

## Μάθημα 6

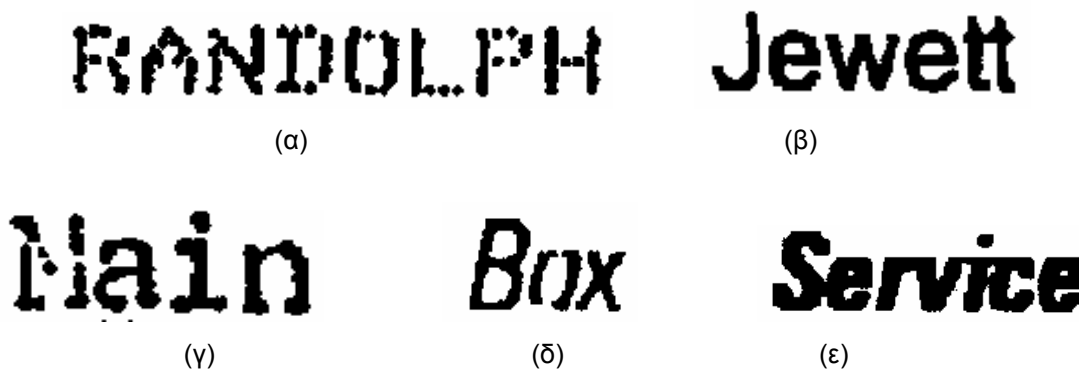
# Εντοπισμός χαρακτήρων

Το στάδιο του εντοπισμού των χαρακτήρων αφορά την επεξεργασία τμημάτων κειμένου ώστε να απομονωθούν οι χαρακτήρες που υπάρχουν σε αυτά. Το στάδιο αυτό επηρεάζει σημαντικά την απόδοση ενός συστήματος οπτικής αναγνώρισης χαρακτήρων γιατί η ύπαρξη χαρακτήρων που διαιρούνται σε δύο κομμάτια ή η συνένωση γειτονικών χαρακτήρων μπορεί να οδηγήσει σε σοβαρά λάθη αναγνώρισης. Τα τμήματα του κειμένου που επεξεργαζόμαστε μπορεί να είναι τμήματα που περιέχουν λέξεις ή γραμμές κειμένου. Οι διάφορες μεθοδολογίες για τον εντοπισμό των χαρακτήρων χωρίζονται σε αυτές που αφορούν τυπωμένο και σε αυτές που αφορούν χειρόγραφο κείμενο. Περιληπτική παρουσίαση των μεθόδων εντοπισμού των χαρακτήρων σε τυπωμένο και χειρόγραφο κείμενο δίδεται από τον (Lu 1994) και (Lu 1995).

### 6.1 Εντοπισμός χαρακτήρων σε τυπωμένο κείμενο

Το τυπωμένο κείμενο μπορεί να αποτελείται από χαρακτήρες οι οποίοι ανήκουν σε μία ή παραπάνω από τις ακόλουθες κατηγορίες (η σειρά αφορά και τον αυξανόμενο βαθμό δυσκολίας για τον εντοπισμό των μεμονωμένων χαρακτήρων):

- Χαρακτήρες ίδιου πλάτους (το νοητό ορθογώνιο παραλληλόγραμμο που περικλείει κάθε χαρακτήρα του κειμένου έχει σταθερό πλάτος) (σχήμα 6.1α).
- Χαρακτήρες που έχουν πλάτος αναλογικό του μεγέθους τους (σχήμα 6.1β).
- Χαρακτήρες με σπασίματα (σχήμα 6.1γ).
- Στραμμένους και επικαλυπτόμενους χαρακτήρες (italic - kerned) (σχήμα 6.1δ).
- Ενωμένους χαρακτήρες (σχήμα 6.1ε).



**Σχήμα 6.1.** Παραδείγματα τυπωμένων χαρακτήρων. (α) Χαρακτήρες ίδιου πλάτους. (β) Χαρακτήρες αναλογικού πλάτους. (γ) Χαρακτήρες με σπασίματα. (δ) Ενωμένοι χαρακτήρες. (ε) Στραμμένοι χαρακτήρες.

Για την αντιμετώπιση των προβλημάτων που έχει κάθε μία από τις παραπάνω κατηγορίες έχουν προταθεί διάφορες μέθοδοι που αναλύονται παρακάτω.

### Χαρακτήρες ίδιου πλάτους

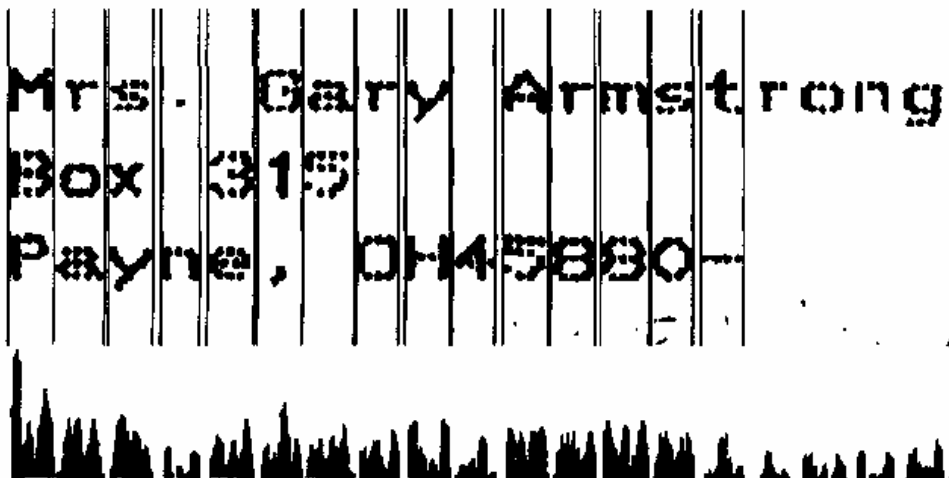
Οι κάθετες προβολές χρησιμοποιούνται σε μεγάλο βαθμό για την κατάτμηση των χαρακτήρων. Η κάθετη προβολή  $V(x)$  είναι το ιστόγραμμα που προκύπτει αθροίζοντας όλα τα σημεία εικόνας σε κατακόρυφη διεύθυνση στην οριζόντια θέση  $x$ . Σε μία γραμμή κειμένου όπου οι χαρακτήρες χωρίζονται σωστά, η  $V(x)$  έχει μηδενικά ανάμεσα στους χαρακτήρες. Επομένως, αν οι χαρακτήρες είναι σε άριστη κατάσταση, δηλαδή διαχωρίζονται τέλεια μεταξύ τους και δεν έχουν σπασίματα, ο διαχωρισμός των χαρακτήρων μπορεί να επιτευχθεί κατευθείαν από τις κάθετες προβολές σε κάθε γραμμή κειμένου. Αν και στις περισσότερες περιπτώσεις το κείμενο δεν είναι σε καλή κατάσταση, εντούτοις οι κάθετες προβολές αποτελούν το βασικό εργαλείο για την κατάτμηση των χαρακτήρων.

Σύμφωνα με τον (Lu 1992), το πρώτο στάδιο αφορά την ομαδοποίηση των γραμμών ανάλογα με την μεταξύ τους απόσταση και το ύψος του γράμματος. Για παράδειγμα, στο σχήμα 6.2 έχουμε δύο ομάδες γραμμών κειμένου, η πρώτη γραμμή κειμένου ανήκει στην πρώτη ομάδα ενώ η δεύτερη και τρίτη γραμμή ανήκουν στην δεύτερη ομάδα.



**Σχήμα 6.2.** Ομαδοποίηση γραμμών κειμένου. Έχουμε δύο ομάδες γραμμών κειμένου (πρώτη γραμμή – δεύτερη και τρίτη γραμμή).

Στη συνέχεια υπολογίζονται οι κάθετες προβολές για κάθε ομάδα γραμμών. Εφόσον οι χαρακτήρες είναι ίδιου πλάτους, τα μηδενικά των προβολών θα αντιστοιχούν στα κενά ανάμεσα στα γράμματα για όλες τις γραμμές της ομάδας (σχήμα 6.3). Τα κενά που μπορεί να εμφανίζονται σε χαρακτήρες με σπασίματα δεν επηρεάζουν την κάθετη προβολή της ομάδας των γραμμών κειμένου. Αντίθετα, τυχών θόρυβος ανάμεσα στους χαρακτήρες μπορεί να εμποδίσει την εμφάνιση μηδενικών στην κάθετη προβολή. Για την αντιμετώπιση ανάλογων προβλημάτων, εξάγουμε στατιστικές τιμές για το πλάτος και την οριζόντια απόσταση των χαρακτήρων (σχήμα 6.4). Για να έχουμε ακριβή υπολογισμό, οι στατιστικές τιμές υπολογίζονται μόνο για περιοχές που αποτελούνται από περισσότερες της μίας γραμμής κειμένου. Για παράδειγμα, στο σχήμα 6.3 εξάγουμε στατιστικά υπολογίζοντας την κάθετη προβολή μέχρι και για το τέλος της δεύτερης μεγαλύτερης γραμμής.



