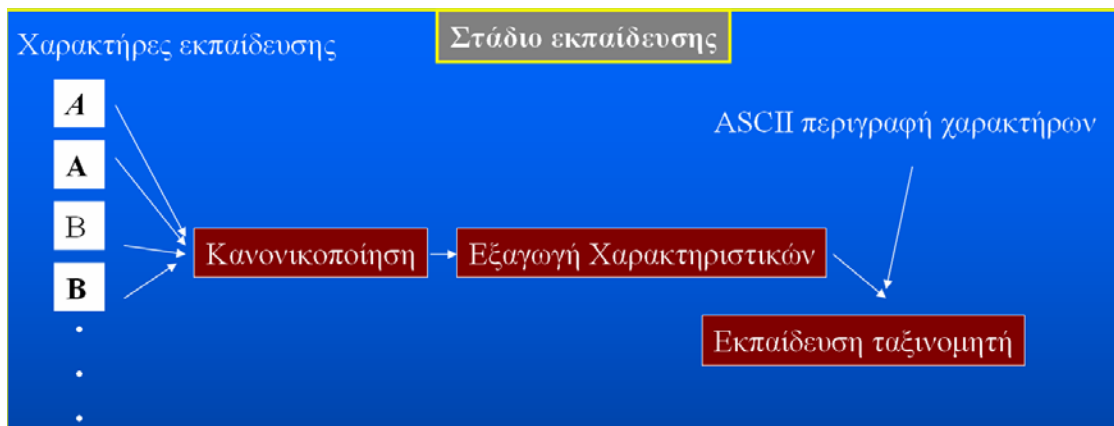


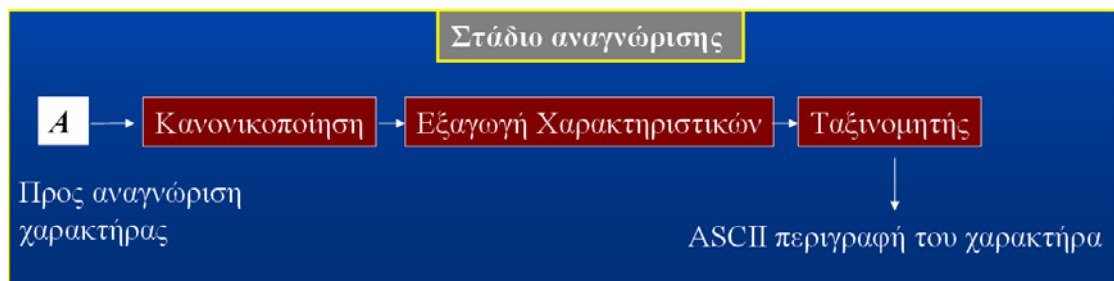
Μάθημα 7

Εξαγωγή Χαρακτηριστικών

Το στάδιο της εξαγωγής χαρακτηριστικών αφορά το πρώτο βήμα για την αναγνώριση των χαρακτήρων και περιλαμβάνει την μετατροπή κάθε χαρακτήρα σε διάνυσμα χαρακτηριστικών μικρής διάστασης το οποίο αποτελεί και την ταυτότητά του. Κάθε χαρακτηριστικό μπορεί να έχει συγκεκριμένη κανονικοποιημένη τιμή, π.χ. περίμετρος του χαρακτήρα, ποσοστό κάλυψης από μαύρα pixels ή να έχει τιμή 0 ή 1 ανάλογα αν πληροί ή όχι κάποια ιδιότητα, π.χ. ύπαρξη οπής, ύπαρξη τελικού σημείου σε κάποιο τεταρτημόριο του χαρακτήρα. Η εξαγωγή χαρακτηριστικών χρησιμοποιείται τόσο στο στάδιο της εκπαίδευσης όσο και στο στάδιο της αναγνώρισης. Στο στάδιο της εκπαίδευσης, ένα σύνολο χαρακτήρων κανονικοποιείται ως προς το μέγεθος, μετατρέπεται σε διανύσματα χαρακτηριστικών και συνδυάζεται με την αντίστοιχη ASCII κωδικοποίησή του για την εκπαίδευση του ταξινομητή αναγνώρισης (σχήμα 7.1α). Στο στάδιο αναγνώρισης, κάθε περιοχή της εικόνας που έχει εντοπιστεί ότι περιέχει χαρακτήρα, κανονικοποιείται, μετατρέπεται σε διάνυσμα χαρακτηριστικών και στη συνέχεια αντιστοιχίζεται σε ένα από τους γνωστούς χαρακτήρες με τη βοήθεια του ταξινομητή αναγνώρισης (σχήμα 7.1β).



(α)



(β)

Σχήμα 7.1. Οπτική αναγνώριση χαρακτήρων. (α) Στάδιο εκπαίδευσης. (β) Στάδιο αναγνώρισης.

Β. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

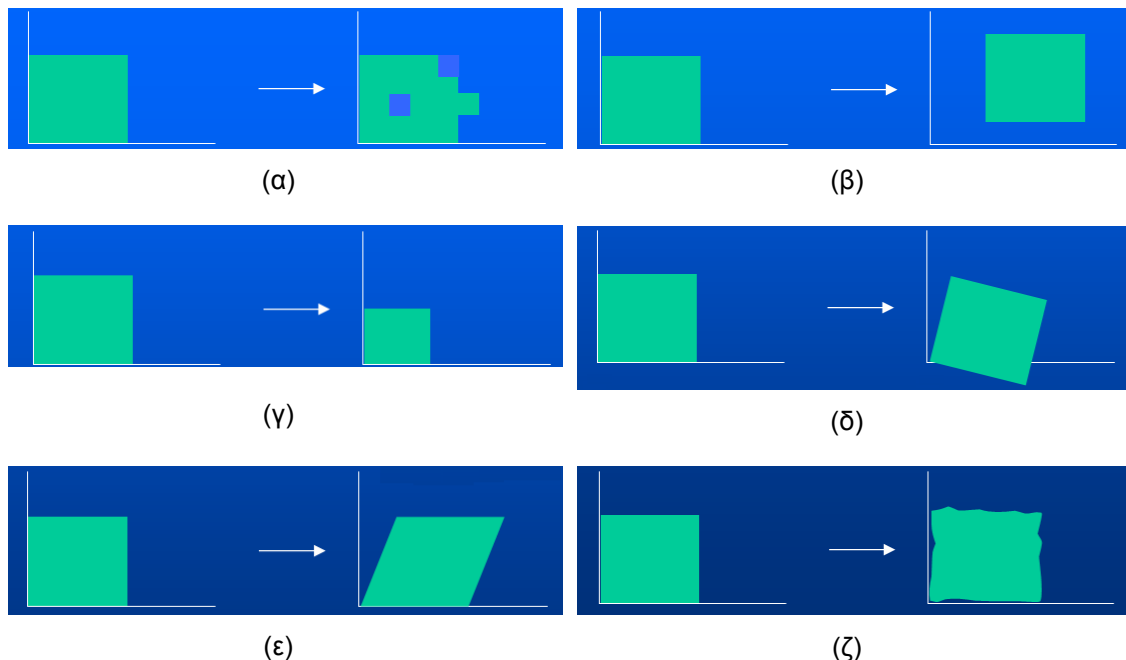
Ένα σύνολο ικανοποιητικών χαρακτηριστικών πρέπει να πληροί ορισμένες προϋποθέσεις όπως:

- Διακριτική ικανότητα. Τα χαρακτηριστικά πρέπει να έχουν αρκετά διαφορετικές τιμές για χαρακτήρες οι οποίοι ανήκουν σε διαφορετικές κλάσεις.
- Αξιοπιστία. Τα χαρακτηριστικά πρέπει να έχουν παραπλήσιες τιμές για χαρακτήρες οι οποίοι ανήκουν στην ίδια κλάση.
- Ανεξαρτησία. Τα χαρακτηριστικά δεν πρέπει να επικαλύπτονται μεταξύ τους.
- Μικρό μέγεθος. Ο αριθμός των χαρακτηριστικών πρέπει να είναι μικρός ώστε η αναγνώριση να γίνεται απλά και γρήγορα.
- Μικρό υπολογιστικό κόστος.
- Μικρή πολυπλοκότητα εξαγωγής.

Επίσης, τα χαρακτηριστικά θα πρέπει να μένουν αναλλοίωτα όταν ο χαρακτήρας εμφανίζει:

- Θόρυβο
- Μετακίνηση
- Αλλαγή στο μέγεθος
- Στροφή
- Κλίση
- Παραμόρφωση

Οι παραπάνω μετασχηματισμοί του χαρακτήρα για τους οποίους θέλουμε τα χαρακτηριστικά να μένουν αναλλοίωτα φαίνονται στο σχήμα 7.2.



