χαρακτηριστικές ιδιότητες .Η διαδικασία που κατακερματίζει την εικόνα σε σημαντικά συστατικά ονομάζεται ανατομή. Η ανατομή είναι μια διαδικασία που αναλύει μια εικόνα χωρίς χρησιμοποίηση συγκεκριμένης κλάσης των πληροφοριών σχήματος. Το κριτήριο για την καλή κατάτμηση είναι η συμφωνία των γενικών ιδιοτήτων των τμημάτων με εκείνες που αναμένονται για τους έγκυρους χαρακτήρες. Οι διαθέσιμες μέθοδοι βασισμένες στην ανατομή μίας εικόνας χρησιμοποιούν το κενό διάστημα και την τονικότητα, κάθετη ανάλυση προβολής, ανάλυση συνδεδεμένων συστατικών και τα ορόσημα . Επιπλέον, η ρητή κατάτμηση μπορεί να υποβληθεί στην αξιολόγηση χρησιμοποιώντας γλωσσικό πλαίσιο .

- **Άδηλη κατάτμηση**

Αυτή η στρατηγική κατάτμησης είναι βασισμένη στην αναγνώριση. Ψάχνει την εικόνα για τα συστατικά που ταιριάζουν με τις προκαθορισμένες κλάσεις. Η κατάτμηση εκτελείται με την χρήση της εμπιστοσύνης αναγνώρισης, συμπεριλαμβανομένης της συντακτικής ή σημασιολογικής ακρίβειας του γενικού αποτελέσματος. Σε αυτήν την προσέγγιση, δύο κλάσεις μεθόδων μπορούν να χρησιμοποιηθούν: μέθοδοι μερικής αναζήτησης και μεθόδους που τέμνουν την εικόνα βάσει αντιπροσωπευτικών χαρακτηριστικών γνωρισμάτων.

Η πρώτη κλάση προσπαθεί να τέμνει τις λέξεις σε γράμματα ή άλλες μονάδες χωρίς τη χρήση βασισμένων στα χαρακτηριστικά γνωρίσματα αλγορίθμων ανατομής. Μάλλον, η εικόνα διαιρείται συστηματικά σε πολλά επικαλυπτόμενα κομμάτια αδιαφορώντας για το περιεχόμενο. Σαν σύλληψη , αυτές οι μέθοδοι προέρχονται από τα σχέδια που αναπτύσσονται για την αναγνώριση των machine-printed λέξεων. Η βασική αρχή είναι να χρησιμοποιηθεί ένα κινητό παράθυρο του μεταβλητού πλάτους για να παρέχει τις ακολουθίες δοκιμαστικών κατατμήσεων, οι οποίες επιβεβαιώνονται από την αναγνώριση χαρακτήρα. Μια άλλη τεχνική συνδυάζει το δυναμικό προγραμματισμό και τα νευρικά δίκτυα . Τέλος, η μέθοδος εκλεκτικής προσοχής προχωρεί τα νευρικά δίκτυα ακόμα περαιτέρω στο χειρισμό του προβλήματος κατάτμησης.

Η δεύτερη κλάση των μεθόδων τέμνει την εικόνα σιωπηρά χρησιμοποιώντας την ταξινόμηση των υποσυνόλων των χωρικών χαρακτηριστικών γνωρισμάτων που συλλέγονται από την εικόνα συνολικά. Αυτή η προσέγγιση μπορεί να διαιρεθεί σε

δύο κατηγορίες: Τις προσεγγίσεις που βασίζονται στο κρυμμένο markov μοντέλο και μη βασισμένες στο markov προσεγγίσεις.

- **Μικτές στρατηγικές**

Συνδυάζουν τη ρητή και την άδηλη κατάτμηση με έναν υβριδικό τρόπο. Ένας αλγόριθμος ανατομής εφαρμόζεται στην εικόνα, αλλά ο σκοπός είναι να κόψει την εικόνα σε αρκετά πολλά σημεία ώστε τα σωστά όρια κατάτμησης να συμπεριλαμβάνονται μεταξύ των αποκοπών που εφαρμόζονται. Μόλις βεβαιωθεί αυτό, η βέλτιστη κατάτμηση επιδιώκεται από την αξιολόγηση των υποσυνόλων των αποκοπών που εφαρμόζονται. Κάθε υποσύνολο υπονοεί μια υπόθεση κατάτμησης, και η ταξινόμηση εφαρμόζεται για να αξιολογήσει τη διαφορετική υπόθεση και να επιλέξει την πιο ελπιδοφόρο κατάτμηση.

2.3 Εξαγωγή και Ανάλυση χαρακτηριστικών

Η εξαγωγή χαρακτηριστικών γνωρισμάτων αφορά την ανάκτηση των ιδιοτήτων καθορισμού. Είδη χαρακτηριστικών γνωρισμάτων που χρησιμοποιούνται είναι ο αριθμός τρυπών στο χαρακτήρα ή ο αριθμός κοιλωμάτων στο εξωτερικό περίγραμμα του, η σχετική προεξοχή των άκρων του χαρακτήρα, και τα τοπικά χαρακτηριστικά γνωρίσματα, όπως οι σχετικές θέσεις των διασταυρώσεων, γραμμών και των γωνιών. Ανάλογα με το τύπο της τεχνικής μας για εξαγωγή χαρακτηριστικών, ομαδοποιούμε τις τεχνικές ανάλυσης χαρακτηριστικών ως:

1. Γενικοί μετασχηματισμοί και ανάπτυξη σειρών
2. Χαρακτηριστικά που προέρχονται από τη στατιστική κατανομή των σημείων
3. Γεωμετρικά και τοπολογικά χαρακτηριστικά

2.3.1 Γενικοί μετασχηματισμοί και ανάπτυξη σειρών

Μετασχηματισμός ενός σήματος είναι η αναπαράσταση του με κάποιο διαφορετικό τρόπο χωρίς όμως να αλλάξουμε την πληροφορία που μεταφέρει.

Η τεχνική γενικών μετασχηματισμών και ανάπτυξης σειρών βοηθάει στη μείωση των διαστάσεων του διανύσματος χαρακτηριστικών και μας δίνει χαρακτηριστικά ανεξάρτητα σε μερικές παραμορφώσεις. Ερευνητές έχουν χρησιμοποιήσει μετασχηματισμούς όπως του Fourier, Walsh, Hadamard και άλλους. Τα χαρακτηριστικά εξάγονται εύκολα με αυτές τις τεχνικές αλλά απαιτούν μεγάλη υπολογιστική ισχύ:

- **Μετασχηματισμοί Fourier**

Η γενική διαδικασία είναι να επιλεχτεί το φάσμα πλάτους του διανύσματος μέτρησης ως χαρακτηριστικό γνώρισμα σε ένα n -διάστατο ευκλείδειο χώρο. Μια από τις ελκυστικότερες ιδιότητες του μετασχηματισμού Φουριέ είναι η δυνατότητα να αναγνωριστούν οι θέση μετατοπισμένων χαρακτήρων, όταν παρατηρηθεί το φάσμα πλάτους και αγνοηθεί η φάση. Οι μετασχηματισμοί Fourier έχουν εφαρμοστεί στην αναγνώριση χαρακτήρα με πολλούς τρόπους.

- **Μετασχηματισμοί Wavelet**

Ένα wavelet είναι μια κυματοειδής ταλάντωση $\psi \in L^2(\mathbb{R})$ με ένα εύρος που αρχίζει από μηδέν, αυξάνει, και έπειτα μειώνεται πίσω σε μηδέν, δηλαδή ένα σήμα ενέργειας.

Ο μετασχηματισμός Wavelet είναι μια τεχνική ανάπτυξης σειράς που μας επιτρέπει να απεικονίσουμε το σήμα σε διαφορετικά επίπεδα ευκρίνειας, δηλαδή να αναλύσουμε διαφορετικές συχνότητες με διαφορετικές κλίμακες.