**Σχήμα 6.3.** Εντοπισμός των χαρακτήρων με χρήση των κάθετων προβολών σε ομάδα γραμμών κειμένου.



**Σχήμα 6.4.** Χαρακτηριστικά γραμμής κειμένου με χαρακτήρες ίδιου πλάτους: Διάστημα ανάμεσα στους χαρακτήρες, κενό ανάμεσα στους χαρακτήρες, πλάτος χαρακτήρων.

Χρησιμοποιώντας τις στατιστικές τιμές για το πλάτος και την οριζόντια απόσταση των χαρακτήρων μπορούμε να διορθώσουμε το αποτέλεσμα ως εξής: (α) Αν από την κάθετη προβολή προκύπτουν γράμματα με πλάτος μεγαλύτερο από το μέσο πλάτος που έχουμε υπολογίσει προχωράμε σε διαίρεση της αντίστοιχης περιοχής. (β) Αν προκύπτουν γράμματα με μικρότερο πλάτος προχωράμε σε ενώσεις γειτονικών περιοχών μέχρι να προκύψει το μέσο πλάτος του γράμματος που έχουμε υπολογίσει. Αφού ολοκληρωθεί η παραπάνω διαδικασία, επαναλαμβάνω τον στατιστικό υπολογισμό για το πλάτος και την οριζόντια απόσταση των χαρακτήρων. Αν υπάρχουν ακόμα γράμματα με διαφορετικό πλάτος, τότε κάθε γραμμή της ομάδας οδηγείται σε αντίστοιχη διαδικασία τμηματοποίησης που αφορά μόνο μία γραμμή κειμένου.

Εναλλακτική μέθοδος είναι η μέθοδος εντοπισμού της περιοδικότητας των κενών ανάμεσα στα γράμματα (GPD: gap periodicity detection algorithm). Βασίζεται στο γεγονός ότι τα κενά ανάμεσα στα γράμματα εμφανίζονται σε περιοδικά διαστήματα όταν το πλάτος των γραμμάτων είναι σταθερό. Για κάθε ομάδα γραμμών κειμένου, υπολογίζουμε τις αντίστροφες κάθετες προβολές (προβολές των σημείων υποβάθρου) και βρίσκουμε τον βέλτιστο συνδυασμό του μήκους των κενών ανάμεσα στους χαρακτήρες και της απόστασης ανάμεσα στα κενά που ταιριάζουν καλύτερα στις αντίστροφες κάθετες προβολές. Αν δεν υπάρχει κατάλληλος συνδυασμός μήκους των κενών και της απόστασης ανάμεσα στα κενά ώστε να δίνει αντίστοιχα διαδοχικά μέγιστα στις αντίστροφες κάθετες προβολές, τότε δεν έχει γίνει σωστή ομαδοποίηση των γραμμών κειμένου ή οι γραμμές κειμένου δεν αποτελούνται από χαρακτήρες ομοίου πλάτους.



(α)



(β)

**Σχήμα 6.5.** Αντίστροφες προβολές γραμμών κειμένου. (α) Υπάρχει κατάλληλος συνδυασμός μήκους των κενών και της απόστασης ανάμεσα στα κενά ώστε να δίνει αντίστοιχα διαδοχικά μέγιστα. (β) Δεν υπάρχει αντίστοιχος συνδυασμός.

### Χαρακτήρες με σπασίματα

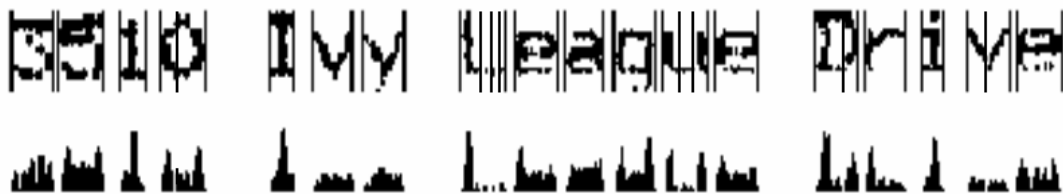
Για να αντιμετωπίσουμε τα σπασίματα που μπορεί να έχουν οι χαρακτήρες (όμοιου ή αναλογικού πλάτους) σε μία γραμμή κειμένου μπορούμε να ακολουθήσουμε την παρακάτω διαδικασία (Lu 1992):

Υπολογίσουμε την κάθετη προβολή της γραμμής κειμένου και στην συνέχεια τα στατιστικά για την απόσταση μεταξύ των χαρακτήρων, το πλάτος τους και το κενό μεταξύ τους. Στη συνέχεια γίνεται μία εκτίμηση για το αν η γραμμή κειμένου αποτελείται από χαρακτήρες ομοίου ή αναλογικού πλάτους. Η εκτίμηση αυτή γίνεται βάσει του αριθμού των περιοχών που παρουσιάζουν μεγάλη απόκλιση από τα στατιστικά για την απόσταση μεταξύ των χαρακτήρων, το πλάτος τους και το κενό μεταξύ τους.

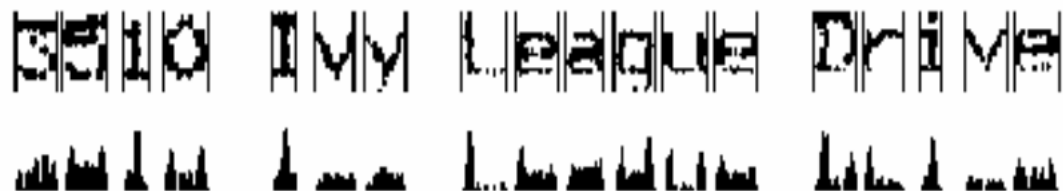
Αν η γραμμή κειμένου αποτελείται από χαρακτήρες ομοίου πλάτους, αρχίζουμε και ενώνουμε τις δύο γειτονικές περιοχές που έχουν το μικρότερο κενό ανάμεσα τους. Η διαδικασία αυτή επαναλαμβάνεται μέχρι τα στατιστικά που υπολογίζονται να είναι σε ένα αποδεκτό όριο.

Αν γραμμή κειμένου αποτελείται από χαρακτήρες αναλογικού πλάτους, τότε χρησιμοποιούμε μία διαδικασία συνενώσεων γειτονικών περιοχών με μικρό κενό ανάμεσά τους. Η διαδικασία συνενώσεως χρησιμοποιεί τα στατιστικά για την απόσταση μεταξύ των χαρακτήρων, το πλάτος τους και το κενό μεταξύ τους. Αν οι γειτονικές περιοχές έχουν μικρό πλάτος και το συνολικό τους πλάτος δεν απέχει πολύ από την στατιστικά υπολογισμένη τιμή για το μέσο πλάτος, τότε οι περιοχές αυτές είναι υποψήφιος για συνένωση. Στο σχήμα 6.6 φαίνεται ένα παράδειγμα της διαδικασίας συνένωσης. Οι χαρακτήρες «O», «L», «u», «D» και «v» έχουν χωριστεί σε παραπάνω τμήματα (oversegmentation) κάνοντας χρήση των κάθετων προβολών (σχήμα 6.6α) ενώ η τμηματοποίηση διορθώνεται με χρήση της διαδικασίας συνένωσης. Μέρος της διαδικασίας συνένωσης είναι και η διαδικασία ανάλυσης των χαρακτήρων η οποία χρησιμοποιείται για την αποφυγή της συνένωσης διαδοχικών χαρακτήρων μικρού πλάτους, όπως οι «ll «ti», «gi», «il», «ff». Η διαδικασία αυτή αναλύει δύο γειτονικές περιοχές και αποφασίζει αν ανήκουν σε ξεχωριστούς χαρακτήρες. Η ανάλυση βασίζεται στην σύγκριση με γνωστούς χαρακτήρες μικρού πλάτους χρησιμοποιώντας τα ύψη καθώς και στο ποσοστό κάλυψης των περιοχών. Η παραπάνω διαδικασία σύγκρισης έχει επιτυχία μόνο όταν επεξεργαζόμαστε χαρακτήρες καλής ποιότητας.

Για την συνένωση των χαρακτήρων με σπασίματα πολύ αποτελεσματική μέθοδος είναι η προεπεξεργασία με χρήση φίλτρων συρρίκνωσης και διόγκωσης (Schilling 2000).



(α)



(β)

**Σχήμα 6.6.** Συνένωση σπασμένων χαρακτήρων. (α) Οι χαρακτήρες έχουν χωριστεί σε παραπάνω τμήματα με χρήση της κάθετης προβολής. (β) Γειτονικές περιοχές έχουν ενωθεί έτσι ώστε να αποκατασταθούν οι χαρακτήρες με σπασίματα.

## Στραμμένοι και επικαλυπτόμενοι χαρακτήρες

Επικαλυπτόμενοι είναι οι χαρακτήρες που επικαλύπτονται με γειτονικούς π.χ. «TA», «PA». Τόσο στους επικαλυπτόμενους όσο και στους στραμμένους χαρακτήρες, οι κάθετες προβολές δεν δίνουν τις απαραίτητες πληροφορίες για σωστό διαχωρισμό (σχήμα 6.7). Για τον λόγο αυτό γίνεται μία δομική ανάλυση (structural analysis) της εικόνας (Lu 1992).