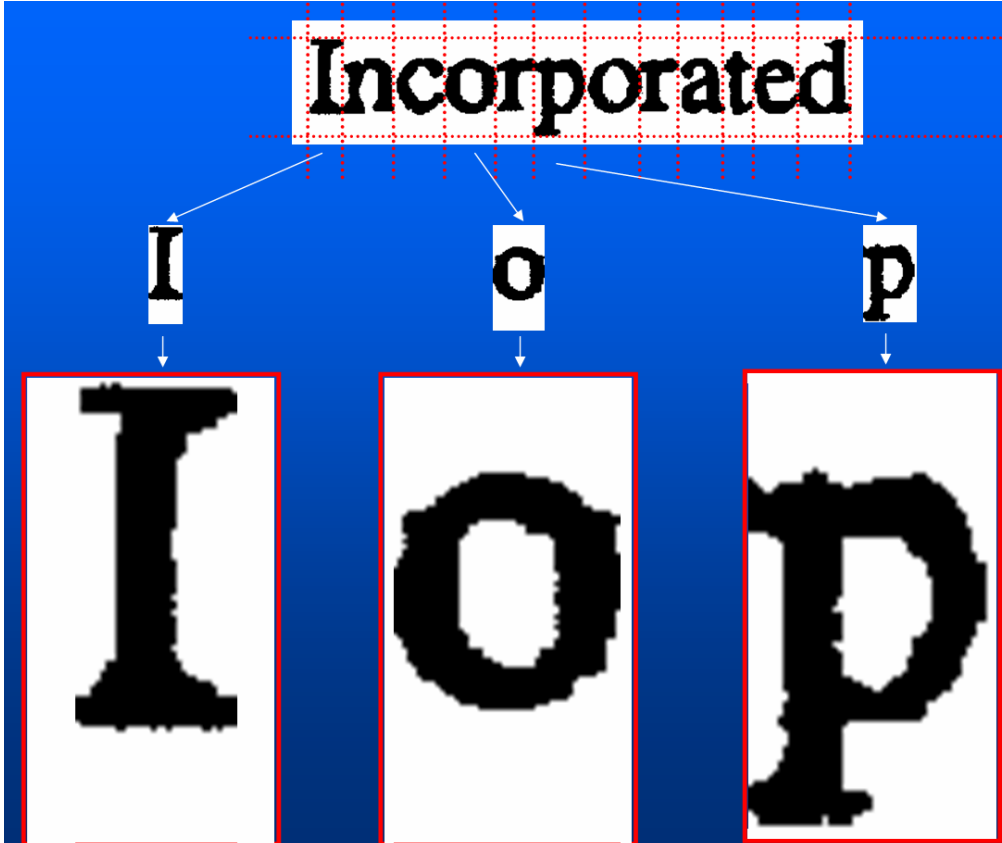
Σχήμα 7.2. Μετασχηματισμοί του χαρακτήρα για τους οποίους θέλουμε τα χαρακτηριστικά να παραμένουν αναλλοίωτα. (α) Θόρυβος. (β) Μετακίνηση. (γ) Αλλαγή στο μέγεθος. (δ) Στροφή. (ε) Κλίση. (ζ) Παραμόρφωση.

Η εξαγωγή χαρακτηριστικών μπορεί να βασίζεται στο σώμα, στο περίγραμμα ή στο σκελετό του χαρακτήρα (Trier 1996). Στη συνέχεια, θα περιγραφεί το στάδιο της κανονικοποίησης των χαρακτήρων και θα αναλυθούν οι κυριότερες μέθοδοι που βασίζονται στις παραπάνω κατηγορίες.

7.1 Κανονικοποίηση χαρακτήρων

Η κανονικοποίηση των χαρακτήρων ως προς το μέγεθος είναι ένα βασικό στάδιο προεπεξεργασίας πριν την εξαγωγή χαρακτηριστικών και την αναγνώριση. Το στάδιο αυτό είναι απαραίτητο γιατί θα πρέπει οι χαρακτήρες που θα συγκριθούν με τους χαρακτήρες της βάσης αναγνώρισης να έχουν το ίδιο μέγεθος. Συνήθως, γίνεται κανονικοποίηση των χαρακτήρων σε πίνακες των οποίων το ύψος είναι σχεδόν διπλάσιο από το πλάτος, π.χ. 24x48. Οι διαστάσεις του παραθύρου κανονικοποίησης εξαρτώνται και από την μέθοδο εξαγωγής χαρακτηριστικών που θα χρησιμοποιήσουμε. Για παράδειγμα, αν θα χρησιμοποιήσουμε τον χωρισμό των χαρακτήρων σε παράθυρα διαστάσεων 5x10 τότε το παράθυρο κανονικοποίησης θα πρέπει να έχει πλάτος πολλαπλάσιο του 5 και ύψος πολλαπλάσιο του 10, π.χ. 30x60.

Για να προχωρήσουμε σε σωστή κανονικοποίηση κάθε χαρακτήρα θα πρέπει να χρησιμοποιήσουμε τις πληροφορίες τόσο για την οριοθέτηση του χαρακτήρα στην γραμμή του κειμένου όσο και για την οριοθέτηση της γραμμής του κειμένου στην συνολική εικόνα (σχήμα 7.3). Το αριστερό και δεξί όριο του χαρακτήρα θεωρείται το όριο διαχωρισμού από το προηγούμενο και το επόμενο γράμμα. Το πάνω και κάτω όριο του χαρακτήρα θα πρέπει να είναι το πάνω και κάτω όριο της αντίστοιχης γραμμής κειμένου. Αυτό γίνεται για να διαχωριστούν τα κεφαλαία από τα μικρά γράμματα (π.χ. «ο» - «Ο») και οι χαρακτήρες που εξαρτώνται από την θέση τους στην γραμμή του κειμένου (π.χ. «.» - «'»). Εναλλακτικά, μπορούμε να κανονικοποιήσουμε κάθε χαρακτήρα στα όρια του, όμως θα πρέπει να αποθηκεύσουμε κάποια επιπλέον πληροφορία την οποία θα την χρησιμοποιήσουμε στην ταξινόμηση και αφορά το μέγεθος και την θέση των χαρακτήρων ως προς την γραμμή του κειμένου.



Σχήμα 7.3. Παράδειγμα κανονικοποίησης χαρακτήρων.

7.2 Χρησιμοποιώντας το σώμα των χαρακτήρων

Οι βασικότεροι μέθοδοι εξαγωγής χαρακτηριστικών που βασίζονται στην πληροφορία που υπάρχει στο συνολικό σώμα των χαρακτήρων είναι οι μέθοδοι απευθείας σύγκρισης (template matching), ο χωρισμός σε ζώνες, οι γεωμετρικές ροπές και οι προβολές. Επιπροσθέτως, από το σώμα των χαρακτήρων μπορούμε να εξάγουμε χαρακτηριστικά που αφορούν το ποσοστό κάλυψης των μαύρων pixels, τον αριθμό των σημείων τομής με ευθείες (crossing points), την θέση του χαρακτήρα ως προς την γραμμή κειμένου καθώς και την ύπαρξη οπών.

Απευθείας σύγκριση (template matching)

Η μεθοδολογία της απευθείας σύγκρισης (template matching) θεωρεί ότι τα pixels της εικόνας του χαρακτήρα είναι και το διάνυσμα των χαρακτηριστικών του (Pratt 1991). Αν θέλουμε να συγκρίνουμε δύο χαρακτήρες X, Y (X είναι το πρότυπο και Y ο χαρακτήρας που εξετάζουμε) οι οποίοι έχουν κανονικοποιηθεί στις ίδιες διαστάσεις και αποτελούνται από n σημεία, τότε ο αριθμός n_{ij} των σημείων όπου ο χαρακτήρας X έχει τιμή i και ο χαρακτήρας Y έχει τιμή j , με $i, j \in \{0, 1\}$ δίδεται από τον τύπο:

$$n_{ij} = \sum_{m=1}^n \delta_m(i, j) \quad (7.1)$$

όπου

$$\delta_m(i, j) = \begin{cases} 1, & \text{if } (x_m = i) \wedge (y_m = j) \\ 0, & \text{otherwise} \end{cases} \quad (7.2)$$

και x_m και y_m είναι τα m -ιστά σημεία των προς σύγκριση χαρακτήρων X και Y .

Χρησιμοποιώντας την απευθείας σύγκριση των pixels έχουν προταθεί οι παρακάτω βασικές τεχνικές για τον εντοπισμό της απόστασης μεταξύ δύο χαρακτήρων (Tubbs 1989):

- XOR μεταξύ των δύο πινάκων:

$$D = n_{10} + n_{01} \quad (7.3)$$

Όσο μεγαλύτερη είναι η απόσταση D τόσο πιο μακριά είναι οι δύο χαρακτήρες.