Αυτή η διαδικασία ονομάζεται πολυδιακριτική ανάλυση. Πιο συγκεκριμένα: Η πολυδιακριτική ανάλυση ενός χώρου $L^2(\mathbb{R})$ αποτελείται από μία ακολουθία υποχώρων

$$\{0\} \cdots \subset V_0 \subset V_1 \subset \cdots \subset V_n \subset V_{n+1} \subset \cdots \subset L^2(\mathbb{R})$$

οι οποίοι ικανοποιούν τις εξής απαιτήσεις:

$$\forall (j, k) \in \mathbb{Z}^2, f(t) \in V_j \Leftrightarrow f(t - 2^j k) \in V_j,$$

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j,$$

$$\forall j \in \mathbb{Z}, f(t) \in V_j \Leftrightarrow f\left(\frac{t}{2}\right) \in V_{j+1},$$

$$\lim_{j \rightarrow +\infty} V_j = \bigcap_{j=-\infty}^{+\infty} V_j = \{0\},$$

$$\lim_{j \rightarrow -\infty} V_j = \text{Closure} \left(\bigcup_{j=-\infty}^{+\infty} V_j \right) = L^2(\mathbb{R}),$$

και υπάρχει θ τέτοιο ώστε $\{\theta(t - n)\}_{n \in \mathbb{Z}}$ να είναι βάση Riesz του V_0 . (Mallat)

Δηλαδή:

1. Κάθε υποχώρος παραμένει αναλλοίωτος όταν μετατεθεί κατά κάποιο πολλαπλάσιο του 2^{-k}
2. Για υποχώρους $V_k \subset V_l, k < l$, ισχύει ότι ο ένας είναι κλίμακα του άλλου κατά ένα συντελεστή καθυστέρησης ίσο με 2^{k-l}
3. Ο υποχώρος V_0 πρέπει να δημιουργείται ως το γραμμικό περίβλημα (linear hull) μίας η περισσότερων συναρτήσεων που ονομάζονται συναρτήσεις κλίμακας.
4. Οι υποχώροι καλύπτουν ολόκληρο το χώρο L^2 .

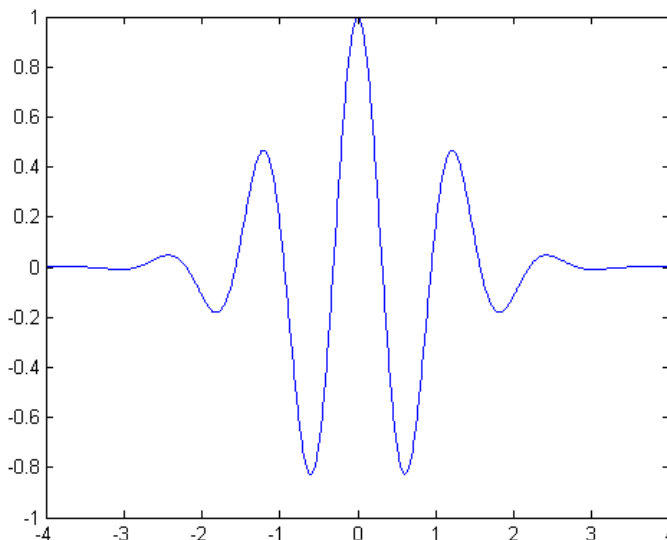
Οπότε μεταθέτοντας και καθυστερώντας το mother wavelet μπορούμε να κατασκευάσουμε την οικογένεια από wavelets ως:

$$\left\{ \psi_{j,n}(t) = \frac{1}{\sqrt{2^j}} \psi \left(\frac{t - 2^j n}{2^j} \right) \right\}_{(j,n) \in \mathbb{Z}^2}$$

Η οποία αποτελεί ορθοκανονική βάση του L^2 .

Τα τμήματα της εικόνας κειμένου, που αντιστοιχούν σε γράμματα ή λέξεις, αντιπροσωπεύονται από τους συντελεστές Wavelet, που αντιστοιχούν στα διάφορα επίπεδα ευκρίνειας. Αυτοί οι συντελεστές προωθούνται έπειτα σε έναν ταξινομητή για την αναγνώριση. Η απεικόνιση στην ανάλυση multiresolution (MRA) με χαμηλή ευκρίνεια μπορεί να απορροφήσει καλύτερα παραλλαγές στη γραφή από ότι σε MRA με υψηλή ευκρίνεια. Εντούτοις, η απεικόνιση στη χαμηλή ευκρίνεια μπορεί να οδηγήσει ώστε σημαντικές λεπτομέρειες για το στάδιο αναγνώρισης να χαθούν.

Morlet wavelet



2.3.2 Χαρακτηριστικά που προέρχονται από τη στατιστική κατανομή των σημείων

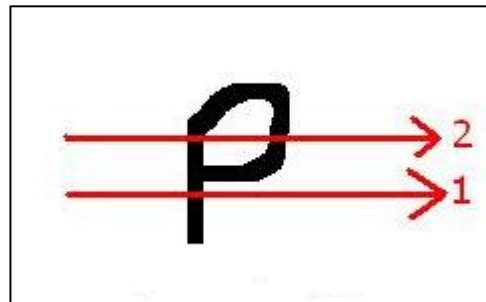
Χαρακτηριστικά που προέρχονται από τη στατιστική κατανομή των σημείων είναι ανθεκτικά σε παραμόρφωση και εφαρμόζονται σε αρκετές διαφορετικές γραμματοσειρές. Παρ' ότι η εφαρμογή τους είναι σχετικά γρήγορη και απλή, η δημιουργία μάσκας είναι δύσκολη. Μερικές από τις χαρακτηριστικές τεχνικές μέσα σε αυτήν την περιοχή παρατίθενται κατωτέρω.

- **Χωρισμός(Zoning)**

Το ορθογώνιο που περιγράφει το χαρακτήρα διαιρείται σε αρκετές περιοχές που επικαλύπτονται, ή όχι, και οι πυκνότητες των μαύρων σημείων μέσα σε αυτές τις περιοχές υπολογίζονται και χρησιμοποιούνται ως χαρακτηριστικά γνωρίσματα.

- **Διασταυρώσεις και αποστάσεις**

Στη τεχνική διασταύρωσης τα χαρακτηριστικά γνωρίσματα βρίσκονται από τον αριθμό φορών που η μορφή χαρακτήρα διασχίζεται από τα διανύσματα κατά μήκος ορισμένων κατευθύνσεων. Αυτή η τεχνική χρησιμοποιείται συχνά από τα



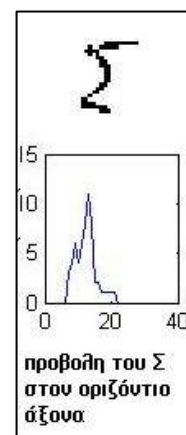
εμπορικά συστήματα επειδή μπορεί να εκτελεσθεί με υψηλή ταχύτητα και απαιτεί χαμηλή πολυπλοκότητα. Κατά τη χρησιμοποίηση της τεχνικής απόστασης ορισμένα τη μήκη μετριοούνται κατά τη διάσχιση των διανυσμάτων από τη μορφή χαρακτήρα. Παραδείγματος χάριν το μήκος των διανυσμάτων μέσα στο όριο των χαρακτήρα.

- **n-tuples**

Η ύπαρξη μαύρων και άσπρων σημείων (πρώτο πλάνο και υπόβαθρο) σε ορισμένες προσδιορισμένες διατάξεις, χρησιμοποιείται ως χαρακτηριστικό γνώρισμα

- **Προβολές**

Η κατανομή σημείων χαρακτήρων προβεβλημένων πάνω σε κάποιο άξονα μπορεί να προσφέρει πληροφορίες για την ταυτότητα του εκάστοτε χαρακτήρα. Στην ουσία μετατρέπουμε ένα διδιάστατο σήμα σε μονοδιάστα



3.3.3 Γεωμετρικά και τοπολογικά χαρακτηριστικά

Η μέθοδος ανάλυσης με γεωμετρικά και τοπολογικά χαρακτηριστικά είναι η πιο διαδεδομένη στους ερευνητικούς κύκλους. Τα χαρακτηριστικά αυτά μπορούν να αντιπροσωπεύουν γενικές και ειδικές ιδιότητες των χαρακτήρων όπως ακμές, κόλπους, κύκλους κτλ. Έχουν επίσης μεγάλη αντοχή σε παραμορφώσεις και διαφορετικά στιλ χαρακτήρων. Με τη βοήθειά τους η επεξεργασία χαρακτήρων γίνεται αρκετά γρήγορη. Από την άλλη μεριά οι διαδικασίες εξαγωγής τους είναι επίπονη και η δημιουργία μασκών δύσκολη. Ομαδοποιώντας τις διαφορετικές μεθόδους ανάλυσης, μπορούμε να τις εντάξουμε σε τέσσερις κατηγορίες:

- **Εξαγωγή τοπολογικών δομών**

Σε αυτή τη κατηγορία, αναζητούμε μια προκαθορισμένη δομή μέσα σε κάποιο χαρακτήρα ή λέξη. Ο αριθμός ή η σχετική με το χαρακτήρα θέση αυτών των δομών, δημιουργεί μια περιγραφική απεικόνιση. Κοινές πρωτογενείς δομές είναι οι ακμές, οι καμπύλες, οι ευθείες γραμμές κτλ.