**Σχήμα 6.7.** Κάθετες προβολές σε γραμμή κειμένου που αποτελείται από επικαλυπτόμενους χαρακτήρες.

Σύμφωνα με την προσέγγιση της δομικής ανάλυσης αναλύονται τα γειτονικά συστατικά της εικόνας. Η μέθοδος αντιμετωπίζει την σωστή κατάτμηση των στραμμένων και επικαλυπτόμενων χαρακτήρων, των στοιχείων «.» «`» «;» «”» καθώς και των χαρακτήρων που αποτελούνται από πολλά τμήματα (π.χ. «i», «j»). Αποτελείται από δύο βασικά βήματα: την εύρεση των συνδεδεμένων συστατικών και την ανάλυση της δομής των γειτονικών συστατικών. Η ανάλυση της δομής αποτελείται από δύο στάδια: Την ανάλυση των στραμμένων και επικαλυπτόμενων χαρακτήρων και τον εντοπισμό των συμβόλων «.» «`» «;» «”». Η ανάλυση των στραμμένων και επικαλυπτόμενων χαρακτήρων επιτυγχάνει την συνένωση των χαρακτήρων με σπασίματα καθώς και των χαρακτήρων που αποτελούνται από πολλά τμήματα ενώ ταυτόχρονα οι στραμμένοι ή επικαλυπτόμενοι χαρακτήρες δεν ενώνονται μεταξύ τους. Κάθε ζευγάρι γειτονικών συνδεδεμένων συστατικών εξετάζονται για κατακόρυφη επικάλυψη. Αν η επικάλυψη ξεπερνάει κάποιο όριο τότε το ζευγάρι των συνδεδεμένων συστατικών είναι υποψήφιο για συνένωση. Στη συνέχεια εξετάζουμε αν το δεύτερο συνδεδεμένο συστατικό είναι ένα από τα σύμβολα «.» «`» «;» «”». Υπολογίζουμε το ποσοστό κάλυψης, το μέγεθος, τον λόγο ύψους προς πλάτος, την απόσταση από το γειτονικό συστατικό καθώς και από το κάτω και πάνω μέρος του ορίου της γραμμής κειμένου. Αν το συνδεδεμένο συστατικό είναι ένα από τα σύμβολα που εξετάζουμε τότε δεν ενώνεται με το γειτονικό του. Ένα παράδειγμα της παραπάνω διαδικασίας δίδεται στο σχήμα 6.8.



**Σχήμα 6.8.** Εντοπισμός των χαρακτήρων χρησιμοποιώντας την μέθοδο που βασίζεται στην δομική ανάλυση της εικόνας.

## Ενωμένοι χαρακτήρες

Οι ενωμένοι χαρακτήρες είναι συνδεδεμένα συστατικά της εικόνας τα οποία ανήκουν σε περισσότερους του ενός χαρακτήρες. Ο σωστός διαχωρισμός των παραπάνω συστατικών αποτελεί ένα από τα δυσκολότερα προβλήματα στην κατάτμηση των χαρακτήρων. Για την αντιμετώπιση του προβλήματος των ενωμένων χαρακτήρων θα πρέπει να ακολουθήσουμε τα παρακάτω βήματα:

- Να εντοπίσουμε τα συνδεδεμένα συστατικά τα οποία περιέχουν δύο ή παραπάνω χαρακτήρες.
- Να βρούμε τα σημεία διαχωρισμού ώστε να προκύψουν οι μεμονωμένοι χαρακτήρες.

Το πιο συνηθισμένο κριτήριο εντοπισμού των ενωμένων χαρακτήρων είναι τα όρια στο πλάτος αλλά και στον λόγο ύψους προς πλάτος των χαρακτήρων. Το μέσο πλάτος των χαρακτήρων υπολογίζεται στατιστικά. Κάθε υποψήφια περιοχή ενωμένων χαρακτήρων εξετάζεται ως προς την απόσταση από το μέσο πλάτος των χαρακτήρων αλλά και ως προς τον λόγο ύψους προς πλάτος. Οι περισσότεροι χαρακτήρες έχουν πλάτη μικρότερα από τα ύψη τους επίσης έχουν πλάτη μικρότερα από το διπλάσιο μέσο πλάτος των χαρακτήρων του κειμένου. Ο συνδυασμός των δύο παραπάνω κριτηρίων λειτουργεί στην πλειοψηφία των περιπτώσεων για τον εντοπισμό των ενωμένων χαρακτήρων. Όπως φαίνεται στο παράδειγμα του σχήματος 6.9α, οι ενωμένοι χαρακτήρες «In», «gr» και «ed» έχουν μεγαλύτερο πλάτος από το μέσο πλάτος των χαρακτήρων της εικόνας και έτσι μπορούμε να τους εντοπίσουμε χρησιμοποιώντας το κριτήριο του πλάτους ή του λόγου ύψους προς πλάτος. Στο παράδειγμα όμως του σχήματος 6.9β, οι ενωμένοι χαρακτήρες «tt» έχουν μικρότερο πλάτος από τους χαρακτήρες «J» και «w», ενώ στο παράδειγμα του σχήματος 6.9γ, οι ενωμένοι χαρακτήρες «LI» έχουν μικρότερο πλάτος από του χαρακτήρες «N» και «G». Επίσης, ο λόγος ύψους προς πλάτος των συμπλεγμάτων «tt» και «LI» δεν διαφέρει από τον λόγο ύψους προς πλάτος των μεμονωμένων χαρακτήρων. Για τον εντοπισμό των παραπάνω προβλημάτων, χρησιμοποιούνται τα μοντέλα των περιγραμμάτων των κάθετων προβολών για τους μεμονωμένους και τους ενωμένους χαρακτήρες. Στο σχήμα 6.10 δίδονται παραδείγματα περιγραμμάτων κάθετων προβολών για μεμονωμένους και ενωμένους χαρακτήρες.

**Incorporated**

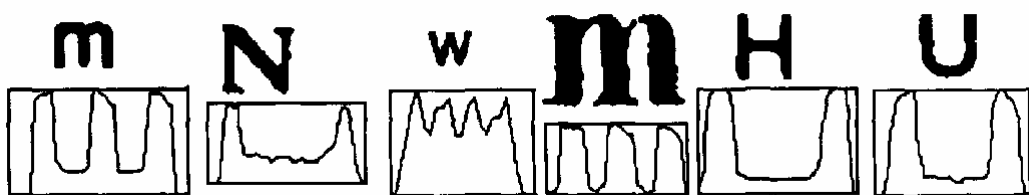
(α)

**Jewett BILLING**

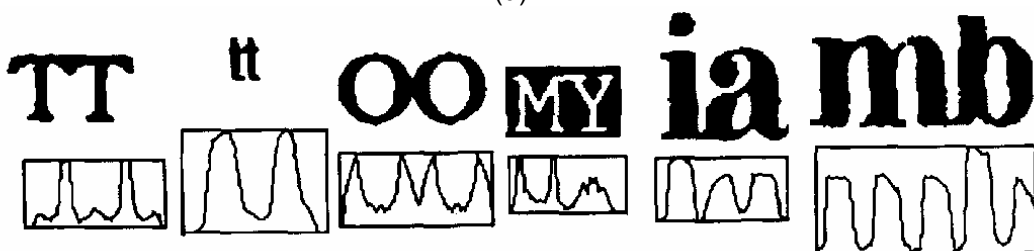
(β)

(γ)

Σχήμα 6.9. Παραδείγματα κειμένων με ενωμένους χαρακτήρες.



(α)



(β)

Σχήμα 6.10. Παραδείγματα περιγραμμάτων κάθετων προβολών για μεμονωμένους χαρακτήρες (α) ή συμπλέγματα χαρακτήρων (β).

## Β. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

Η κάθετη προβολή των ενωμένων χαρακτήρων έχει τουλάχιστον δύο τοπικά μέγιστα. Όμως, και οι χαρακτήρες «H», «h», «N», «u» και «U» έχουν δύο τοπικά μέγιστα στις κάθετες προβολές ενώ οι χαρακτήρες «m», «w» και «M» έχουν τρία τοπικά μέγιστα. Για να διαχωρίσουμε τους χαρακτήρες που έχουν δύο ή περισσότερα τοπικά μέγιστα στις κάθετες προβολές από τους ενωμένους χαρακτήρες, δημιουργούμε μοντέλα με τα περιγράμματα των κάθετων προβολών των χαρακτήρων αυτών. Τα χαρακτηριστικά των κάθετων προβολών που χρησιμοποιούμε για την κατασκευή των μοντέλων είναι:

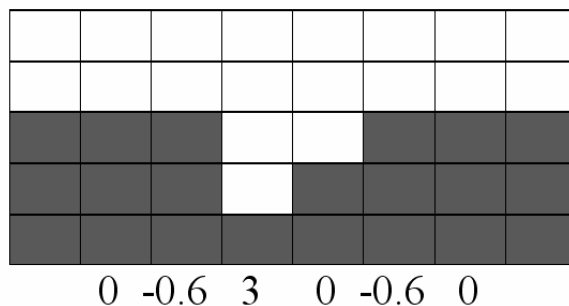
- Ο αριθμός των κορυφών
- Το ύψος και το άπλωμα των κορυφών
- Το πλάτος, το βάθος και η ομαλότητα των κοιλάδων
- Η οριζόντια συμμετρία

Συγκρίνοντας με τα παραπάνω μοντέλα μπορούμε να αποφασίζουμε εάν ένα συνδεδεμένο συστατικό ανήκει σε έναν ή πολλούς χαρακτήρες. Για τον εντοπισμό των σημείων διαχωρισμού των συνδεδεμένων συστατικών που αντιστοιχούν σε ενωμένους χαρακτήρες, μπορούμε να χρησιμοποιήσουμε τις παρακάτω τεχνικές:

α. Μεγιστοποίηση της συνάρτησης κατάτμησης (Kahan 1987):

$$\frac{V(x-1) - 2V(x) + V(x+1)}{V(x)} \quad (6.1)$$

όπου  $V(x)$  είναι η κάθετη προβολή. Η συνάρτηση κατάτμησης είναι η δεύτερη παράγωγος της συνάρτησης της κάθετης προβολής. Ένα παράδειγμα εύρεσης του σημείου διαχωρισμού με μεγιστοποίηση της συνάρτησης κατάτμησης δίνεται στο σχήμα 6.11.