- Οι αποστάσεις Jaccard και Yule:

$$d_J = \frac{n_{11}}{n_{11} + n_{10} + n_{01}} \quad (7.4)$$

$$d_Y = \frac{n_{11}n_{00} - n_{10}n_{01}}{n_{11}n_{00} + n_{10}n_{01}} \quad (7.5)$$

Οι παραπάνω αποστάσεις όσο πιο κοντά στο 1 είναι τόσο κοννότερα είναι οι δύο χαρακτήρες. Στο παράδειγμα του πίνακα 7.1 υπολογίζονται οι παραπάνω αποστάσεις για την απευθείας σύγκριση ενός προτύπου με τρεις χαρακτήρες.

- Η χρήση βαρών: Ο τύπος υπολογισμού του n_{ij} γίνεται:

$$n_{ij} = \sum_{m=1}^n p_m(k/i) \delta_m(i, j) \quad (7.6)$$

B. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

όπου $p_m(k/i)$ είναι η πιθανότητα η εικόνα Y να ταιριάζει με το πρότυπο X_k , δεδομένου ότι το m -ιοστό σημείο του προτύπου X_k είναι i . Η τιμή του $p_m(k/i)$ προσεγγίζεται σαν ο αριθμός των προτύπων που έχουν την ίδια m -ιοστή τιμή σημείου όπως το πρότυπο X_k δια τον συνολικό αριθμό των προτύπων. Η χρήση βαρών βοηθάει στο να δίνουμε μεγαλύτερη βαρύτητα στα pixels τα οποία ανήκουν με μεγαλύτερη σιγουριά σε συγκεκριμένη κλάση χαρακτήρα. Στο σχήμα 7.4 δίδεται παράδειγμα υπολογισμού των πιθανοτήτων $p_m(k/i)$ για κλάση χαρακτήρων η οποία αποτελείται από 3 πρότυπα. Στο συγκεκριμένο παράδειγμα, το n_{11} της σύγκρισης με τον πρώτο χαρακτήρα από 11 γίνεται $(9 \cdot 1 + 2 \cdot 2/3) = 10,33$

Πίνακας 7.1. Εντοπισμός της απόστασης χρησιμοποιώντας απευθείας σύγκριση (template matching). D η απόσταση βάσει του XOR, d_J η Jaccard απόσταση και d_Y η Yule απόσταση.

| | | | |
|---|-------|-------|-------|
| | | | |
| n_{11} | 11 | 10 | 10 |
| n_{10} | 1 | 2 | 2 |
| n_{01} | 1 | 1 | 1 |
| n_{00} | 2 | 2 | 2 |
| $D = n_{10} + n_{01}$ | 2 | 3 | 3 |
| $d_J = \frac{n_{11}}{n_{11} + n_{10} + n_{01}}$ | 0,846 | 0,769 | 0,769 |
| $d_Y = \frac{n_{11}n_{00} - n_{10}n_{01}}{n_{11}n_{00} + n_{10}n_{01}}$ | 0,913 | 0,818 | 0,818 |

$p_1(k/0) = 0/3 = 0$
 $p_1(k/1) = 3/3 = 1$

$p_3(k/0) = 1/3$
 $p_3(k/1) = 2/3$

$p_5(k/0) = 3/3 = 1$
 $p_5(k/1) = 0/3 = 0$

(α)

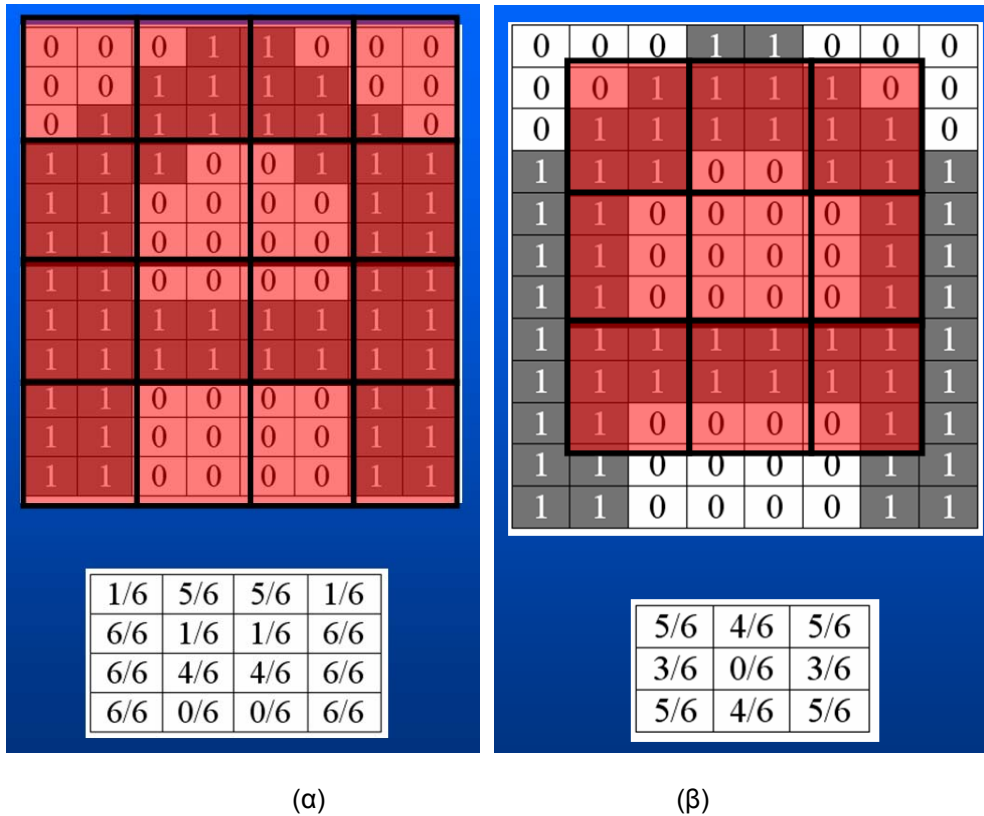
| $p_m(k/0)$ | | | $p_m(k/1)$ | | |
|------------|---|-----|------------|---|-----|
| 0 | 0 | 1/3 | 1 | 1 | 2/3 |
| 0 | 1 | 0 | 1 | 0 | 1 |
| 1/3 | 0 | 0 | 2/3 | 1 | 1 |
| 0 | 1 | 0 | 1 | 0 | 1 |
| 0 | 1 | 0 | 1 | 0 | 1 |

(β)

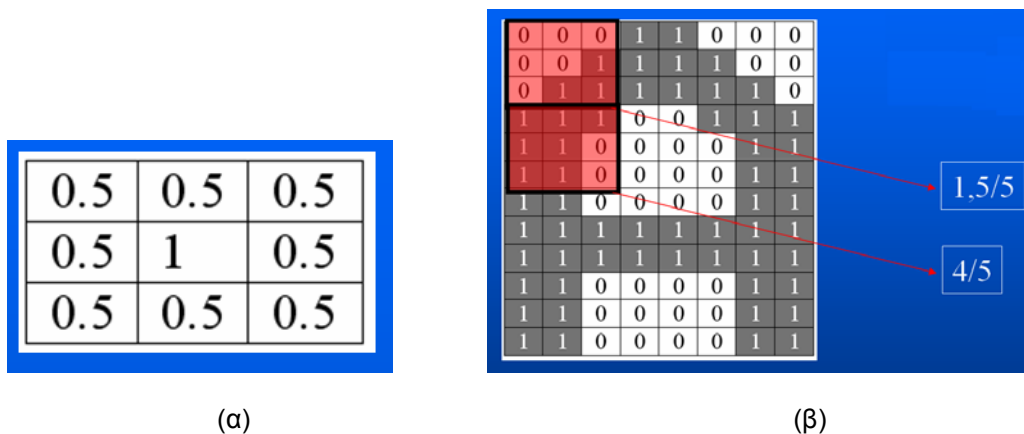
Σχήμα 7.4 Παράδειγμα υπολογισμού των πιθανοτήτων $p_m(k/i)$ για κλάση χαρακτήρων η οποία αποτελείται από 3 πρότυπα. (α) Η πιθανότητα $p_1(k/0)$ είναι $0/3$ γιατί δεν έχουμε κανένα πρότυπο που να έχει 0 στην θέση 1 (πάνω αριστερό pixel). Ομοίως υπολογίζονται οι όλα τα $p_m(k/i)$. (β) Οι τελικές τιμές για τα $p_m(k/0)$ και $p_m(k/1)$ της συγκεκριμένης κλάσης προτύπων.