- **Μέτρηση και την προσέγγιση των γεωμετρικών ιδιοτήτων**

Σε πολλές μελέτες, οι μετρήσεις γεωμετρικών ποσοτήτων απεικονίζουν τους χαρακτήρες. Τέτοιες είναι, η αναλογία μεταξύ πλάτους και ύψους του χαρακτήρα, η σχετική οριζόντια και κάθετη απόσταση μεταξύ πρώτου και τελευταίου σημείου, η απόσταση μεταξύ δύο σημείων, το μήκος μιας λέξης. Ένα πολύ σημαντικό χαρακτηριστικό μέτρο είναι η καμπυλότητα ή μεταβολή της καμπυλότητας.

- **Κωδικοποίηση**

Μία πολύ δημοφιλής μέθοδος κωδικοποίησης είναι η μέθοδος Freeman. Αυτή η κωδικοποίηση επιτυγχάνεται κυρίως με τη χαρτογράφηση ενός χαρακτήρα σε ένα 2-διαστάσεων παραμετρικό χώρο.

2.4 Ταξινόμηση και Εκπαίδευση

Τελικά, η διαδικασία οπτικής αναγνώρισης χαρακτήρων κατατάσσει μια εικόνα χαρακτήρα σε μια κατηγορία με τη χρησιμοποίηση ενός αλγορίθμου ταξινόμησης βασισμένου στα χαρακτηριστικά γνωρίσματα που εξάγονται και τις σχέσεις μεταξύ των χαρακτηριστικών γνωρισμάτων. Δεδομένου ότι τα μέλη μιας κατηγορίας χαρακτήρα είναι ισοδύναμα ή παρόμοια από τη στιγμή που μοιράζονται

καθοριστικές ιδιότητες, η μέτρηση της ομοιότητας είναι σημαντική σε οποιοδήποτε ταξινομητή.

2.4.1 Template matching

- **Απευθείας Σύγκριση**

Ένας χαρακτήρας εισόδου συγκρίνεται άμεσα με ένα πρότυπο σύνολο καταχωρημένων πρωτοτύπων. Σύμφωνα με ένα μέτρο ομοιότητας γίνεται ταίριασμα πρωτοτύπων για την αναγνώριση. Οι τεχνικές αντιστοίχισης μπορούν να είναι απλές όπως η ένα προς ένα σύγκριση ή σύνθετες όσο η ανάλυση δέντρων απόφασης στην οποία μόνο τα επιλεγμένα εικονοκύτταρα εξετάζονται. Αν και η άμεση αντιστοίχιση έχει σταθερές μαθηματικές βάσεις, το ποσοστό αναγνώρισης αυτής της μεθόδου είναι πολύ ευαίσθητο στο θόρυβο.

- **Παραμορφώσιμα πρωτότυπα και ελαστική αντιστοίχιση**

Μια εναλλακτική μέθοδος είναι η χρήση των παραμορφώσιμων πρωτοτύπων, όπου μια παραμόρφωση εικόνας χρησιμοποιείται για να ταιριάζει με μια άγνωστη εικόνα με μια βάση δεδομένων των γνωστών εικόνων. Ένα μέτρο ανομοιότητας προκύπτει από το ποσό παραμόρφωσης που απαιτείται.

- **Χαλαρή αντιστοίχιση**

Είναι μια τεχνική αντιστοίχισης εικόνας συμβολικού επιπέδου που χρησιμοποιεί μία χαρακτηριστική περιγραφή για την εικόνα χαρακτήρα. Κατ' αρχάς, προσδιορίζονται οι περιοχές αντιστοίχισης. Κατόπιν, τα στοιχεία εικόνας συγκρίνονται με το μοντέλο με βάση μερικές καθορισμένες εκτιμήσεις των αναθέσεων. Αυτή η διαδικασία απαιτεί μια τεχνική αναζήτησης σε ένα πολυδιάστατο διάστημα, για την εύρεση του σφαιρικού μεγίστου μερικών λειτουργιών

2.4.2 Στατιστικές Τεχνικές

Στις προσεγγίσεις στατιστικής ταξινόμησης, τα σχέδια εικόνας χαρακτήρα αντιπροσωπεύονται από σημεία σε ένα πολυδιάστατο χώρο χαρακτηριστικών γνωρισμάτων. Κάθε συστατικό του χώρου χαρακτηριστικών γνωρισμάτων είναι μια αξία μέτρησης, η οποία λειτουργεί ως μια τυχαία μεταβλητή που απεικονίζει την έμφυτη μεταβλητότητα ανάμεσα στις κατηγορίες. Ένας ταξινομητής χωρίζει το χώρο χαρακτηριστικών γνωρισμάτων σε περιοχές που συνδέονται με κάθε κατηγορία, Αντιστοιχίζοντας το παρατηρηθέν σχέδιο με κάποιο χαρακτήρα σύμφωνα με την περιοχή κατηγορίας στην οποία περιέχεται. Οι σημαντικότεροι τρόποι προσέγγισης είναι οι ακόλουθοι:

- **Μη-παραμετρική αναγνώριση**

Αυτή η μέθοδος χρησιμοποιείται για το διαχωρισμό, σε διαφορετικές κλάσεις, προτύπων . Η πιο γνωστή μέθοδος non parametric ταξινόμησης είναι ο κοντινότερος γείτονας (NN) και χρησιμοποιείται εκτενώς στην αναγνώριση χαρακτήρα. Δεν προϋποθέτει a priori πληροφορίες για τα στοιχεία. Ένα εισερχόμενο πρότυπο ταξινομείται χρησιμοποιώντας τον τομέα(cluster), του οποίου το κέντρο έχει την ελάχιστη απόσταση από το πρότυπο σε σχέση με όλους τους άλλους τομείς.

- **Παραμετρική αναγνώριση**

Δεδομένου ότι οι a priori πληροφορίες είναι διαθέσιμες για τους χαρακτήρες στα δεδομένα εκπαίδευσης(training data), είναι δυνατό να ληφθεί ένα παραμετρικό μοντέλο για κάθε χαρακτήρα .Από τη στιγμή που οι παράμετροι του μοντέλου, που είναι βασισμένο σε μερικές πιθανότητες, λαμβάνονται, οι χαρακτήρες ταξινομούνται σύμφωνα με μερικούς κανόνες απόφασης όπως η μέγιστη πιθανότητα ή η μέθοδος Bayes.

- **Ανάλυση Ομαδοποίησης(clustering)**

Οι τομείς των χαρακτηριστικών γνωρισμάτων χαρακτήρα, που αντιπροσωπεύουν τις ευδιάκριτες κλάσεις, αναλύονται με μεθόδους ομαδοποίησης. Η διαδικασία ομαδοποίησης είναι η εξής:

- ✓ Επιλογή του μέτρου ομοιότητας ή ανομοιότητας
- ✓ Επιλογή του κριτηρίου ομαδοποίησης
- ✓ Ορισμός του αλγορίθμου βελτιστοποίησης

Η ομαδοποίηση μπορεί να πραγματοποιηθεί είτε με συσσωρευτικούς είτε διαχωριστικούς αλγορίθμους. Οι συσσωρευτικοί αλγόριθμοι λειτουργούν βαθμιαία και συγχωνεύουν τους μικρούς τομείς σε μεγαλύτερους με βάση το κριτήριο ομαδοποίησης. Ενώ, οι διαχωριστικές μέθοδοι χωρίζουν τις κλάσεις χαρακτήρων.