**Σχήμα 6.11.** Εύρεση του σημείου διαχωρισμού με μεγιστοποίηση της συνάρτησης κατάτμησης. Η μέγιστη τιμή της συνάρτησης κατάτμησης είναι 3 και αντιστοιχεί στο τέταρτο στοιχείο της κάθετης προβολής της εικόνας.

β. Μεγιστοποίηση της συνάρτησης λόγου κορυφής-κοιλάδας (Lu 1992): Η κάθετη προβολή της περιοχής προβάλλεται σε μια νέα συνάρτηση που ονομάζεται λόγος κορυφής-κοιλάδας (peak-to-valley ratio) εφαρμόζοντας μία τεχνική γενικευμένης διαφοροποίησης. Παράδειγμα αυτής της συνάρτησης δίνεται στο σχήμα 6.12. Απότομα ελάχιστα στην κάθετη προβολή δημιουργούν μέγιστα στην συνάρτηση λόγου κορυφής-κοιλάδας (σχήμα 6.13). Αυτά τα μέγιστα θεωρούνται σαν πιθανά σημεία διαχωρισμού των ενωμένων χαρακτήρων. Η τελική απόφαση για τα σημεία διαχωρισμού λαμβάνει υπόψη και το μέσο πλάτος των χαρακτήρων. Σύμφωνα με την διαδικασία διαχωρισμού, πρώτα υπολογίζεται μία περιοχή πιθανού διαχωρισμού σύμφωνα με το μέσο πλάτος των χαρακτήρων. Το μέγιστο της συνάρτησης λόγου κορυφής-κοιλάδας μέσα στην περιοχή αυτή είναι το σημείο διαχωρισμού. Στην συνέχεια, εξετάζονται τα πλάτη των δύο υποπεριοχών που προκύπτουν. Αν κάθε υποπεριοχή έχει επιτρεπτό μήκος τότε θεωρείται ότι περιέχει έναν μόνο χαρακτήρα. Αν το πλάτος της υποπεριοχής επιτρέπει την ύπαρξη πολλών χαρακτήρων, η διαδικασία αναζήτησης σημείου

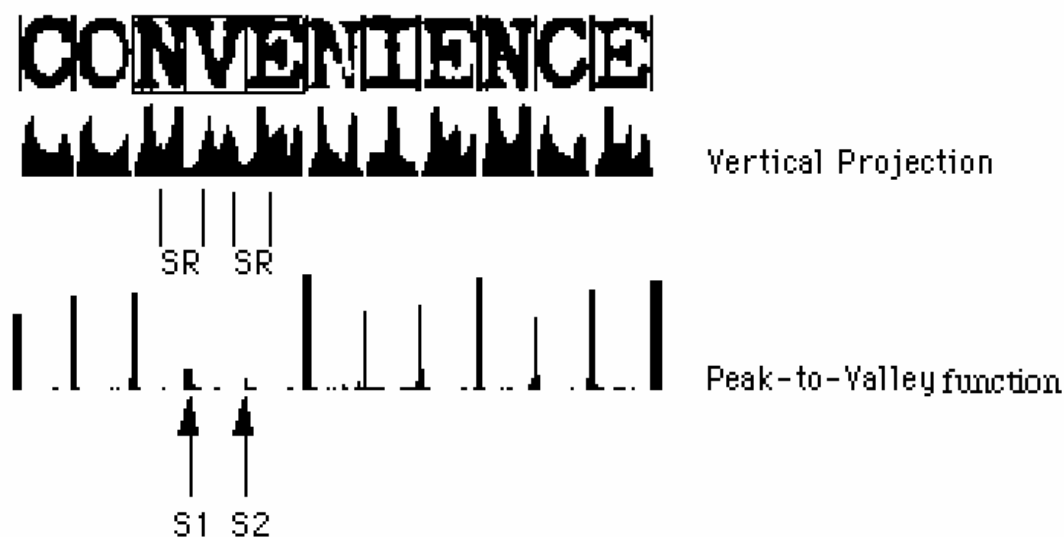
διαχωρισμού επαναλαμβάνεται. Αν δεν εμφανίζονται μέγιστα στις εξεταζόμενες περιοχές, τότε η διαδικασία εύρεσης σημείου διαχωρισμού σταματάει. Στο σχήμα 6.14 δίδεται ένα παράδειγμα κατάτμησης. Εμφανίζονται δύο επιτρεπόμενες περιοχές κατάτμησης  $SR$  βάσει του μέσου πλάτους των χαρακτήρων που αντιστοιχούν σε δύο μέγιστα  $S1$  και  $S2$ .



Σχήμα 6.12. Δημιουργία συνάρτησης λόγου κορυφής-κοιλιάδας (peak-to-valley ratio) από την κάθετη προβολή της περιοχής των ενωμένων χαρακτήρων.



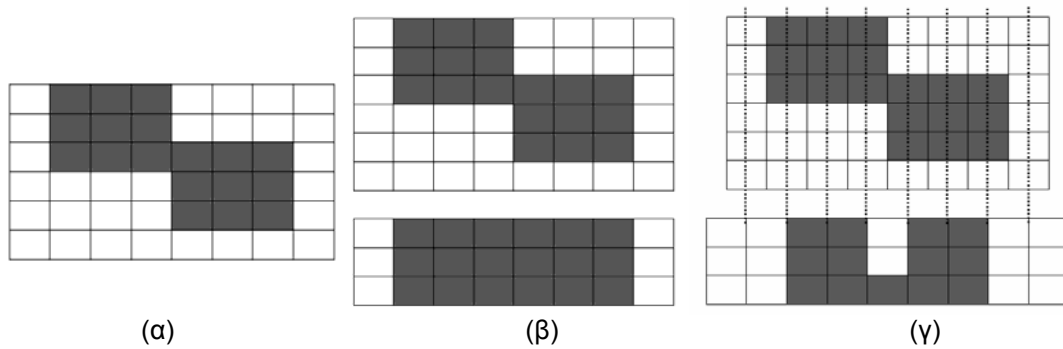
Σχήμα 6.13. Πιθανά σημεία διαχωρισμού όπως προκύπτουν από την συνάρτηση λόγου κορυφής-κοιλιάδας (peak-to-valley ratio).



Σχήμα 6.14. Κατάτμηση των χαρακτήρων βάσει της μεγιστοποίησης του λόγου κορυφής-κοιλιάδας.  $SR$  είναι οι επιτρεπόμενες περιοχές κατάτμησης και  $S1$ ,  $S2$  τα εντοπισμένα μέγιστα στις περιοχές  $SR$ .