Χωρισμός σε ζώνες

Σύμφωνα με την προσέγγιση χωρισμού σε ζώνες, υπολογίζεται η κατανομή των μαύρων pixels σε διάφορες ζώνες στις οποίες χωρίζουμε τον χαρακτήρα (Bokser 1992). Για την εξαγωγή χαρακτηριστικών με χρήση των ζωνών του χαρακτήρα, η πιο απλή εφαρμογή είναι ο χωρισμός του χαρακτήρα σε $n \times m$ ζώνες και ο υπολογισμός της πυκνότητας των pixels σε κάθε ζώνη. Στο παράδειγμα του σχήματος 7.5 υπολογίζεται η πυκνότητα των pixels σε 16 ζώνες καθώς και σε 9 ζώνες που επικαλύπτουν τις πρώτες στις οποίες έχουμε χωρίσει τον χαρακτήρα. Εναλλακτικά, μπορούμε να δώσουμε βάρη στα pixels κάθε παραθύρου που έχουμε χωρίσει τον χαρακτήρα τονίζοντας περισσότερο τα pixels που βρίσκονται στο κέντρο του παραθύρου (σχήμα 7.6) (Gatos 1993). Με αυτό τον τρόπο επιτυγχάνουμε μεγαλύτερη αντοχή στον θόρυβο.



Σχήμα 7.5 Υπολογισμός της πυκνότητας των pixels σε ζώνες στις οποίες έχουμε χωρίσει τον χαρακτήρα. (α) 16 ζώνες. (β) 9 ζώνες που επικαλύπτουν τις πρώτες.



Σχήμα 7.6 Χρήση βαρών για τον υπολογισμό της πυκνότητας των pixels σε κάθε παράθυρο. (α) Τα βάρη σε κάθε παράθυρο. (β) Οι τιμές των χαρακτηριστικών που προκύπτουν.

Προβολές

Οι προβολές σε οριζόντια και κάθετη διεύθυνση συνήθως χρησιμοποιούνται για την κατάτμηση των εγγράφων σε γραμμές κειμένου, λέξεις, χαρακτήρες κ.λ.π. καθώς και για τον εντοπισμό της στροφής των εγγράφων. Όμως έχουν προταθεί για την εξαγωγή χαρακτηριστικών από το 1956 (Glauberman 1956). Ένα παράδειγμα οριζόντιας και κάθετης προβολής ενός χαρακτήρα δίδεται στο σχήμα 7.7. Το ιστόγραμμα της προβολής συνήθως κανονικοποιείται με την ομαδοποίηση γειτονικών γραμμών ή στηλών ώστε να χρησιμοποιείται σταθερός αριθμός περιοχών σε οριζόντια και κάθετη διεύθυνση. Όμως, οι προβολές είναι πολύ ευαίσθητες στην στροφή και στον τύπο γραψίματος. Επιπροσθέτως, βασική πληροφορία σχετικά με την μορφή του χαρακτήρα χάνεται.

Η κάθετη προβολή $y(x_i)$ αντιστοιχεί στον αριθμό των pixels για το οποίο ισχύει $x=x_i$. Για να μετρήσουμε την ομοιότητα μεταξύ δύο ιστογραμμάτων αρκεί να υπολογίσουμε το:

$$d = \sum_{i=1}^n |y_1(x_i) - y_2(x_i)| \quad (7.7)$$

όπου n ο αριθμός των υποδιαίρεσεων της οριζόντιας διάστασης του χαρακτήρα και y_1 και y_2 είναι τα δύο ιστογράμματα που θέλουμε να συγκρίνουμε. Συνήθως, χρησιμοποιούμε το αθροιστικό ιστόγραμμα Y το οποίο αθροίζει τις k πρώτες τιμές του ιστογράμματος:

$$Y(x_k) = \sum_{i=1}^k y(x_i) \quad (7.8)$$

Η απόσταση μεταξύ των δύο ιστογραμμάτων γίνεται:

$$D = \sum_{i=1}^n |Y_1(x_i) - Y_2(x_i)| \quad (7.9)$$

Η απόσταση D δεν είναι τόσο ευαίσθητη σε τυχών μετατοπίσεις των επικρατέστερων κορυφών των οριζόντιων και κάθετων προβολών.



Σχήμα 7.7 Εξαγωγή χαρακτηριστικών που βασίζεται στις οριζόντιες και κάθετες προβολές του χαρακτήρα.

Γεωμετρικές ροπές (geometric moments)

Οι αναλλοίωτες ροπές έχουν προταθεί σαν χαρακτηριστικά για την αναγνώριση προτύπων από την δεκαετία του 1960 (Hu 1962). Με την χρήση των ροπών μπορούμε να εξαγάγουμε χαρακτηριστικά των αντικειμένων τα οποία μπορεί να είναι αναλλοίωτα ως προς την μετακίνηση, το μέγεθος και την στροφή. Οι κανονικές ροπές (regular moments) τάξης $(p+q)$ ορίζονται για μια ασπρόμαυρη εικόνα ως εξής:

$$m_{pq} = \sum_{i=1}^M (x_i)^p (y_i)^q \quad (7.10)$$

όπου M είναι το σύνολο των σημείων της εικόνας και (x_i, y_i) οι συντεταγμένες κάθε σημείου.

Οι κεντρικές ροπές (central moments) οι οποίες είναι ανεξάρτητες μετακίνησης δίδονται από το τύπο:

$$\mu_{pq} = \sum_{i=1}^M (x_i - \bar{x})^p (y_i - \bar{y})^q \quad (7.11)$$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}} \quad (7.12)$$

Οι κανονικοποιημένες κεντρικές ροπές (normalized central moments) οι οποίες είναι ανεξάρτητες της μετακίνησης και του μεγέθους δίδονται από τον τύπο:

$$n_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \quad \gamma = \frac{p+q}{2} + 1, \quad p+q = 2, 3, \dots \quad (7.13)$$

Τέλος, οι αναλλοίωτες ροπές H_u (Hu moments invariants) οι οποίες είναι ανεξάρτητες μετακίνησης, μεγέθους και στροφής, δίδονται από τους τύπους (οι ροπές δεύτερης τάξης):

$$\Phi_1 = n_{20} + n_{02} \quad (7.14)$$

$$\Phi_2 = (n_{20} - n_{02})^2 + 4n_{11}^2$$

Στο σχήμα 7.8 δίδεται παράδειγμα υπολογισμού των κανονικών ροπών, των κεντρικών ροπών, των κανονικοποιημένων κεντρικών ροπών και των αναλλοίωτων ροπών H_u , για χαρακτήρα οποίος έχει υποστεί μετακίνηση, μεγέθυνση και στροφή. Παρατηρούμε ότι οι κεντρικές ροπές έχουν ίδιες τιμές για τους δύο πρώτους χαρακτήρες (αναλλοίωτες σε μετακίνηση), οι κανονικοποιημένες κεντρικές ροπές έχουν ίδιες τιμές για τους τρεις πρώτους χαρακτήρες (αναλλοίωτες σε μετακίνηση και αλλαγή στο μέγεθος) ενώ οι αναλλοίωτες ροπές H_u έχουν περίπου ίδιες τιμές για όλους τους χαρακτήρες (αναλλοίωτες σε μετακίνηση, αλλαγή στο μέγεθος και περιστροφή).

Ποσοστό κάλυψης των μαύρων pixels

Το ποσοστό κάλυψης των μαύρων pixels για έναν χαρακτήρα δίδεται από τον τύπο:

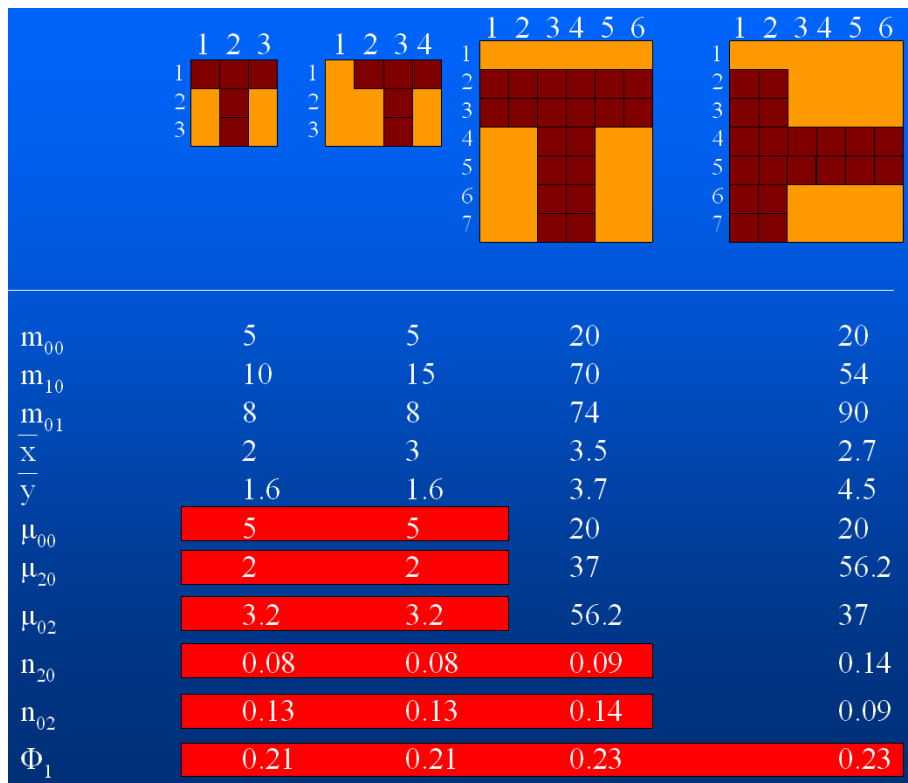
$$P = \frac{\sum_{x=1}^X \sum_{y=1}^Y f(x, y)}{XY} \quad (7.15)$$