2.4.3 Δομικές τεχνικές

Η επαναλαμβανόμενη περιγραφή ενός σύνθετου προτύπου μέσω απλούστερων βασισμένων στη μορφή του αντικειμένου ήταν η αρχική ιδέα για τη δημιουργία της δομικής αναγνώρισης προτύπων. Αυτά τα πρότυπα χρησιμοποιούνται για να περιγράψουν και να ταξινομήσουν τους χαρακτήρες στα συστήματα αναγνώρισης χαρακτήρα. Οι χαρακτήρες απεικονίζονται ως η ένωση των δομικών πρωτευόντων

στοιχείων. Τα πρωτεύοντα στοιχεία χαρακτήρα που εξάγονται από το γράψιμο πρέπει είναι ποσοτικά προσδιορίσιμα και συσχετίσιμα. Οι ακόλουθες δομικές μέθοδοι εφαρμόζονται στα προβλήματα αναγνώρισης χαρακτήρα.

- **Γραμματικές μέθοδοι**

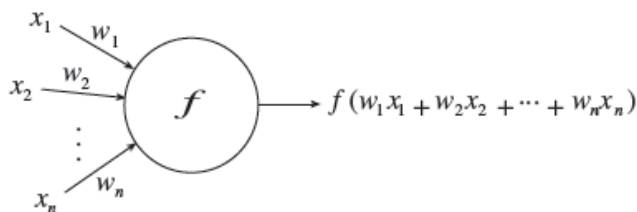
Οι γραμματικές μέθοδοι δημιουργούν μερικούς κανόνες παραγωγής προκειμένου να διαμορφωθούν οι χαρακτήρες από ένα σύνολο πρωτόγονων μέσω των επίσημων γραμματικών. Αυτές οι μέθοδοι μπορούν να συνδυάσουν οποιοδήποτε τύπο τοπολογικών και στατιστικών χαρακτηριστικών γνωρισμάτων σύμφωνα με μερικούς συντακτικούς ή/και σημασιολογικούς κανόνες

- **Γραφικές μέθοδοι**

Οι γραφικές μονάδες αντιπροσωπεύονται από τα δέντρα, τους γράφους ή τις γραφικές παραστάσεις. Τα πρωτεύοντα στοιχεία χαρακτήρα (π.χ. καμπύλες) επιλέγονται από μια δομική προσέγγιση, ανεξάρτητα από το πώς γίνεται η λήψη τελικής απόφασης στην αναγνώριση. Για κάθε κλάση, ένας γράφος ή ένα δέντρο διαμορφώνεται στο στάδιο κατάρτισης για να αντιπροσωπεύσει τις καμπύλες, τα γράμματα ή τις λέξεις. Το στάδιο αναγνώρισης κατατάσσει την άγνωστη γραφική παράσταση σε μια από τις κλάσεις με τη χρησιμοποίηση ενός μέτρου ομοιότητας

2.4.4 Νευρωνικά Δίκτυα

Ένα τεχνητό νευρωνικό δίκτυο (ΤΝΔ) είναι ένα μοντέλο επεξεργασίας πληροφοριών που εμπνέεται από τον τρόπο που επεξεργάζονται την πληροφορία τα βιολογικά νευρικά συστήματα, όπως ο εγκέφαλος. Το βασικό στοιχείο αυτού του μοντέλου είναι η δομή του συστήματος επεξεργασίας πληροφοριών, αποτελείται από τη μαζικά παράλληλη διασύνδεση των προσαρμοστικών "νευρικών" επεξεργαστών(νευρώνων). Στην εικόνα βλέπουμε τη δομική μονάδα ενός νευρωνικού δικτύου, ένα νευρώνα.



Κάθε είσοδος του νευρώνα είναι πολλαπλασιασμένη με ένα μέγεθος που ονομάζεται βάρος. Η f είναι μία προεπιλεγμένη συνάρτηση που μας υπολογίζει την έξοδο του νευρώνα.

Λόγω της παράλληλης φύσης της δομής ενός ΤΝΔ, αυτό μπορεί να εκτελέσει τους υπολογισμούς ταχύτερα από τις κλασσικές τεχνικές. Ένα ΤΝΔ διαμορφώνεται για μια συγκεκριμένη εφαρμογή, όπως η αναγνώριση προτύπων ή η ταξινόμηση στοιχείων, μέσω μιας διαδικασίας εκμάθησης. Λόγω της προσαρμοστικής φύσης του, μπορεί να προσαρμοστεί στις αλλαγές των στοιχείων και να μάθει τα χαρακτηριστικά του σήματος εισόδου.

Ένα νευρωνικό δίκτυο περιέχει πολλούς κόμβους. Η έξοδος από έναν κόμβο γίνεται είσοδος σε έναν άλλο στο δίκτυο και η τελική απόφαση εξαρτάται από τη σύνθετη αλληλεπίδραση όλων των κόμβων

Οι γενικές μαθηματικές σχέσεις που συμβολίζουν τη λειτουργία ενός νευρωνικού δικτύου είναι οι εξής:

$$Y=f(x)$$

$$f(x) = K \left(\sum_i w_i g_i(x) \right)$$

Όπου

K: η συνάρτηση ενεργοποίησης

Y: η έξοδος του νευρωνικού δικτύου

g_i : η έξοδος κάθε νευρώνα

w_i : τα βάρη των νευρώνων

Πλεονεκτήματα:

- Ένα νευρωνικό δίκτυο μπορεί να εκτελέσει εργασίες που ένα γραμμικό πρόγραμμα δεν μπορεί.
- Όταν ένα στοιχείο του νευρικού δικτύου αποτυγχάνει, η διαδικασία μπορεί να συνεχιστεί χωρίς πρόβλημα λόγω της παράλληλης διάταξης των νευρώνων.
- Ένα νευρικό δίκτυο μαθαίνει και δεν χρειάζεται να επαναπρογραμματιστεί. Μπορεί να εφαρμοστεί σε οποιαδήποτε εφαρμογή.

Μειονεκτήματα:

1. Το δίκτυο είναι αναγκαίο να εκπαιδευτεί για να λειτουργήσει.
2. Η αρχιτεκτονική ενός νευρωνικού δικτύου είναι διαφορετική από την αρχιτεκτονική των μικροεπεξεργαστών επομένως πρέπει να εξομοιωθεί.
3. Απαιτείται υψηλός χρόνος επεξεργασίας για τα μεγάλα νευρωνικά δίκτυα.

- **Τοπολογία δικτύων**

Οι νευρωνικές δικτυακές αρχιτεκτονικές μπορούν να ταξινομηθούν σε δύο σημαντικές ομάδες: feed-forward δίκτυα και ανατροφοδότησης .

1. Στα feed forward δίκτυα η ροή δεδομένων γίνεται αυστηρά μόνο με κατεύθυνση από την είσοδο προς την έξοδο.
2. Στα δίκτυα ανατροφοδότησης η ροή μπορεί να γίνεται και αντίστροφα, με ανατροφοδότηση της εξόδου στην είσοδο στο ίδιο layer η και προηγούμενο

Τα πιο κοινά νευρωνικά δίκτυα που χρησιμοποιούνται στα συστήματα αναγνώρισης χαρακτήρα είναι το πολυστρωματικό perceptron των feed-forward δικτύων και ο αυτοοργανωτικός χάρτης του Kohonen (SOM) των δικτύων ανατροφοδότησης.

- **Εκπαίδευση**

Διάφορες προσεγγίσεις υπάρχουν για την εκπαίδευση των νευρωνικών δικτύων . Αυτές περιλαμβάνουν τη δοκιμή και σφάλμα, τη πίσω διάδοση, τη μηχανή Boltzman, το δίκτυο Hebbian και την ανταγωνιστική εκμάθηση, οι οποίες καλύπτουν τη δυαδική και συνεχή εκτιμημένη είσοδο, καθώς επίσης και την εποπτευμένη και ανεπίβλεπτη εκμάθηση.

Μπορούμε επίσης να προγραμματίσουμε ένα νευρωνικό δίκτυο ώστε να πραγματοποιήσει μια διεργασία δημιουργώντας χειροκίνητα την τοπολογία του και θέτοντας τα βάρη των νευρώνων. Όμως έτσι, παραμερίζουμε τη θεμελιώδη ιδιότητα των νευρωνικών δικτύων η οποία είναι η δυνατότητα να προγραμματίζονται μόνα τους.