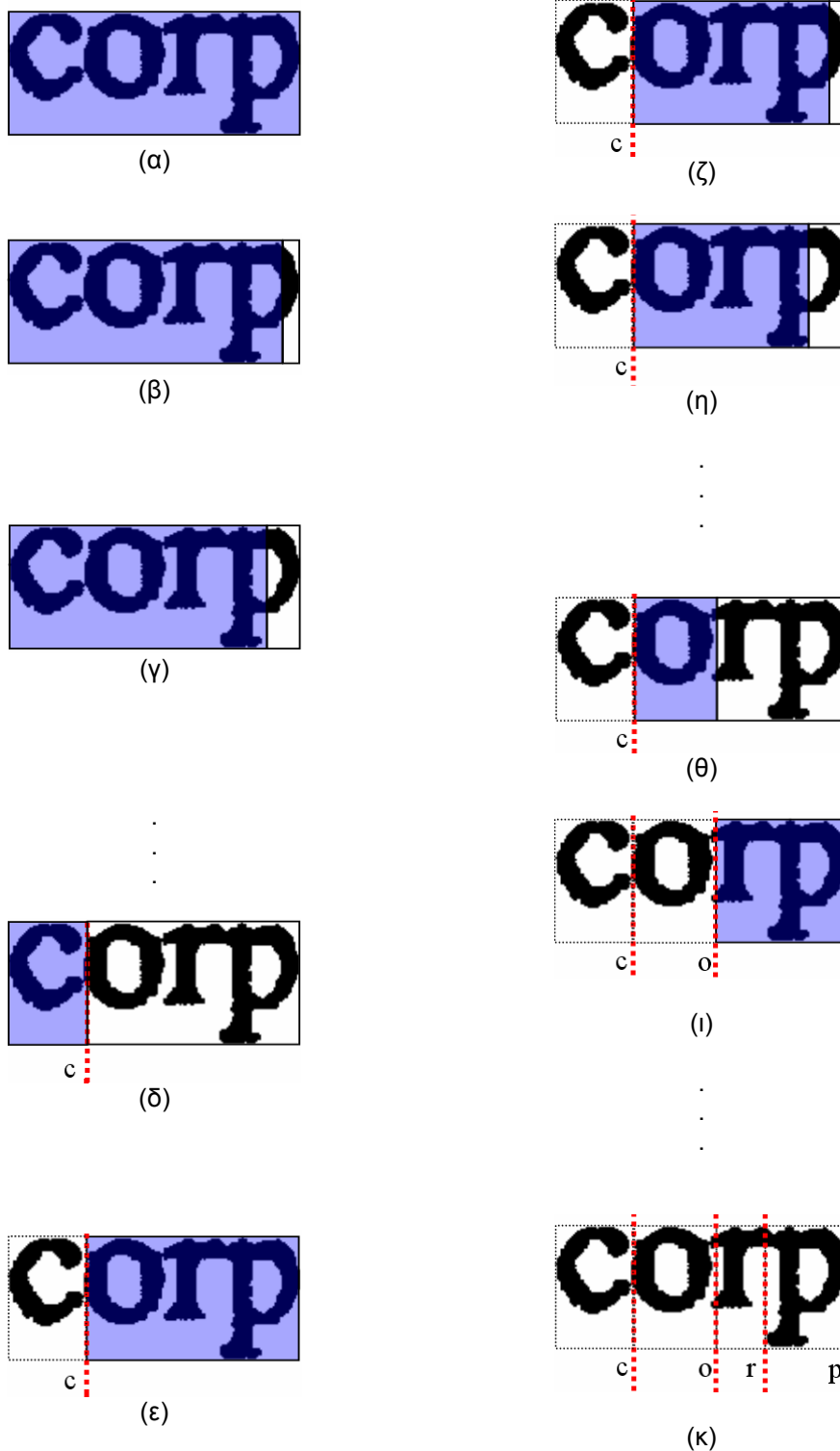
γ. Ελαχιστοποίηση της συνάρτησης κόστους διάσπασης (Tsujimoto 1991): Η συνάρτηση κόστους διάσπασης (break cost function) προτάθηκε για τον εντοπισμό του βαθμού ένωσης μιας περιοχής. Ορίζεται για κάθε θέση ανάμεσα στις γειτονικές στήλες της περιοχής και υπολογίζεται αθροίζοντας τα μαύρα pixels της εικόνας τα οποία έχουν προκύψει από εφαρμογή της πράξης AND μεταξύ γειτονικών στηλών. Τα τελικά υποψήφια σημεία διαχωρισμού προκύπτουν εντοπίζοντας τα τοπικά ελάχιστα της συνάρτησης κόστους διάσπασης. Η συνάρτηση αυτή αποδεικνύεται πιο αποτελεσματική για τον εντοπισμό των σημείων διαχωρισμού σε σχέση με τις κάθετες προβολές. Ένα παράδειγμα δίδεται στο σχήμα 6.15. Η συνάρτηση κόστους διάσπασης παρουσιάζει τοπικό ελάχιστο σε αντίθεση με τις κάθετες προβολές.



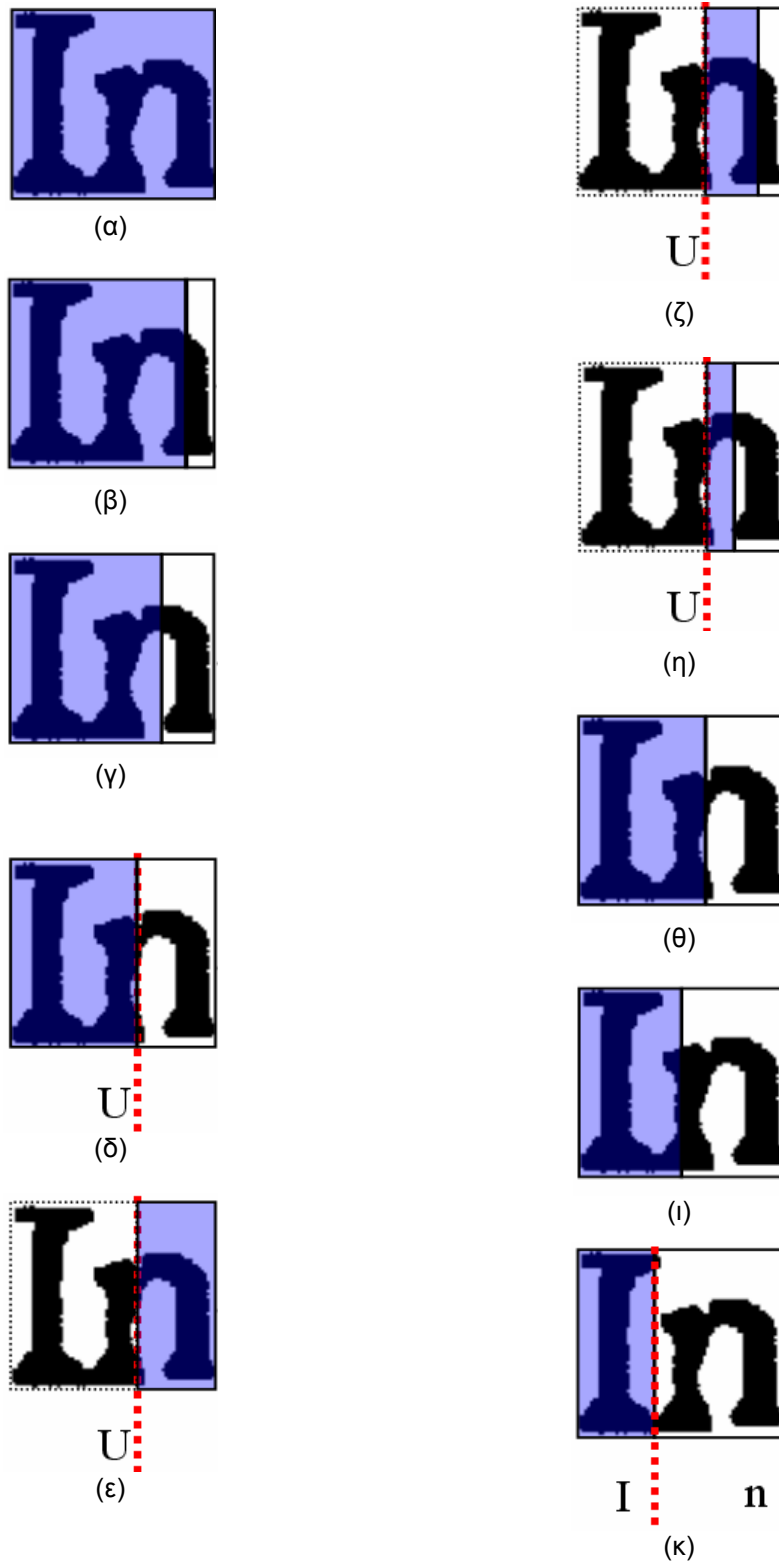
**Σχήμα 6.15.** Υπολογισμός της συνάρτησης κόστους διάσπασης. (α) Αρχική εικόνα. (β) Κάθετες προβολές. (γ) Συνάρτηση κόστους διάσπασης.