όπου $X \times Y$ είναι οι διαστάσεις του κανονικοποιημένου παραθύρου του χαρακτήρα και $f(x, y)$ η εικόνα του χαρακτήρα με τιμές 1 για τα μαύρα pixels και 0 για τα pixels φόντου. Το ποσοστό

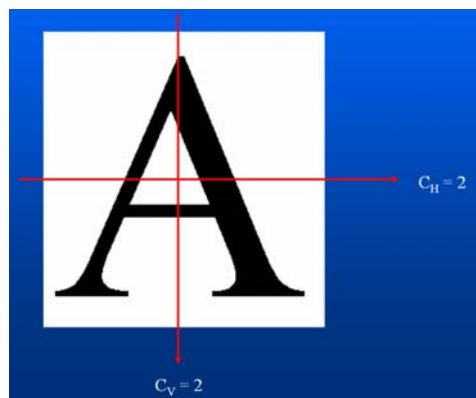
κάλυψης των μαύρων pixels είναι στατιστικό χαρακτηριστικό το οποίο είναι αναλλοίωτο κυρίως ως προς την μετακίνηση και την στροφή. Μπορεί να χρησιμοποιηθεί σαν ένα κριτήριο διαχωρισμού των χαρακτήρων σε βασικές κατηγορίες.

Σημεία τομής με ευθείες (crossing points)

Αν μετρήσουμε τις εναλλαγές 01 και 10 στις οριζόντιες και κάθετες μεσοκάθετους στο κανονικοποιημένο παράθυρο του χαρακτήρα τότε τα σημεία τομής με τις δύο μεσοκάθετους προκύπτουν αν διαιρέσουμε τις εναλλαγές δια 2 (σχήμα 7.9). Ο αριθμός των σημείων τομής με τις μεσοκάθετους χρησιμοποιείται σαν χαρακτηριστικό (Impedono 1991) το οποίο είναι αναλλοίωτο κυρίως ως προς τον θόρυβο, την μετατόπιση, το μέγεθος και την παραμόρφωση. Χρησιμοποιείται κυρίως σαν ένα κριτήριο διαχωρισμού των χαρακτήρων σε βασικές κατηγορίες.



Σχήμα 7.8 Υπολογισμός των κανονικών ροπών, των κεντρικών ροπών, των κανονικοποιημένων κεντρικών ροπών και των αναλλοίωτων ροπών H_u , για χαρακτήρα οποίος έχει υποστεί μετακίνηση, μεγέθυνση και στροφή.



Σχήμα 7.9 Υπολογισμός των σημείων τομής με τις οριζόντιες και κάθετες μεσοκάθετους.

Υπαρξη οπών

Οι οπές που υπάρχουν στην εικόνα του χαρακτήρα μπορούν να υπολογιστούν με τον παρακάτω αλγόριθμο ο οποίος βασίζεται στα διαδοχικά σημεία φόντου (white runs) της ασπρόμαυρης εικόνας.

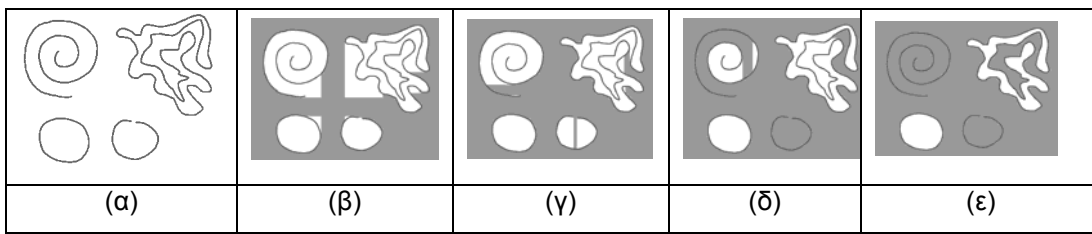
Βήμα 1. Όλα τα οριζόντια και κάθετα διαδοχικά σημεία φόντου που συνορεύουν με τα όρια της εικόνας μαρκάρονται.

Βήμα 2. Όλα τα οριζόντια και κάθετα διαδοχικά μη μαρκαρισμένα σημεία φόντου που συνορεύουν με μαρκαρισμένα σημεία του βήματος 1 μαρκάρονται επίσης.

Βήμα 3. Επαναλαμβάνουμε το βήμα 2 μέχρι να μην υπάρχουν άλλα σημεία για μαρκάρισμα στην τελευταία επανάληψη.

Βήμα 4. Όλα τα μη μαρκαρισμένα σημεία φόντου ανήκουν σε οπές. Χρησιμοποιώντας ονοματισμό των συνδεδεμένων συστατικών για τα σημεία αυτά μπορούμε να εντοπίσουμε τον αριθμό και την θέση των οπών.

Παράδειγμα του παραπάνω αλγορίθμου δίδεται στο σχήμα 7.10. Ο αριθμός των οπών χρησιμοποιείται σαν χαρακτηριστικό το οποίο είναι αναλλοίωτο ως προς τους περισσότερους δυνατούς μετασχηματισμούς του χαρακτήρα. Χρησιμοποιείται κυρίως σαν ένα κριτήριο διαχωρισμού των χαρακτήρων σε βασικές κατηγορίες.



Σχήμα 7.10 Εντοπισμός οπών σε εικόνα. (α) Αρχική εικόνα. (β)-(ε) Αποτελέσματα μετά από 1,2,5 και 19 επαναλήψεις αντίστοιχα.

Θέση ως προς στην γραμμή κειμένου

Αφού εντοπιστούν οι τρεις ζώνες (πάνω, μεσαία, κάτω) από τις οποίες αποτελείται η γραμμή του κειμένου (σχήμα 7.11), υπολογίζουμε αν ο χαρακτήρας έχει τμήμα του σε κάθε μία από αυτές τις ζώνες (Gatos 1997). Τα όρια των τριών ζωνών των γραμμών κειμένου αντιστοιχούν στα τοπικά ελάχιστα της οριζόντιας προβολής της γραμμής κειμένου. Η θέση του χαρακτήρα ως προς την γραμμή κειμένου χρησιμοποιείται σαν χαρακτηριστικό το οποίο είναι αναλλοίωτο ως προς τους περισσότερους δυνατούς μετασχηματισμούς του χαρακτήρα. Χρησιμοποιείται κυρίως σαν ένα κριτήριο διαχωρισμού των χαρακτήρων σε βασικές κατηγορίες.



Σχήμα 7.11 Οι τρεις ζώνες από τις οποίες αποτελείται η γραμμή κειμένου.

7.2 Χρησιμοποιώντας το περίγραμμα των χαρακτήρων

Οι βασικότεροι μέθοδοι εξαγωγής χαρακτηριστικών που βασίζονται στην επεξεργασία του περιγράμματος είναι οι προβολές του περιγράμματος, ο χωρισμός του περιγράμματος σε ζώνες και οι Fourier περιγραφείς (Fourier Descriptors). Επιπροσθέτως, σαν χαρακτηριστικό μπορούμε να θεωρήσουμε και το μήκος του περιγράμματος κάθε χαρακτήρα.

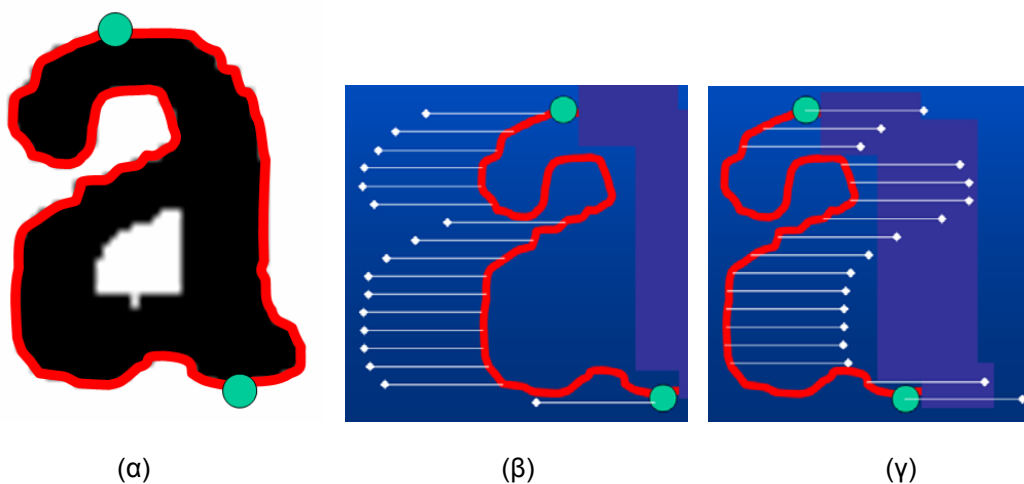
Προβολές του περιγράμματος

Το κίνητρο για την χρήση των προβολών του περιγράμματος είναι ότι κάθε μισό του περιγράμματος (σχήμα 7.12) μπορεί να προσεγγιστεί με ξεχωριστή συνάρτηση ως προς x ή y . Μπορούμε να χρησιμοποιήσουμε οριζόντιες και κάθετες, εσωτερικές ή εξωτερικές προβολές του περιγράμματος των χαρακτήρων.