Εκπαίδευση με επίβλεψη

Η φιλοσοφία αυτής της μεθόδου είναι πως για κάθε συγκεκριμένη είσοδο στο νευρωνικό δίκτυο εισάγουμε εμείς εκ των προτέρων μια επιθυμητή έξοδο. Με αυτό το τρόπο το νευρωνικό δίκτυο υπολογίζει το σφάλμα μεταξύ πραγματικής και επιθυμητής εξόδου και αλλάζοντας τα βάρη σύμφωνα με ένα καθορισμένο αλγόριθμο επιχειρεί να μειώσει αυτό το σφάλμα και να βρει τις τιμές βαρών που το καθιστούν ελάχιστο, με άλλα λόγια εκπαιδεύεται.

- **Δοκιμή και σφάλμα**

Η πιο βασική μέθοδος της εκπαίδευσης ενός νευρωνικού δικτύου είναι η δοκιμή και σφάλμα. Αν το δίκτυο δεν συμπεριφέρεται με τον τρόπο που πρέπει, αλλάζουμε το βάρος ενός τυχαίου συνδέσμου κατά ένα τυχαίο ποσό. Εάν η ακρίβεια του δικτύου μειώνεται, αναιρούμε τις αλλαγές και δοκιμάζουμε άλλες. Χρειάζεται χρόνος, αλλά η μέθοδος αυτή παράγει αποτελέσματα.

- **Πίσω-διάδοση**

Ο αλγόριθμος πίσω-διάδοσης συγκρίνει τα αποτελέσματα που λαμβάνονται με το αποτέλεσμα που αναμενόταν. Στη συνέχεια, χρησιμοποιεί αυτές τις πληροφορίες να τροποποιήσει συστηματικά τα βάρη όλου του νευρωνικού δικτύου. Η εκπαίδευση αυτή διαρκεί μόνο ένα κλάσμα του χρόνου που απαιτεί η μέθοδος δοκιμή και σφάλμα . Μπορεί επίσης να χρησιμοποιηθεί αξιόπιστα για την κατάρτιση των δικτύων με ένα μόνο μέρος των δεδομένων. Τα δίκτυα που προκύπτουν είναι συνήθως σωστά και μπορούν να λύνουν προβλήματα για τα οποία δεν έχουν ποτέ εκπαιδευτεί ειδικά (Neil Fraser,1998) .

Εκπαίδευση χωρίς επίβλεψη

Σε αυτή τη μέθοδο το σύστημα προσπαθεί μόνο του να οργανώσει και να ομαδοποιήσει τα δεδομένα αναγνωρίζοντας τις ιδιότητές τους. Συνήθως πραγματοποιείται κάποιου είδους προεπεξεργασία όπως το data mining ώστε να προκύψουν οι ιδιότητες των δεδομένων. Παραδείγματα τέτοιων μεθόδων είναι η ομαδοποίηση και η μείωση διαστάσεων.

- **Ομαδοποίηση(clustering)**

Ομαδοποίηση ονομάζεται η ταξινόμηση των δεδομένων σε υποομάδες σύμφωνα με κάποιο στοιχείο ομοιότητας μεταξύ τους. Για την πραγματοποίηση της ,το σύστημα πρέπει να εξάγει με κάποιο αλγόριθμο την ιδιότητα των δεδομένων προς σύγκριση. Έπειτα η ομοιότητα υπολογίζεται σύμφωνα με κάποια απόσταση την οποία έχουμε ορίσει.

- **Μείωση διαστάσεων(dimensionality reduction)**

Με αυτή τη μέθοδο μετασχηματίζουμε τα δεδομένα ώστε να περάσουμε από ένα χώρο με πολλές διαστάσεις σε κάποιον με λιγότερες. Σκοπός μας είναι να διατηρήσουμε όσο το δυνατόν πιο σταθερή τη διασπορά των δεδομένων ώστε να παραμείνουν σταθερές οι μεταξύ τους σχετικές αποστάσεις

2.5 Μετεπεξεργασία

Το αποτέλεσμα της οπτικής αναγνώρισης του τυπωμένου ή χειρόγραφου κειμένου επηρεάζεται συχνά από ένα μη αμελητέο ποσό από λάθη και αβεβαιότητα, και είναι επομένως ουσιαστική η εφαρμογή ενός αλγορίθμου διορθώσεων. Μέχρι αυτό το σημείο καμία σημασιολογική πληροφορία δεν εξετάζεται κατά τη διάρκεια των

σταδίων της αναγνώρισης χαρακτήρα. Ένας κοινός στόχος των μεθόδων μεταεπεξεργασίας οπτικής αναγνώρισης χαρακτήρων είναι να εγγυηθεί (ή, τουλάχιστον, να μεγιστοποιήσει την πιθανότητα) ότι οι λέξεις ή οι προτάσεις που παράγονται αυτή είναι σωστές υπό την έννοια ότι ανήκουν στη γλώσσα του στόχου. Το κριτήριο μπορεί να είναι τόσο απλό όσο ένα μικρό σύνολο έγκυρων λέξεων (π.χ. Πιθανά ονόματα χωρών) ή τόσο σύνθετη όσο μια αβίαστη πρόταση σε μια φυσική γλώσσα. Αντίστοιχα, διάφορες προσεγγίσεις υπάρχουν στο πρόβλημα, το οποίο μπορεί να ταξινομηθεί σε δύο κύριες ομάδες: ντετερμινιστικές και μη ντετερμινιστικές μέθοδοι.

Κεφάλαιο 3

Πείραμα

Στη παρούσα διπλωματική εργασία υλοποιήσαμε σε περιβάλλον Matlab ένα σύστημα Οπτικής αναγνώρισης ελληνικών χειρόγραφων χαρακτήρων.

Το σύστημα χρησιμοποιεί μετασχηματισμούς Wavelet για την εξαγωγή των παραμέτρων και RBF νευρωνικό δίκτυο για την αναγνώριση των χαρακτήρων.

Για την εκπαίδευση των νευρωνικού δικτύου χρησιμοποιήσαμε τη βάση δεδομένων GCDB (John Margaronis, Minas Christou, Ergina Kavallieratou, Theodoros Tzouramanis).

3.1 Διαδικασία

1. Καταστρώνουμε τον πίνακα εισόδου χρησιμοποιώντας τη συνάρτηση `features2` που δημιουργήσαμε στο matlab.

2. Καταστρώνουμε τον πίνακα εξόδου (επιθυμητές τιμές)

3. Δημιουργούμε το RBF νευρωνικό δίκτυο χρησιμοποιώντας τη συνάρτηση `newrb` του matlab θέτοντας τις κατάλληλες παραμέτρους (π.χ. `spread=10`, `max. Neurons=30`).

4. Δημιουργούμε διανύσματα αγνώστων χαρακτήρων προς αναγνώριση χρησιμοποιώντας τη συνάρτηση `features3`.

5. Πραγματοποιούμε εξομοίωση του κάθε ενός δικτύου με τη συνάρτηση `sim` θέτοντας για είσοδο τα διανύσματα προς αναγνώριση και λαμβάνουμε τα αποτελέσματα αυτής.

6. Κάνουμε αναγνώριση του χαρακτήρα αντιστοιχώντας τη μέγιστη τιμή εξόδου στο αντίστοιχο χαρακτήρα.

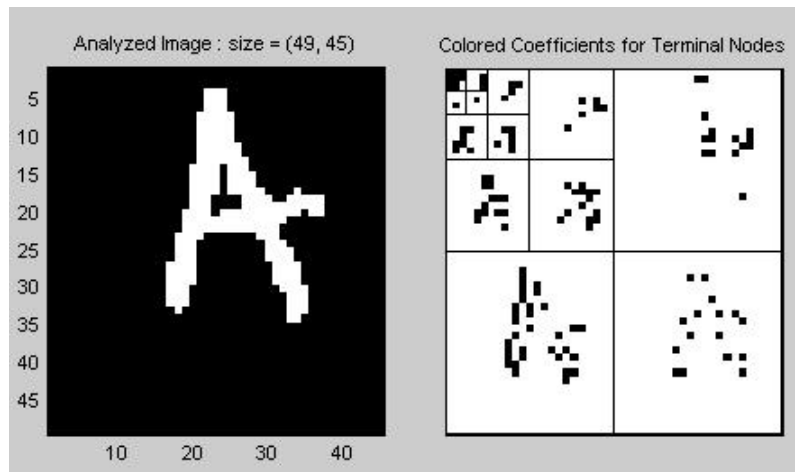
7. Εξάγουμε το ποσοστό αναγνώρισης και επαναλαμβάνουμε τη διαδικασία για διαφορετικές τιμές παραμέτρων.