δ. Βάσει αναγνώρισης (Casey 1982): Ένα συνδεδεμένο συστατικό εικόνας διασπάται έτσι ώστε να προκύψει σύνολο χαρακτήρων που αναγνωρίζονται με ασφάλεια. Το συνδεδεμένο συστατικό οδηγείται στην είσοδο του συστήματος επίβλεψης της κατάτμησης (Segmentation Supervisor - SS). Το παράθυρο αναγνώρισης αρχικοποιείται στο πλήρες πλάτος του συνδεδεμένου συστατικού έτσι ώστε αν αποτελείται από έναν χαρακτήρα να αναγνωριστεί με ένα πέρασμα. Αν το σύστημα αναγνώρισης απορρίπτει τον χαρακτήρα, τότε το παράθυρο αναγνώρισης μικραίνει από το δεξί μέρος και το σύστημα αναγνώρισης εφαρμόζεται στο νέο παράθυρο. Αυτή η διαδικασία επαναλαμβάνεται μέχρι να αναγνωριστεί με επιτυχία ο χαρακτήρας στο παράθυρο αναγνώρισης ή το πλάτος του παραθύρου έχει μικρύνει τόσο ώστε αποκλείεται να περιέχει κάποιο χαρακτήρα. Αν το πρότυπο μέσα στο παράθυρο αναγνωρίζεται με ασφάλεια, το σύστημα καταγράφει τον κωδικό του χαρακτήρα και το σημείο αποκοπής από δεξιά. Στη συνέχεια, το παράθυρο αρχικοποιείται στο υπόλοιπο δεξί μέρος της εικόνας και η παραπάνω διαδικασία επαναλαμβάνεται. Η διαδικασία κατάτμησης τερματίζεται αν το παράθυρο της τελευταίας επανάληψης αναγνωριστεί χωρίς να γίνει αποκοπή από δεξιά ή έχει πολύ μικρό πλάτος. Στην περίπτωση που δεν μπορεί να αναγνωριστεί με ασφάλεια το παράθυρο για καμία θέση του δεξιού ορίου, αν είμαστε στην πρώτη επανάληψη τότε το συνδεδεμένο συστατικό απορρίπτεται. Αλλιώς, θεωρούμε την προηγούμενη αναγνώριση λανθασμένη και αρχικοποιούμε το παράθυρο στο αντίστοιχο όπου έγινε η λανθασμένη αναγνώριση. Στη συνέχεια, αρχίζουμε και στενεύουμε το παράθυρο από το δεξί μέρος και προσπαθούμε να αναγνωρίσουμε ξανά τον χαρακτήρα του παραθύρου. Η παραπάνω διαδικασία τερματίζει με επιτυχία μόνο αν το συνδεδεμένο συστατικό χωρίζεται σε επιμέρους κομμάτια τα οποία αναγνωρίζονται με ασφάλεια.

Στα σχήματα 6.16 και 6.17 δίνονται δύο παραδείγματα εφαρμογής της παραπάνω μεθοδολογίας για τον εντοπισμό των σημείων διαχωρισμού των ενωμένων χαρακτήρων. Στο πρώτο παράδειγμα, επιλέγουμε τα σωστά σημεία διαχωρισμού εντοπίζοντας κάθε φορά το σωστό αποτέλεσμα αναγνώρισης. Στο δεύτερο παράδειγμα, η μέθοδος φθάνει σε λανθασμένο αποτέλεσμα (σχήμα 6.17δ). Το γεγονός αυτό το εντοπίζουμε γιατί το υπόλοιπο συνδεδεμένο συστατικό δεν μπορεί να αναγνωριστεί. Η διόρθωση γίνεται με ακύρωση του αποτελέσματος και αρχικοποίηση του παραθύρου στο αντίστοιχο όπου έγινε η λανθασμένη αναγνώριση. Αν συνεχίσουμε την διαδικασία μικραίνοντας το παράθυρο αναγνώρισης από δεξιά καταλήγουμε στο σωστό αποτέλεσμα διαχωρισμού και αναγνώρισης του συνδεδεμένου συστατικού.



**Σχήμα 6.16.** Βήματα διαχωρισμού γραμμάτων βάσει αναγνώρισης. (α)-(δ) Το παράθυρο αναγνώρισης μικραίνει από δεξιά μέχρι να πετύχουμε αναγνώριση. (ε) Το παράθυρο αναγνώρισης αρχικοποιείται στο υπόλοιπο μέρος του συνδεδεμένου συστατικού. (ζ)-(θ) Το παράθυρο αναγνώρισης μικραίνει από δεξιά μέχρι να πετύχουμε αναγνώριση. (ι) Το παράθυρο αναγνώρισης αρχικοποιείται στο υπόλοιπο μέρος του συνδεδεμένου συστατικού. (κ) Το τελικό αποτέλεσμα της κατάτμησης.



**Σχήμα 6.17.** Διόρθωση λανθασμένης αναγνώρισης κατά τον διαχωρισμό γραμμάτων βάσει αναγνώρισης. (α)-(δ) Το παράθυρο αναγνώρισης μικραίνει από δεξιά μέχρι να πετύχουμε αναγνώριση. Η αναγνώριση αυτή είναι λανθασμένη. (ε) Το παράθυρο αναγνώρισης αρχικοποιείται στο υπόλοιπο μέρος του συνδεδεμένου συστατικού. (ζ)-(η) Το παράθυρο αναγνώρισης μικραίνει από δεξιά χωρίς να μπορούμε να πετύχουμε αναγνώριση. (θ) Θεωρούμε την προηγούμενη αναγνώριση λανθασμένη και αρχικοποιούμε το παράθυρο στο αντίστοιχο όπου έγινε η λανθασμένη αναγνώριση. (ι), (κ) Το παράθυρο αναγνώρισης μικραίνει από δεξιά μέχρι να πετύχουμε την σωστή αναγνώριση.

## 6.2 Εντοπισμός χαρακτήρων σε χειρόγραφο κείμενο

Για τον διαχωρισμό των χαρακτήρων σε χειρόγραφο κείμενο, η χρήση των περισσότερων μεθοδολογιών που ήδη αναφέρθηκαν και χρησιμοποιούνται σε τυπωμένο κείμενο (κάθετες προβολές, ανάλυση των συνδεδεμένων συστατικών, συνάρτηση κόστους διάσπασης) δεν είναι αποτελεσματική. Οι βασικές τεχνικές που χρησιμοποιούνται είναι η ανάλυση του περιγράμματος και η επεξεργασία του σκελετού του τμήματος των ενωμένων χαρακτήρων. Ακολουθούν δύο παραδείγματα εφαρμογών διαχωρισμού χειρόγραφων χαρακτήρων που χρησιμοποιούν τις παραπάνω τεχνικές.

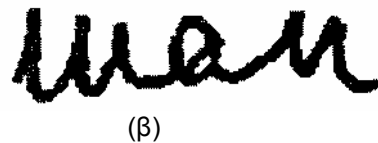
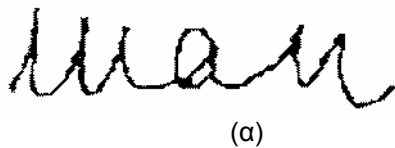
### Ανάλυση του περιγράμματος

Η μέθοδος αυτή αφορά την ανάλυση του περιγράμματος της λέξης ώστε να εντοπιστούν υποψήφιες διαδρομές κατάτμησης οι οποίες επιβεβαιώνονται με την χρήση λεξικού (Bozinovic 1989). Τα διάφορα στάδια της μεθόδου είναι τα ακόλουθα:

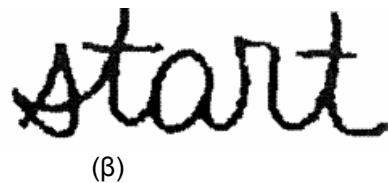
**ΒΗΜΑ 1:** Εξομάλυνση της επιφάνειας της λέξης με χρήση φίλτρου διόγκωσης (swell filter) (Schilling 2000). Παράδειγμα δίδεται στο σχήμα 6.18.

**ΒΗΜΑ 2:** Εντοπισμός και διόρθωση της κλίσης της λέξης (Bozinovic 1989). Παράδειγμα δίδεται στο σχήμα 6.19.

**ΒΗΜΑ 3:** Εντοπισμός των τριών ζωνών (πάνω, μεσαία, κάτω) από τις οποίες αποτελείται η λέξη (σχήμα 6.20). Τα όρια μεταξύ των τριών ζωνών αντιστοιχούν στα τοπικά ελάχιστα της οριζόντιας προβολής της λέξης.



**Σχήμα 6.18.** Εξομάλυνση της επιφάνειας της λέξης με χρήση φίλτρου διόγκωσης. (α) Αρχική εικόνα. (β) Εικόνα μετά την εξομάλυνση



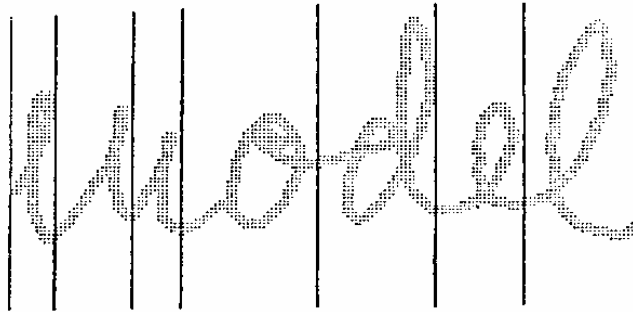
**Σχήμα 6.19.** Διόρθωση της κλίσης της λέξης. (α) Αρχική εικόνα. (β) Εικόνα μετά την διόρθωση της κλίσης.



**Σχήμα 6.20.** Οι τρεις ζώνες από τις οποίες αποτελείται η λέξη.

**ΒΗΜΑ 4:** Ακολουθήση του περιγράμματος της λέξης (Pavlidis 1981).

**ΒΗΜΑ 5:** Εύρεση αρχικών τμημάτων κατάτμησης (presegments). Στο στάδιο αυτό αναζητούνται κάθετες γραμμές οι οποίες δίνουν ένα πρώτο αποτέλεσμα κατάτμησης της λέξης. Τα τμήματα που θα προκύψουν θέλουμε να είναι όσο το δυνατόν μικρότερα ώστε να περιλαμβάνονται όλοι οι συνδυασμοί χαρακτήρων. Για παράδειγμα στο σχήμα 6.21, το πρώτο γράμμα της λέξης «model» μπορεί να είναι «m» ή «ui». Θέλουμε να υπάρχουν όλες οι δυνατές θέσεις του διαχωριστικού κάθετου ορίου που αντιστοιχούν σε δυνατά γράμματα.