Για να υπολογίσουμε τις οριζόντιες προβολές του περιγράμματος, πρώτα υπολογίζουμε το ανώτατο και κατώτατο σημείο του περιγράμματος του γράμματος (σχήμα 7.13α). Το περίγραμμα χωρίζεται στα δύο σύμφωνα με τα δύο αυτά σημεία. Για να εξάγουμε τις εξωτερικές προβολές του περιγράμματος, για κάθε τιμή του y , επιλέγουμε τα σημεία του περιγράμματος με τις περισσότερες προς τα έξω τιμές για κάθε μισό του περιγράμματος (σχήμα 7.13β). Για να εξάγουμε τις εσωτερικές προβολές του περιγράμματος, για κάθε τιμή του y , επιλέγουμε τα σημεία του περιγράμματος με τις περισσότερες προς τα μέσα μεγαλύτερες τιμές για κάθε μισό του περιγράμματος (σχήμα 7.13γ). Τα χαρακτηριστικά που προκύπτουν έχουν αρκετή ανεξαρτησία ως προς τον θόρυβο και την παραμόρφωση, όμως εξαρτώνται από την στροφή του χαρακτήρα.



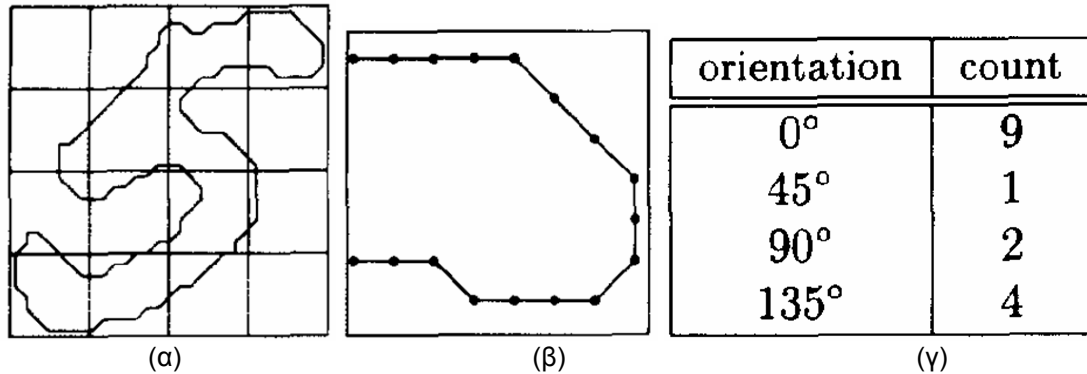
Σχήμα 7.12 Η προβολή προς τα αριστερά και προς τα δεξιά του περιγράμματος του χαρακτήρα.



Σχήμα 7.13 Εξωτερικές και εσωτερικές προβολές του περιγράμματος. (α) Χωρισμός του περιγράμματος σε δύο τμήματα εντοπίζοντας το ανώτατο και κατώτατο σημείο του. (β) Οι εξωτερικές προβολές του αριστερού μέρους των οριζόντιων προβολών του περιγράμματος. (γ) Οι αντίστοιχες εσωτερικές προβολές.

Χωρισμός του περιγράμματος σε ζώνες

Ο χωρισμός του περιγράμματος σε ζώνες έχει χρησιμοποιηθεί με επιτυχία για την εξαγωγή χαρακτηριστικών (Kimura 1991). Σε κάθε ζώνη, τα ευθύγραμμα τμήματα ανάμεσα σε διαδοχικά σημεία του περιγράμματος ομαδοποιούνται ανάλογα με την κλίση τους: Οριζόντια κλίση (0°), κάθετη κλίση (90°) και δύο διαγώνιες κλίσεις (45° , 135°). Τα χαρακτηριστικά προκύπτουν από τον συνολικό αριθμό των τμημάτων κάθε ομάδας. Αν και τα χαρακτηριστικά που προκύπτουν έχουν αρκετή ανεξαρτησία ως προς τον θόρυβο και την παραμόρφωση, όμως εξαρτώνται από την στροφή του χαρακτήρα.



Σχήμα 7.14 Χωρισμός του περιγράμματος σε ζώνες. (α) Χωρισμός σε 16 ζώνες. (β) Τα διαδοχικά σημεία του περιγράμματος για την πάνω δεξιά ζώνη. (γ) Τα χαρακτηριστικά που προκύπτουν ανάλογα με τις κλίσεις των σημείων.

Περιγραφείς Fourier (Fourier Descriptors)

Αν προσεγγίσουμε το περίγραμμα των χαρακτήρων με κάποια συνάρτηση, τότε οι συντελεστές της συνάρτησης αυτής μπορούν να χρησιμοποιηθούν ως χαρακτηριστικά. Η ακολουθία των σημείων του κλειστού περιγράμματος των χαρακτήρων μπορεί να προσεγγιστεί με χρήση κάποιας περιοδικής συνάρτησης. Οι σειρές Fourier έχουν χρησιμοποιηθεί με επιτυχία για τον σκοπό αυτό (Kuhl 1982).

Το κλειστό περίγραμμα, $(x(t), y(t))$, $t=1, \dots, m$, προσεγγίζεται ως εξής:

$$\hat{x}(t) = A_0 + \sum_{n=1}^N \left[a_n \cos \frac{2n\pi t}{T} + b_n \sin \frac{2n\pi t}{T} \right] \quad (7.16)$$

$$\hat{y}(t) = C_0 + \sum_{n=1}^N \left[c_n \cos \frac{2n\pi t}{T} + d_n \sin \frac{2n\pi t}{T} \right]$$

όπου T είναι το συνολικό μήκος του περιγράμματος και $\hat{x}(t) \equiv x(t)$, $\hat{y}(t) \equiv y(t)$ όταν $N \rightarrow \infty$. Οι συντελεστές είναι:

$$A_0 = \frac{1}{T} \int_0^T x(t) dt$$

$$C_0 = \frac{1}{T} \int_0^T y(t) dt$$

$$\begin{aligned}
 a_n &= \frac{2}{T} \int_0^T x(t) \cos \frac{2n\pi t}{T} dt \\
 b_n &= \frac{2}{T} \int_0^T x(t) \sin \frac{2n\pi t}{T} dt \\
 c_n &= \frac{2}{T} \int_0^T y(t) \cos \frac{2n\pi t}{T} dt \\
 d_n &= \frac{2}{T} \int_0^T y(t) \sin \frac{2n\pi t}{T} dt.
 \end{aligned} \tag{7.17}$$

Οι συναρτήσεις $x(t)$ και $y(t)$ μπορούν να οριστούν χρησιμοποιώντας αθροίσματα αντί για ολοκληρώματα στον υπολογισμό των συντελεστών. Οι συντελεστές a_n , b_n , c_n και d_n οι οποίοι είναι και τα εξαγόμενα χαρακτηριστικά, μπορούν να οριστούν ως εξής:

$$\begin{aligned}
 a_n &= \frac{T}{2n^2\pi^2} \sum_{i=1}^m \frac{\Delta x_i}{\Delta t_i} [\cos \phi_i - \cos \phi_{i-1}] \\
 b_n &= \frac{T}{2n^2\pi^2} \sum_{i=1}^m \frac{\Delta x_i}{\Delta t_i} [\sin \phi_i - \sin \phi_{i-1}] \\
 c_n &= \frac{T}{2n^2\pi^2} \sum_{i=1}^m \frac{\Delta y_i}{\Delta t_i} [\cos \phi_i - \cos \phi_{i-1}] \\
 d_n &= \frac{T}{2n^2\pi^2} \sum_{i=1}^m \frac{\Delta y_i}{\Delta t_i} [\sin \phi_i - \sin \phi_{i-1}].
 \end{aligned} \tag{7.18}$$

όπου

$$\begin{aligned}
 \phi_i &= 2n\pi t_i/T, \\
 \Delta x_i &= x_i - x_{i-1}, \quad \Delta y_i = y_i - y_{i-1}, \\
 \Delta t_i &= \sqrt{\Delta x_i^2 + \Delta y_i^2}, \quad t_i = \sum_{j=1}^i \Delta t_j, \\
 T = t_m &= \sum_{j=1}^m \Delta t_j,
 \end{aligned} \tag{7.19}$$