Αναλυτικότερα:

3.2 Εξαγωγή παραμέτρων

Οι εικόνες των χαρακτήρων που χρησιμοποιούμε είναι δυαδικές και έχουν μέγεθος 49x45 pixels. Που σημαίνει πως κάθε γράμμα αποτελείται από 2205 στοιχεία.

Ως γνωστόν για να λειτουργήσει σωστά



και να είναι δυνατόν να εκπαιδύσουμε το νευρωνικό μας δίκτυο πρέπει να μειωθεί ο όγκος των δεδομένων. Για αυτό το λόγο καταφεύγουμε στο μετασχηματισμό και συμπίεση της εικόνας με wavelet.

Συγκεκριμένα, πραγματοποιήσαμε Multiresolution ανάλυση σε 4 επίπεδα χρησιμοποιώντας το wavelet 'sym4' και στη συνέχεια το διακριτό μετασχηματισμό Meyer('dmey'). Στη συνέχεια υπολογίσαμε τη μέση τιμή και τη διακύμανση κάθε πίνακα συντελεστών. Αποτέλεσμα ήταν να λάβουμε 26 διαφορετικές τιμές για κάθε εικόνα χαρακτήρα οι οποίες αποτελούν τις παραμέτρους αναγνώρισης.

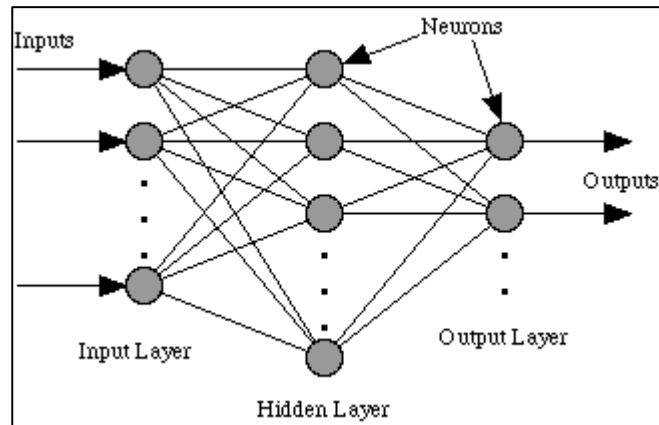
Πραγματοποιούμε το πείραμα για 100 και 200 χειρόγραφους χαρακτήρες εκπαίδευσης από κάθε γράμμα και εξετάζουμε τα ποσοστά αναγνώρισης σε κάθε περίπτωση για διαφορετικές τιμές διασποράς των νευρώνων του συστήματος.

3.3 Αναγνώριση

Για το στάδιο της αναγνώρισης επιλέξαμε να χρησιμοποιήσουμε ένα RBF νευρωνικό δίκτυο.

Το RBF νευρωνικό δίκτυο είναι ένα feed forward δίκτυο που αποτελείται από 3 στρώματα (είσοδος-κρυφό επίπεδο-έξοδος).

Αρχικά πρέπει να εκπαιδύσουμε το δίκτυο και για αυτό το λόγο τοποθετούμε στην είσοδο την μήτρα εκπαίδευσης, η οποία έχει για στήλες τα διανύσματα παραμέτρων κάθε γράμματος που χρησιμοποιείται για εκπαίδευση και στην έξοδο τη μήτρα θεμιτών αποτελεσμάτων.



Οι νευρώνες του κρυφού επιπέδου περιέχουν την Γκαουσιανή συνάρτηση μεταφοράς και οι έξοδοι τους είναι αντιστρόφως ανάλογες προς την απόσταση των εισόδων από το κέντρο του νευρώνα. Δηλαδή η μαθηματική σχέση που μας δίνει την έξοδο του κρυφού επιπέδου είναι η εξής:

$$h_j = \exp\left\{\frac{-\|X^i - U_j\|}{2\sigma_j^2}\right\}$$

Όπου:

X: η είσοδος του νευρώνα

σ: η διασπορά

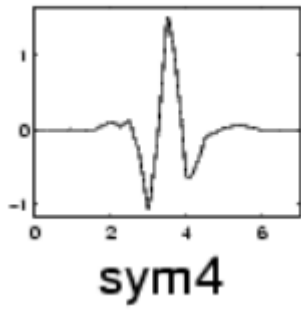
U: το κέντρο του νευρώνα

Το spread της Γκαουσιανής συνάρτησης το ορίζουμε πειραματικά και είναι το ίδιο για όλους τους νευρώνες στο δικό μας πείραμα.

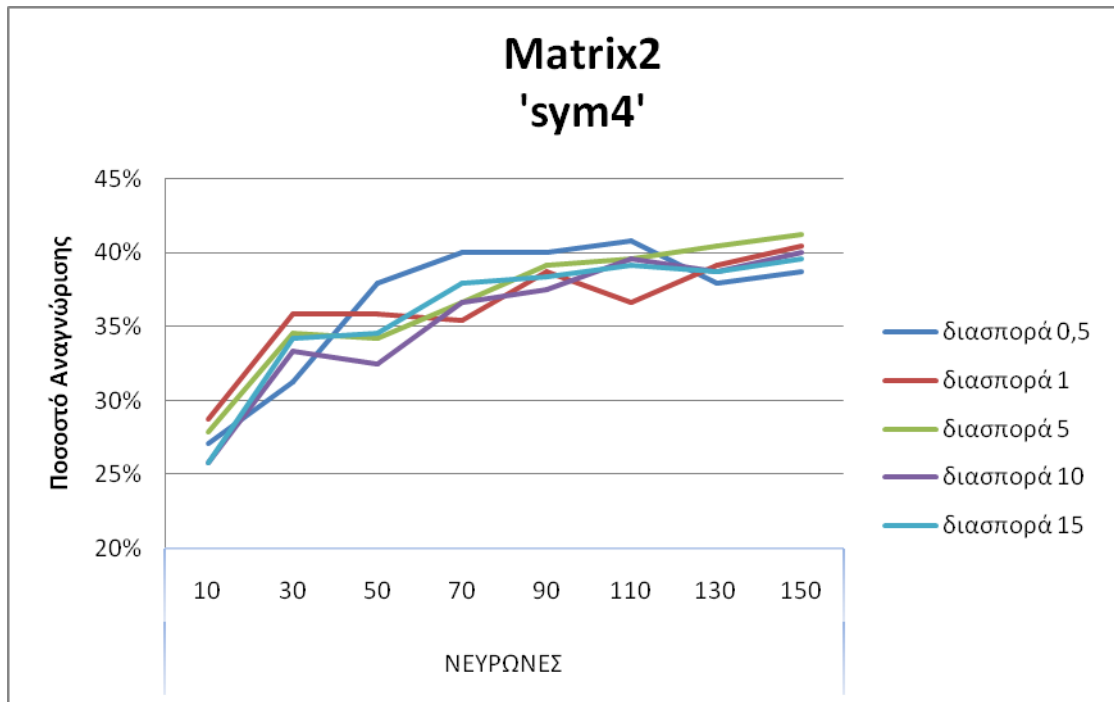
3.4 Πείραμα 1

3.4.1 Πείραμα 1.α

Πίνακας εκπαίδευσης (Matrix2)=100x24=2400 χαρακτήρες
Πίνακας αναγνώρισης(test)=(10x24)=240 χαρακτήρες
Μετασχηματισμός wavelet='sym4'



Αποτελέσματα



Αριθμός αναγνώρισεων(στα 240):

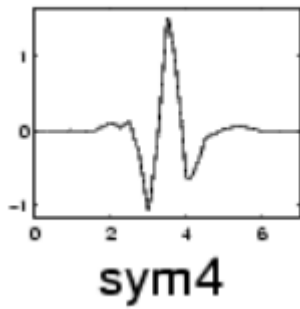
		Αριθμός νευρώνων							
		10	30	50	70	90	110	130	150
Διασπορά	0,5	65	75	91	96	96	98	91	93
	1	69	86	86	85	93	88	94	97
	5	67	83	82	88	94	95	97	99
	10	62	80	78	88	90	95	93	96
	15	62	82	83	91	92	94	93	95