**Σχήμα 6.21.** Εύρεση αρχικών τμημάτων κατάτμησης (presegments).

Ο αλγόριθμος αναζητά τοπικά ελάχιστα του κάτω περιγράμματος της λέξης στην μεσαία ζώνη της λέξης:

```

presegment
  trace lower contour left-to-right
  and produce a chain code sequence (chain);
  start with first chain element, cc;
  repeat
    if (cc is at local min) then
      generate-PSP-left;
      generate-PSP-right;
      if (no PSP generated in this iteration) then
        emergency-generate-PSP-right;
      move to next chain element cc;
  until chain exhausted;
  compact-PSP;
    
```

όπου PSP είναι τα αρχικά σημεία κατάτμησης (Presegment Points) τα οποία ορίζονται από την  $y$ -συντεταγμένη τους. Η διαδικασία δημιουργίας σημείου κατάτμησης από αριστερά (PSP-left) είναι η ακόλουθη:

```

generate-PSP-left
  singlerun turned off;
  set cy to  $y_0$ ;
  repeat
    if (((runs(y) > 1)  $\vee$  (density(y) > md)) & singlerun) then
      exit loop;
    else if ((runs(y) = 1) & (density(y)  $\leq$  md) & (not singlerun)) then
      set singlerun;
      decrease cy by 1;
      until (cy = 0  $\vee$  cy < last PSP);
      if singlerun then
        set PSP to middle value y of singlerun stretch;
    
```

όπου  $y_0$  είναι η τρέχουσα τιμή του  $y$ , *runs* είναι ο αριθμός των κατακόρυφων τμημάτων που αποτελούνται από διαδοχικά 1, *density* είναι η κάθετη προβολή και *md* είναι η μέγιστη τιμή της κάθετης προβολής που αντιστοιχεί στο πάχος της γραμμής του χαρακτήρα. Η διαδικασία δημιουργίας σημείου κατάτμησης από δεξιά (PSP-right) είναι αντίστοιχη με τις παρακάτω αλλαγές:

## B. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

- Η επανάληψη τερματίζεται όταν  $(y - y_0 < ms)$
- Το  $cy$  αυξάνεται αντί να μειώνεται

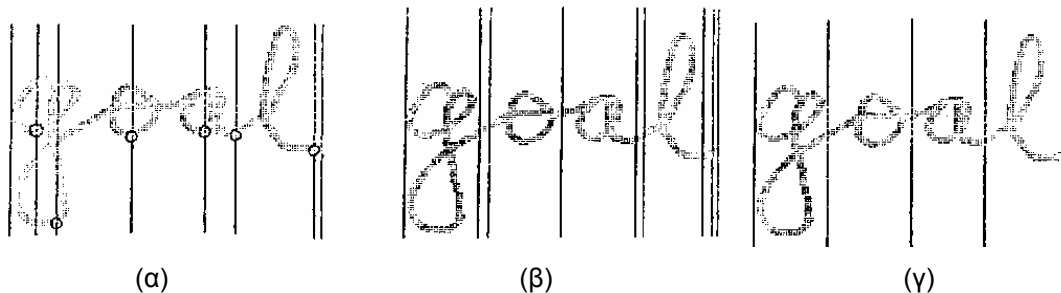
όπου  $ms$  είναι η μέγιστη μετακίνηση προς τα δεξιά. Η εναλλακτική διαδικασία δημιουργίας σημείου κατάτμησης από δεξιά είναι η εξής:

```
emergency-generate-PSP-right  
set  $cy$  to  $y_0$ ;  
for ( $cy$  from  $y_0$  to  $y_0 + ms$ ) do  
  find  $cy$  with lowest density ( $cy$ );
```

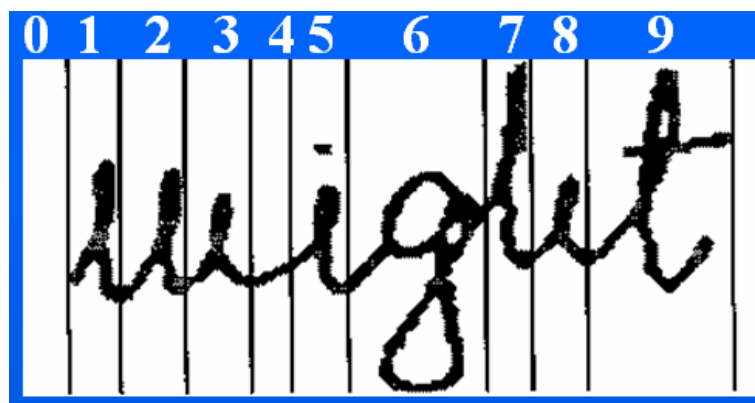
Τέλος, η διαδικασία επιλογής των σημείων είναι η εξής:

```
compact-PSP  
set  $cps$  to first(leftmost) PSP;  
for current  $cps$  do  
  for set {  $cps$ , all  $cps$ 's such that  $cps' - cps < marg$  }  
    find element  $mcp$  with minimum density ( $mcp$ );  
  
  if (first iteration) then  
    output  $cps$ ;  
  set  $cps$  to first  $cps'$  such that  $cps' - cps \geq marg$ ;  
  if ( $cps \neq$  last PSP) & (not first iteration) then  
    output  $mcp$ ;  
output last PSP;
```

όπου  $marg$  είναι το πλάτος του παραθύρου επιλογής. Παράδειγμα εύρεσης των αρχικών σημείων κατάτμησης δίδονται στα σχήματα 6.22 και 6.23.



**Σχήμα 6.22.** Στάδια εύρεσης αρχικών τμημάτων κατάτμησης (presegments). (α) Τα τοπικά ελάχιστα στο κάτω περίγραμμα του χαρακτήρα. (β) Τα εντοπισμένα τμήματα πριν την τελική επιλογή των σημείων. (γ) Τα εντοπισμένα τμήματα μετά την τελική επιλογή των σημείων.



**Σχήμα 6.23.** Εύρεση 10 αρχικών τμημάτων κατάτμησης (presegments) για την λέξη «might».

## Β. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

**ΒΗΜΑ 6:** Εντοπισμός λέξεων. Ξεκινώντας από αριστερά προς τα δεξιά της λέξης (από το τμήμα 0) δημιουργούμε μία λίστα πιθανών γραμμάτων. Κάθε εισαγωγή στην λίστα έχει μία τιμή στο διάστημα (-1,1) ανάλογα με τον βαθμό εμπιστοσύνης της αναγνώρισης και τον αριθμό των τμημάτων που καλύπτει. Οι εισαγωγές ταξινομούνται ως προς τον μεγαλύτερο βαθμό εμπιστοσύνης. Για το παράδειγμα του σχήματος 6.23, η λίστα είναι της μορφής:

(( m 0.84 3) (n 0.73 2) (u 0.73 2) (w 0.60 3) (w 0.51 2) (c 0.30 1) (j 0.30 1) ... (f -0.74 1))

Η λίστα αυτή οδηγείται για έλεγχο στο λεξικό αφού το πρώτο (καλύτερο) αποτέλεσμα (m 0.84 3) αντικατασταθεί με το συνδυασμό του με το επόμενο γράμμα. Η λίστα υπόθεσης μετά το τμήμα 3 είναι η ακόλουθη:

(( ' 0.60 1) (u -0.46 2) (y 0.66 2) (v -0.73 2) (n -0.74 2) (h -0.82 2) (k -0.83 2) ... (m -0.93 3))

Κάθε αποτέλεσμα της παραπάνω λίστας συνδυάζεται με το γράμμα m, το δίγραμμο που προκύπτει ελέγχεται από λεξικό και στη συνέχεια γίνεται ξανά υπολογισμός του βαθμού εμπιστοσύνης για κάθε δίγραμμο. Όλα τα αποτελέσματα αντικαθιστούν την εισαγωγή (m 0.84 3) στην προηγούμενη λίστα η οποία ταξινομείται ξανά ως προς τον μεγαλύτερο βαθμό εμπιστοσύνης:

(( (n 0.73 2 2) (u 0.73 2 2) (m' 0.72 4(3 1)) (w 0.60 3 3) (w 0.51 2 2) (c 0.30 1 1) (j 0.30 1 1) ... (s 0.25 1 1) (e 0.21 1 1) (mu 0.19 5(3 2)) ...

όπου μία εισαγωγή της μορφής (m' 0.72 4(3 1)) σημαίνει ότι το δίγραμμο «m'» έχει βαθμό εμπιστοσύνης 0.72 και συνολικό μήκος 4 τμημάτων από τα οποία τα 3 πρώτα αφορούν το γράμμα «m» και το επόμενο 1 τον χαρακτήρα «'». Βλέπουμε πως στην δεύτερη επανάληψη η σωστή λύση «m'» έχει πέσει στην τρίτη καλύτερη θέση της λίστας. Αν συνεχίσουμε τις επαναλήψεις, κατά την τέταρτη επανάληψη η σωστή λύση βρίσκεται στην πρώτη θέση:

(( (m'ig 0.77 6(3 1 1 1)) (un 0.71 4(2 2)) (nu 0.71 4(2 2)) ... (mu 0.19 5 (3 2)))

για να καταλήξουμε μετά από 20 επαναλήψεις στην βέλτιστη λύση:

((m'ight 0.48 9(3 1 1 2 1)), ...