και m είναι ο αριθμός των σημείων του περιγράμματος. Επειδή η επιλογή του αρχικού σημείου (x_i, y_i) επηρεάζει τις τιμές των χαρακτηριστικών, υπολογίζουμε την γωνία μετατόπισης από τον πρώτο κύριο άξονα:

$$\theta_1 = \frac{1}{2} \tan^{-1} \frac{2(a_1 b_1 + c_1 d_1)}{\sqrt{a_1^2 - b_1^2 + c_1^2 - d_1^2}}. \tag{7.20}$$

Οι συντελεστές οι οποίοι είναι ανεξάρτητοι από την επιλογή του αρχικού σημείου υπολογίζονται ως εξής:

$$\begin{bmatrix} a_n^* & b_n^* \\ c_n^* & d_n^* \end{bmatrix} = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} \cos n\theta_1 & -\sin n\theta_1 \\ \sin n\theta_1 & \cos n\theta_1 \end{bmatrix} \quad (7.21)$$

Για να υπολογίσουμε συντελεστές ανεξάρτητους από την γωνία στροφής του αντικειμένου, αρχικά υπολογίζουμε την γωνία του κεντρικού άξονα του αντικειμένου (σχήμα 7.15):

$$\psi_1 = \tan^{-1} \frac{c_1^*}{a_1^*} \quad (7.22)$$

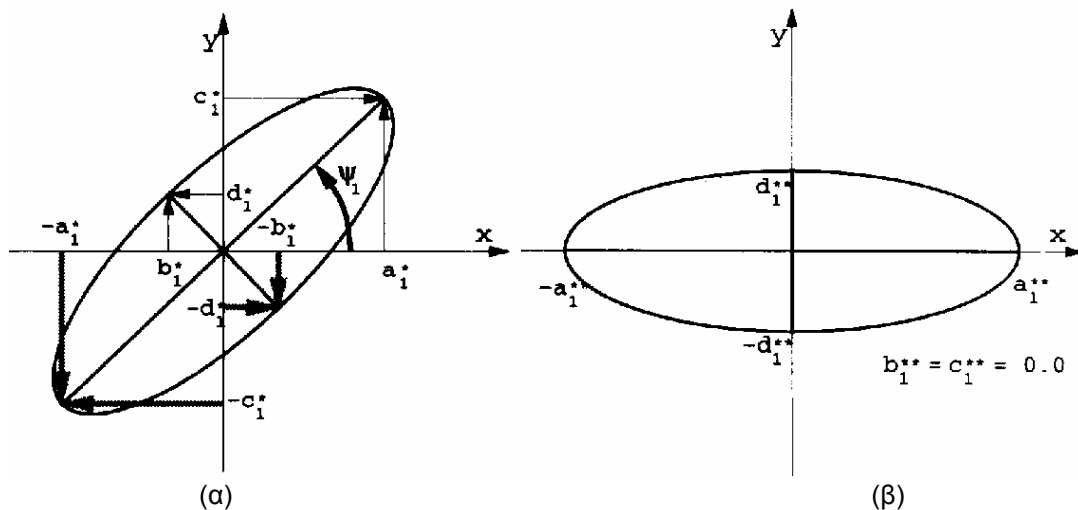
Τα νέα χαρακτηριστικά τα οποία είναι ανεξάρτητα από την στροφή του αντικειμένου δίδονται από τον τύπο:

$$\begin{bmatrix} a_n^{**} & b_n^{**} \\ c_n^{**} & d_n^{**} \end{bmatrix} = \begin{bmatrix} \cos \psi_1 & \sin \psi_1 \\ -\sin \psi_1 & \cos \psi_1 \end{bmatrix} \begin{bmatrix} a_n^* & b_n^* \\ c_n^* & d_n^* \end{bmatrix} \quad (7.23)$$

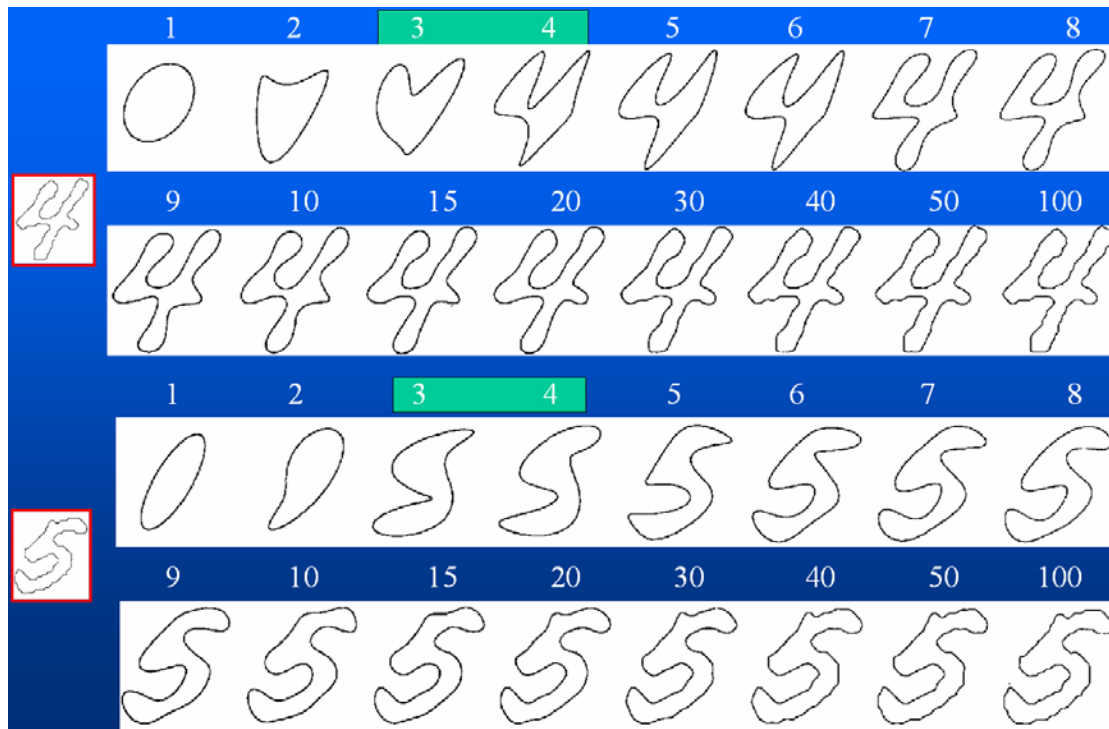
Για να έχουμε επιπλέον ανεξαρτησία από το μέγεθος, αρκεί να διαιρέσουμε του συντελεστές με την ποσότητα E :

$$E = \sqrt{a_1^{*2} + c_1^{*2}} = a_1^{**} \quad (7.24)$$

Στο σχήμα 7.16 δίδονται δύο παραδείγματα αναδημιουργίας των χαρακτήρων χρησιμοποιώντας τους συντελεστές Fourier n τάξης. Χρησιμοποιώντας τους συντελεστές τάξης 3 ή 4 σαν χαρακτηριστικά μπορούμε να δημιουργήσουμε ταξινομητή με ικανοποιητική διαχωριστική ικανότητα.



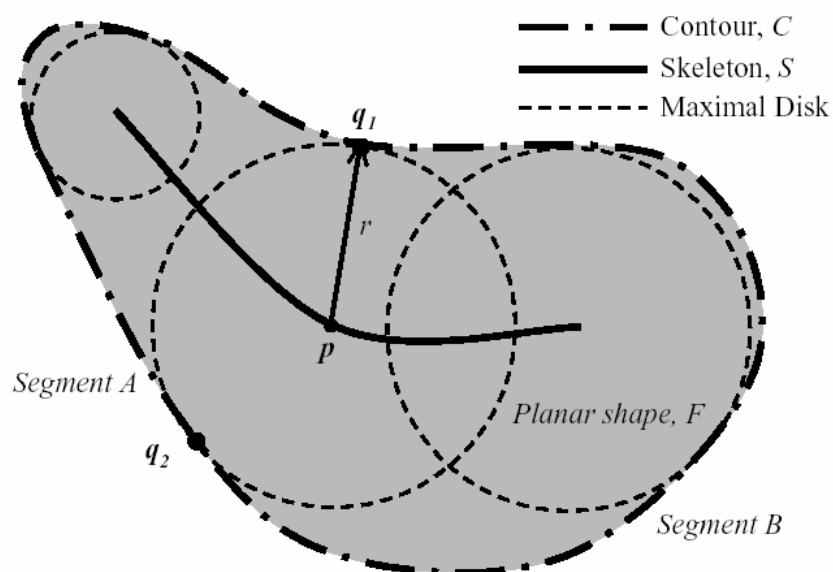
Σχήμα 7.15 Η στροφή των ελλείψεων πρώτης τάξης που χρησιμοποιούνται στους Fourier περιγραφείς ώστε να πετύχουμε ανεξαρτησία από στροφή. (α) Πριν την στροφή. (β) Μετά την στροφή.