3.4.2 Πείραμα 1.β

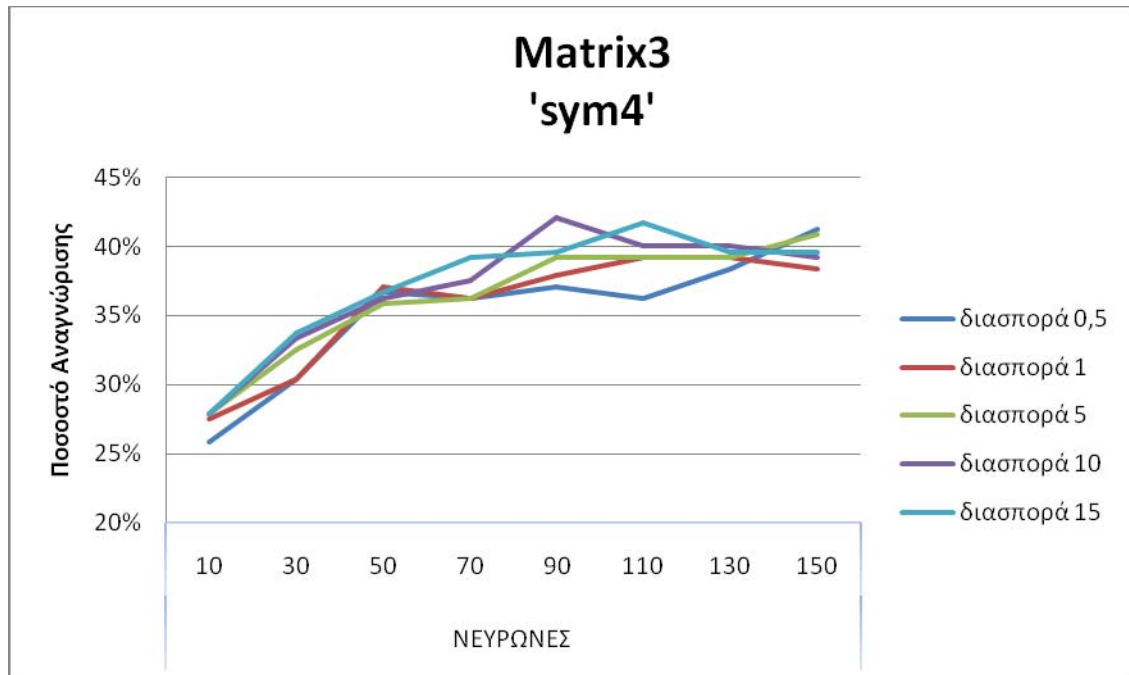
Πίνακας εκπαίδευσης (Matrix2)=200x24=4800 χαρακτήρες

Πίνακας αναγνώρισης(test)=(10x24)=240 χαρακτήρες

Μετασχηματισμός wavelet='sym4'



Αποτελέσματα



		ΝΕΥΡΩΝΕΣ							
		10	30	50	70	90	110	130	150
Διασπορά	0,5	62	73	88	87	89	87	92	99
	1	66	73	89	87	91	94	94	92
	5	67	78	86	87	94	94	94	98
	10	67	80	87	90	101	96	96	94
	15	67	81	88	94	95	100	95	95

3.4.3 Συμπεράσματα

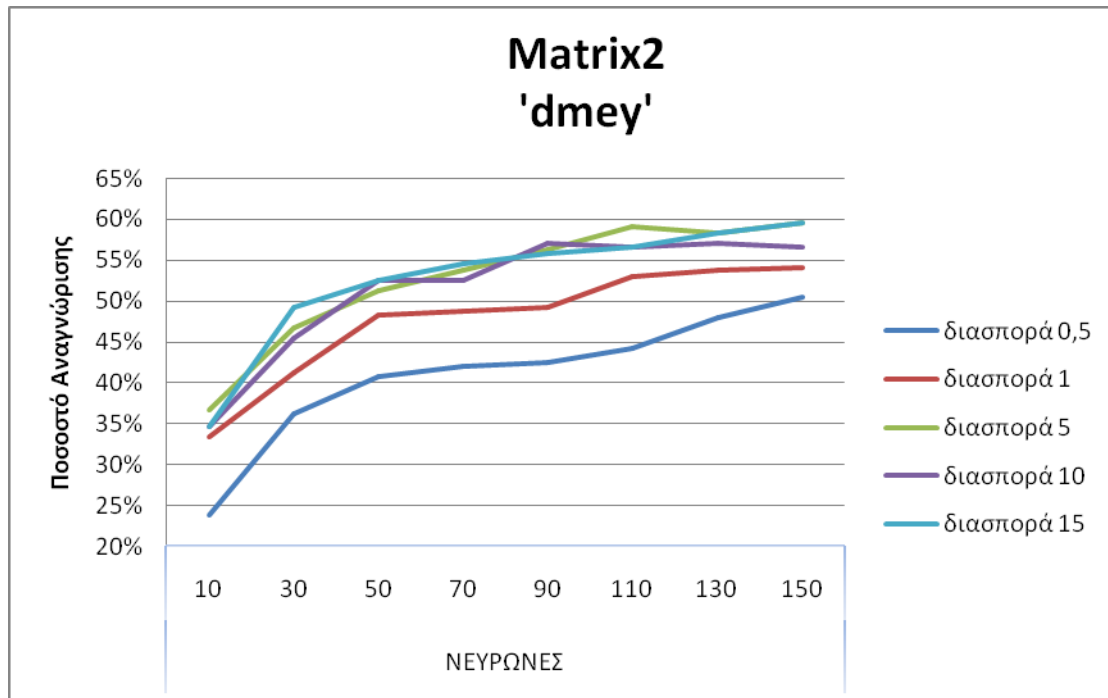
1. Πριν απ' όλα παρατηρούμε το αρκετά χαμηλό ποσοστό αναγνώρισης σε όλες τις περιπτώσεις το οποίο δεν ξεπερνάει το 42%. Πράγμα που καθιστά το σύστημά μας χρήσιμο μόνο για συγκριτική μελέτη των αποτελεσμάτων και εξαγωγή κάποιων συμπερασμάτων που αφορούν τις αλλαγές στις παραμέτρους που χρησιμοποιήσαμε. Τα αίτια του χαμηλού ποσοστού αναγνώρισης θα εξεταστούν αργότερα.
2. Όσο αφορά τον αριθμό των νευρώνων παρατηρούμε ότι: Αυξάνοντας τον αριθμό των νευρώνων μέχρι να φτάσουμε τους 90, σημειώνεται μια σημαντική αύξηση στο ποσοστό αναγνώρισης η οποία φτάνει και το 15% σε σχέση με αυτό που προέκυψε όταν χρησιμοποιήσαμε 10 νευρώνες. Από τους 90 νευρώνες και μετά όμως, παρατηρούμε ότι η αύξηση στον αριθμό τους δεν προκαλεί και αύξηση του ποσοστού αναγνώρισης το οποίο μένει σταθερό ή μεταβάλλεται ελαφρώς. Αυτό συμβαίνει διότι υπερβολικός αριθμός νευρώνων μπορεί να οδηγήσει σε κατάσταση overlearning όπου το σύστημα γίνεται πολύ ευαίσθητο στο θόρυβο οπότε αρχίζει κ μειώνεται η αποτελεσματικότητά του.
3. Συγκρίνοντας τα 2 σχεδιαγράμματα μεταξύ τους (Matrix2 και Matrix3) παρατηρούμε πως τα ποσοστά αναγνώρισης δεν βελτιώνονται παρότι μεγαλώσαμε τη βάση δεδομένων που χρησιμοποιήσαμε για την εκπαίδευση του συστήματος πράγμα που σημαίνει ότι η πρώτη βάση δεδομένων είναι υπεραρκετή και πιθανότατα θα μπορούσε να είναι και μικρότερη.
4. Διαφορά βλέπουμε στη συμπεριφορά του συστήματος αλλάζοντας τη βάση δεδομένων για τιμή διασποράς ίση με 0.5. Στη πρώτη περίπτωση η βέλτιστη επίδοση παρατηρείται για 90 νευρώνες ενώ στη δεύτερη περίπτωση για 150.