### Επεξεργασία του σκελετού

Μία χαρακτηριστική εφαρμογή χρήσης του σκελετού για τον διαχωρισμό των χειρόγραφων χαρακτήρων δίδεται από τον (Chen 2000) και αφορά τον διαχωρισμό ενωμένων χειρόγραφων αριθμών. Τα βασικά βήματα του προτεινόμενου αλγόριθμου είναι τα ακόλουθα:

**ΒΗΜΑ 1:** Εντοπισμός εσωτερικού και εξωτερικού σκελετού του συνδεδεμένου συστατικού. (σχήμα 6.24)



(α)



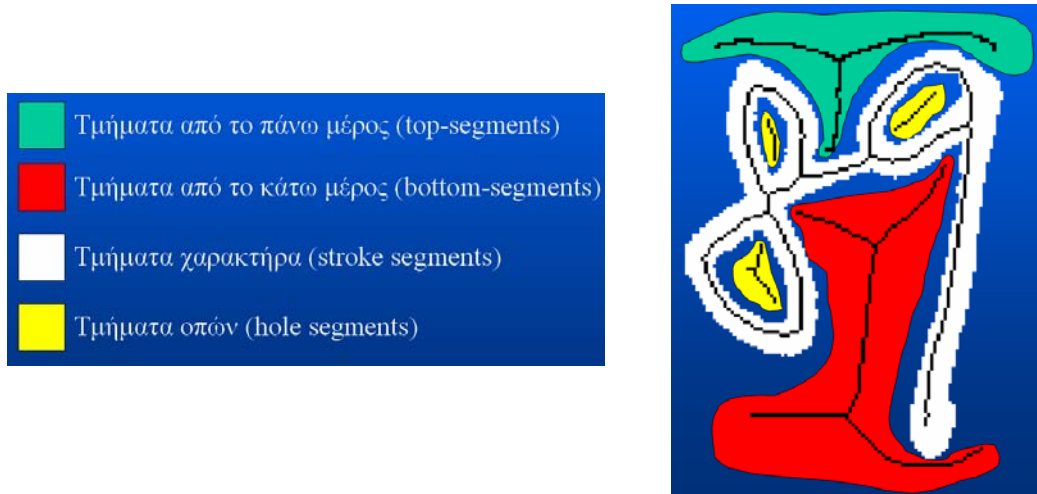
(β)

**Σχήμα 6.24.** Εσωτερικός και εξωτερικός σκελετός των ενωμένων χαρακτήρων. (α) Αρχική εικόνα. (β) Ο εσωτερικός και εξωτερικός σκελετός.

**ΒΗΜΑ 2:** Διαχωρισμός των τμημάτων του σκελετού σε:

- Τμήματα στο πάνω μέρος
- Τμήματα στο κάτω μέρος
- Τμήματα χαρακτήρα
- Τμήματα οπών

Στο σχήμα 6.25 δίδεται ένα παράδειγμα κατάταξης των τμημάτων του εσωτερικού και εξωτερικού σκελετού σε μία από τις παραπάνω κατηγορίες.

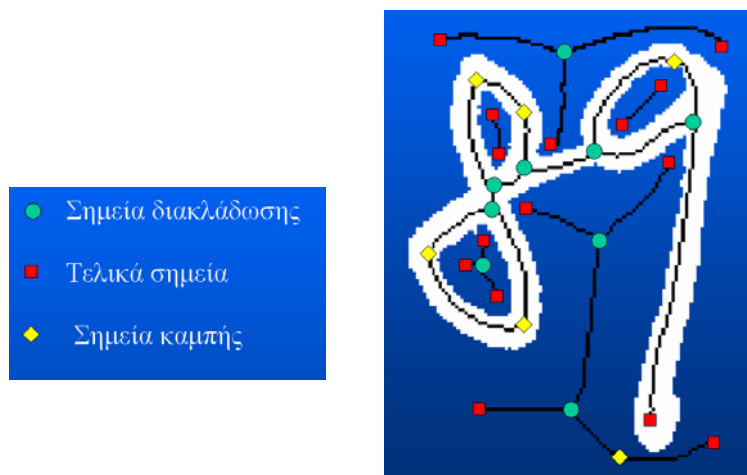


**Σχήμα 6.25.** Κατάταξη των εσωτερικών και εξωτερικών τμημάτων του σκελετού των ενωμένων χαρακτήρων.

**ΒΗΜΑ 3:** Εντοπισμός σημείων σκελετού. Τα σημεία αυτά μπορεί να είναι:

- Σημεία διακλάδωσης
- Τελικά σημεία
- Σημεία καμπής

Στο σχήμα 6.26 δίνεται ένα παράδειγμα των εξαγόμενων σημείων του σκελετού.



**Σχήμα 6.26.** Εντοπισμός σημείων σκελετού.



**ΒΗΜΑ 4:** Εύρεση διαδρομών κατάτμησης. Μεταξύ των ενωμένων χαρακτήρων μπορεί να έχουμε ένωση σε ένα ή σε πολλά σημεία.

Αν έχουμε ένωση σε ένα σημείο, τότε ισχύουν τα παρακάτω:

- Υπάρχει μόνο μία μετάβαση 0->1 και μία 1->0 κατά την διαδρομή κατάτμησης
- Η διαδρομή δεν περιλαμβάνει τμήματα οπών

Για να εντοπίσουμε κάθε διαδρομή κατάτμησης αν έχουμε ένωση σε ένα σημείο, ακολουθούμε τον παρακάτω αλγόριθμο: Ξεκινάμε από ένα σημείο στο πάνω μέρος, περνάμε από ένα σημείο χαρακτήρα και καταλήγουμε σε ένα σημείο στο κάτω μέρος. Αν δεν υπάρχει κατάλληλο σημείο χαρακτήρα ή στο κάτω μέρος η διαδρομή συνεχίζει μέχρι να συναντήσει τμήμα σκελετού στο κάτω μέρος.

Αντίστοιχα, αν έχουμε ένωση σε πολλά σημεία, η διαδρομή κατάτμησης περιλαμβάνει μία ή περισσότερες περιοχές οπών. Ο αλγόριθμος είναι ο ακόλουθος: Ξεκινάμε από ένα σημείο στο πάνω μέρος, περνάμε από ένα σημείο χαρακτήρα, μετά από ένα σημείο οπής, ξανά σημείο χαρακτήρα και καταλήγουμε σε ένα σημείο στο κάτω μέρος. Παραδείγματα εύρεσης διαδρομών κατάτμησης δίνονται στο σχήμα 6.27.



Σχήμα 6.27. Παραδείγματα διαδρομών κατάτμησης σε ενωμένους χειρόγραφους αριθμούς.

## Βιβλιογραφία

- (Bozinovic 1989)** Bozinovic, R.M., Srihari, S.N.: Off-Line Cursive Script Word Recognition. IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 11, No 1 (1989) 68-83
- (Casey 1982)** Casey, R.G., Nagy, G.: Recursive segmentation and classification of composite character patterns. 6<sup>th</sup> Int. Conf. On Pattern Recognition (1982) 1023-1026
- (Chen 2000)** Chen, Y., Wang, J.: Segmentation of Single – or Multiple – Touching Handwritten Numeral String Using Background and Foureground Analysis. IEEE Trans. Pattern Anal. Mach. Intell. PAMI Vol. 22, No. 11 (2000) 1304-1317
- (Kahan 1987)** Kahan, S., Pavlidis, T.: On the recognition of printed characters of any font and size. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-9 (1987) 274-287
- (Lu 1992)** Lu, Y., Haist, B., Harmon, L., Trenkle, J., Vogt, R.: An accurate and efficient system for segmenting machine-printed text. Proc. of the U.S. Postal Service 5<sup>th</sup> Advanced Technology Conference, Vol. 3 (1992) 93-105
- (Lu 1995)** Lu, Y.: Machine Printed Character Segmentation – An Overview. Pattern Recognition, Vol. 28, No. 1 (1995) 67-80
- (Lu 1996)** Lu, Y., Shridhar, M.: Character Segmentation in Handwritten Words – An Overview. Pattern Recognition, Vol. 29, No. 1 (1996) 77-96
- (Pavlidis 1981)** Pavlidis, T.: Algorithms for Graphics and Image Processing. Springer-Verlag (1981)
- (Schilling 2000)** Schilling, R. J.: Fundamental of Robotics Analysis and Control. Prentice-Hall, Englewood Cliffs, NJ (1990)
- (Tsujiimoto 1991)** Tsujimoto, S., Asada, H.: Resolving ambiguity in segmenting touching characters. First Int. Conf. On Document Analysis and Recognition (1991) 701-709