3.5 Πείραμα 2

3.5.1 Πείραμα 2.α

Πίνακας εκπαίδευσης (Matrix2)=100x24=2400 χαρακτήρες
Πίνακας αναγνώρισης(test)=(10x24)=240 χαρακτήρες
Μετασχηματισμός wavelet='dmey'

Αποτελέσματα



Αριθμός Αναγνωρίσεων(στα 240)

		ΝΕΥΡΩΝΕΣ							
		10	30	50	70	90	110	130	150
Διασπορά	0,5	57	87	98	101	102	106	115	121
	1	80	99	116	117	118	127	129	130
	5	88	112	123	129	135	142	140	143
	10	83	109	126	126	137	136	137	136
	15	83	118	126	131	134	136	140	143

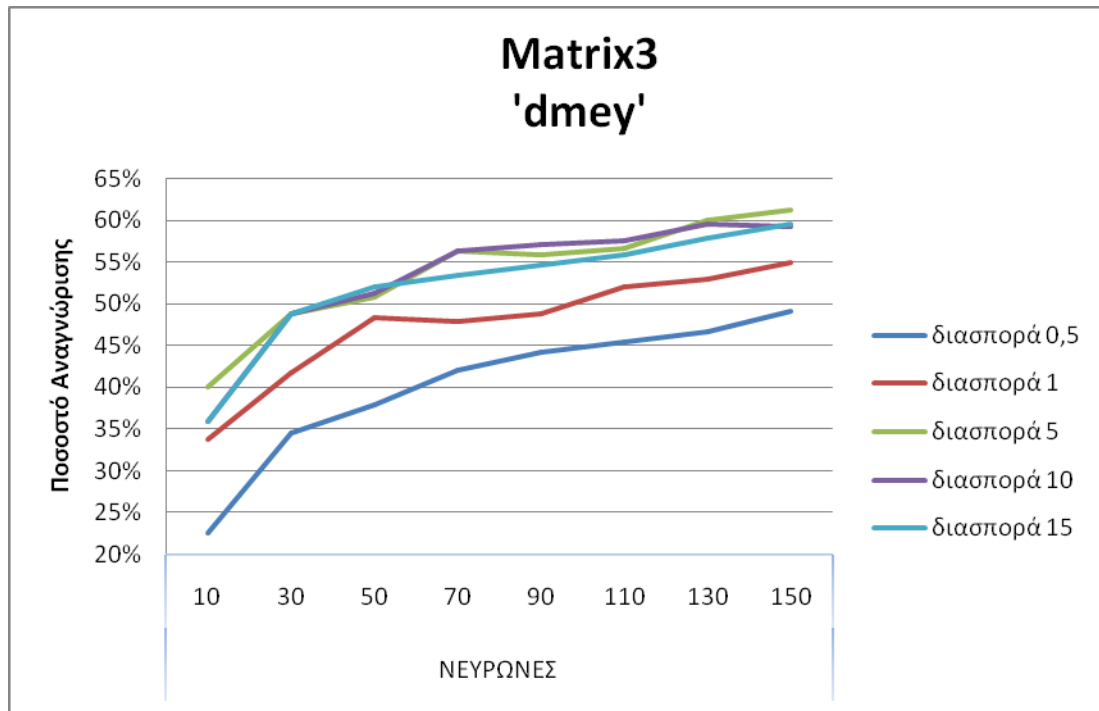
3.5.2 Πείραμα 2.β

Πίνακας εκπαίδευσης (Matrix2)=200x24=4800 χαρακτήρες

Πίνακας αναγνώρισης(test)=(10x24)=240 χαρακτήρες

Μετασχηματισμός wavelet='dmey'

Αποτελέσματα



Αριθμός Αναγνώρισεων(στα 240)

		NEΥΡΩΝΕΣ							
		10	30	50	70	90	110	130	150
Διασπορά	0,5	54	83	91	101	106	109	112	118
	1	81	100	116	115	117	125	127	132
	5	96	117	122	135	134	136	144	147
	10	86	117	123	135	137	138	143	142
	15	86	117	125	128	131	134	139	143

3.5.3 Συμπεράσματα

1. Πριν απ' όλα η σημαντικότερη παρατήρηση που κάνουμε είναι ότι χρησιμοποιώντας το μετασχηματισμό 'dmev' αντί του 'sym4' το ποσοστό αναγνώρισης βελτιώθηκε δραματικά, περίπου κατά 20%, φτάνοντας το στο 60% που είναι όμως επίσης αρκετά χαμηλό. Αυτή η διαφορά όμως είναι αρκετή ώστε να βγάλουμε το συμπέρασμα πως χρησιμοποιώντας σαν παραμέτρους για αναγνώριση χαρακτήρα τη μέση τιμή και τη διασπορά, ο διακριτός μετασχηματισμός Meyer είναι προτιμότερος από το wavelet 'sym4'.
2. Παρατηρούμε για άλλη μια φορά πως η μεγαλύτερη μήτρα εκπαίδευσης δε μας προσέφερε τίποτα σε θέματα αποτελεσματικότητας. Αντιθέτως έκανε τους υπολογισμούς πιο πολύπλοκους και το σύστημά μας πιο αργό.
3. Όσον αφορά τη διασπορά σε αυτό το πείραμα φαίνεται πως μέχρι αυτή να φτάσει την τιμή 5 παρουσιάζεται βελτίωση στην απόδοση. Στη συνέχεια οποιοδήποτε μεταβολή δεν παρουσιάζει ουσιαστική βελτίωση.

3.6 Βελτιώσεις

Όπως διαπιστώσαμε από τα πειράματα το σύστημα έχει χαμηλή απόδοση όμως θα μπορούσε να αποτελέσει βάση για ένα πιο ανεπτυγμένο σύστημα αναγνώρισης χαρακτήρα.

Πρώτη παράμετρος που μπορεί να βελτιωθεί είναι ο υπολογισμός του spread κάθε νευρώνα ξεχωριστά. Υπάρχουν μέθοδοι υπολογισμού του spread ώστε ο κάθε νευρώνας και ολόκληρο το δίκτυο να έχουν βέλτιστη συμπεριφορά. Με αυτό τον τρόπο το ποσοστό αναγνώρισης θα μπορούσε να ανέβει αρκετά.

Άλλος τρόπος βελτίωσης θα μπορούσε να είναι η χρήση υψηλότερης ανάλυσης στις εικόνες χαρακτήρων. Αυτό θα μας επέτρεπε να αναλύσουμε τους χαρακτήρες σε περισσότερα από 4 επίπεδα όπως κάναμε σε αυτή την εργασία. Έτσι θα είχαμε περισσότερες και ακριβέστερες παραμέτρους.

Επίσης θα μπορούσαν να χρησιμοποιηθούν επιπλέον, άλλων ειδών παράμετροι, όπως θα ήταν οι διασταυρώσεις ή οι τα τοπολογικά χαρακτηριστικά. Με αυτό τον τρόπο τα κριτήρια αναγνώρισης θα γίνονταν περισσότερα και θα οδηγούμασταν σε καλύτερα αποτελέσματα.

Βιβλιογραφία

R. Rojas: *Neural Networks*, Springer-Verlag, Berlin, 1996

Rafael C. Gonzalez, Richard E. Wood, Steven, L. Eddins- *Digital Image Processing Using Matlab*

Bernd Jähne-*Handbook of Computer Vision and Applications*

James S. Walker-*Wavelet-based image processing*

Stephane Mallat – *A wavelet tour of signal processing*

V.K. Govindan, A.P. Shivaprasad- *Character Recognition: A Review*

J. Mantas- *An Overview of Character Recognition Methodologies*

R H Davis and J Lyall- *Recognition of handwritten characters a review*

Oivind Due Trier, Anil K. Jain, Torfinn Taxt- *Feature extraction methods for character recognition—a survey*

Yi Lu and M. Shridhar-*Character segmentation in handwritten words-an overview*

Nafiz Arica- *An Overview Of Character Recognition Focused On Off-line Handwriting*

Homayoon S.M. Beiji- *An overview on handwriting recognition*

Line Eikvil-*Optical Character Recognition*