Σχήμα 7.16 Αναδημιουργία χαρακτήρων χρησιμοποιώντας τους περιγραφείς Fourier τάξης 1,2,3,4,5,6,7,8,9,10,15,20,30,40,50 και 100. Στο αριστερό μέρος είναι οι αρχικοί χαρακτήρες «4» και «5». Χρησιμοποιώντας τους συντελεστές τάξης 3 ή 4 σαν χαρακτηριστικά μπορούμε να περιγράψουμε ικανοποιητικά τους χαρακτήρες.

7.3 Χρήση σκελετού

Ο σκελετός ενός αντικειμένου ορίζεται σαν ο γεωμετρικός τόπος των κέντρων των μέγιστων κυκλικών δίσκων που περιέχονται μέσα στο αντικείμενο (Choi 2003) (σχήμα 7.17). Για τον εντοπισμό του σκελετού αρκεί να εντοπίσουμε τους κυκλικούς δίσκους οι οποίοι έχουν τουλάχιστον δύο σημεία επαφής με το περίγραμμα του αντικειμένου.



Σχήμα 7.17 Ορισμός του σκελετού ενός αντικειμένου.

Β. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

Αν οι γειτονίες γύρω από ένα pixel p_1 ορίζονται από το σχήμα 7.18, ένας απλός αλγόριθμος υπολογισμού του σκελετού είναι ο ακόλουθος (Gonzalez 1997):

ΒΗΜΑ 1: Ένα pixel μαρκάρεται για απομάκρυνση αν:

$$(\alpha) 2 \leq N(p_1) \leq 6$$

$$(\beta) S(p_1) = 1$$

$$(\gamma) p_2 p_4 p_6 = 0$$

$$(\delta) p_4 p_6 p_8 = 0$$

όπου $N(p_1)$ ο αριθμός των μη μηδενικών γειτόνων του p_1 και $S(p_1)$ ο αριθμός των 0-1 μεταπτώσεων των διαδοχικών pixels p_1, p_2, \dots, p_9 .

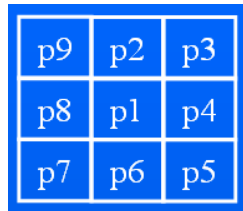
ΒΗΜΑ 2: Μετά την πρώτη από πάνω προς τα κάτω σάρωση σβήνονται τα μαρκαρισμένα pixels και επαναλαμβάνεται η διαδικασία με:

$$(\gamma) p_2 p_4 p_8 = 0$$

$$(\delta) p_2 p_6 p_8 = 0$$

ΒΗΜΑ 3: Η παραπάνω διαδικασία επαναλαμβάνεται μέχρι να μην υπάρχει αλλαγή σε pixel.

Ένα αποτέλεσμα του παραπάνω αλγορίθμου σκελετοποίησης δίδεται στο σχήμα 7.19.



Σχήμα 7.18 Οι γειτονίες γύρω από ένα pixel p_1



Σχήμα 7.19 Σκελετός αντικειμένου χρησιμοποιώντας τον αλγόριθμο του (Gonzalez 1997).

Β. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

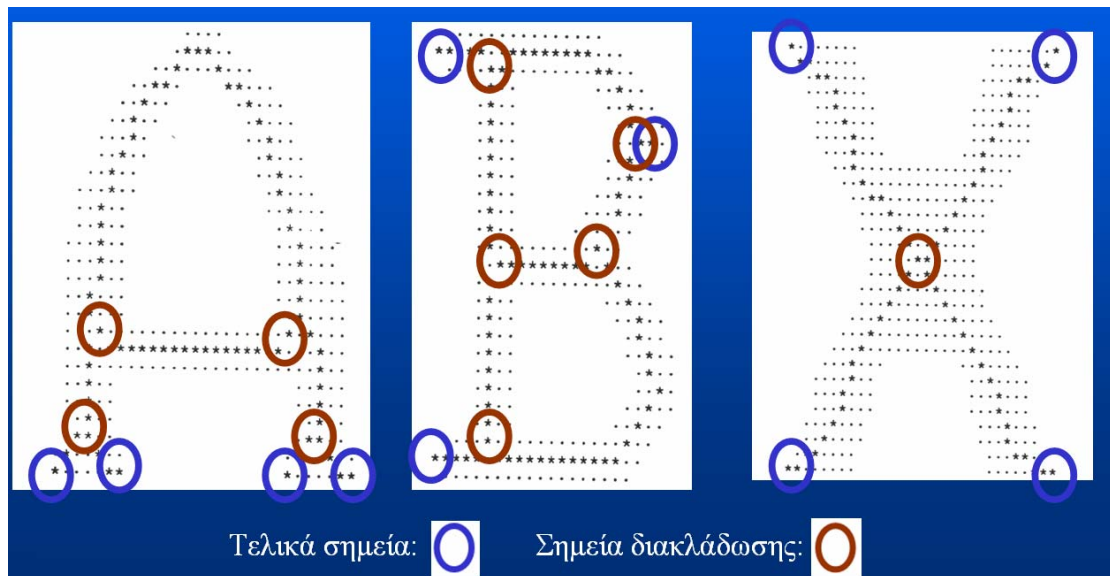
Αν έχουμε εξαίρει τον σκελετό του γράμματος, τότε μπορούμε να χρησιμοποιήσουμε σαν χαρακτηριστικά την θέση των τελικών σημείων και των σημείων διακλάδωσης του σκελετού. Ένα σημείο του σκελετού είναι τελικό αν ισχύει:

$$N(p_1)=1 \quad (7.25)$$

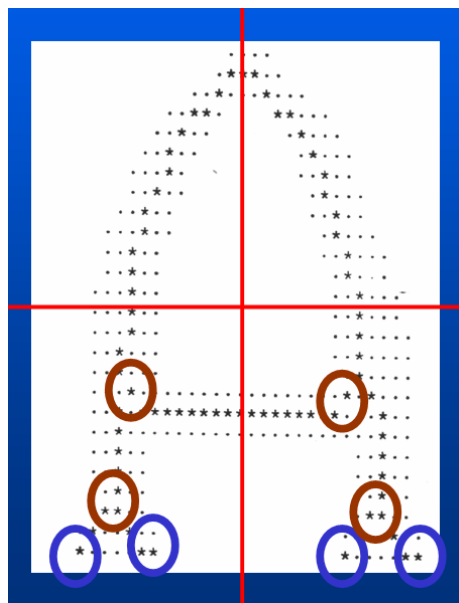
Ένα σημείο του σκελετού είναι σημείο διακλάδωσης αν ισχύει:

$$N(p_1) \geq 3 \quad (7.26)$$

Παραδείγματα τελικών σημείων και σημείων διακλάδωσης δίδονται στο σχήμα 7.20. Συνήθως ως χαρακτηριστικά χρησιμοποιούνται ο αριθμός και η ύπαρξη των σημείων αυτών στα 4 τεταρτημόρια του χαρακτήρα (Gatos 1997).



Σχήμα 7.20 Εντοπισμός τελικών σημείων και σημείων διακλάδωσης στον σκελετό των χαρακτήρων.



Σχήμα 7.20 Εντοπισμός τελικών σημείων και σημείων διακλάδωσης στα 4 τεταρτημόρια του χαρακτήρα.

Βιβλιογραφία

- (Bokser 1992)** Bokser, M.: Omnidocument technologies. Proc. IEEE 80 (1992) 1066-1078
- (Choi 2003)** Choi, W., Lam, K., Siu, W.: Extraction of the Euclidean skeleton based on a connectivity criterion. Pattern Recognition 36 (2003) 721-729
- (Gatos 1993)** Gatos, B., Karras, D., Perantonis, S.: Optical Character Recognition Using Novel Feature Extraction & Neural Network Classification Techniques. Proc. of the Workshop on Neural Network Application and Tools, IEEE Computer Society Press (1993) 65-72
- (Gatos 1997)** Gatos, B., Papamarkos N., Chamzas, C.: A binary tree based OCR technique for machine printed characters", Engineering Applications of Artificial Intelligence, vol. 10, No. 4 (1997) 403-412
- (Glauberman 1956)** Glauberman, M.H.: Character Recognition for Business Machines, Electronics (1956) 132-136
- (Gonzalez 1997)** Gonzalez, R.C.: Digital Image Processing, Second Edition (1997)
- (Hu 1962)** Hu, M.K.: Visual pattern recognition by moment invariants. IRE Trans. Inf. Theory 8 (1962) 179-187
- (Impedovo 1991)** Impedovo, S., Ottaviano, L., Occhinegro, S.: Optical Character Recognition – A survey. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 5 (1991) 1-23
- (Kimura 1991)** Kimura, F., Shridhar, M.: Handwritten numerical recognition based on multiple algorithms. Pattern Recognition, No. 24, Vol. 10 (1991) 969-984
- (Kuhl 1982)** Kuhl, F.P., Giardina, C.R.: Elliptic Fourier features of a closed contour. Comput. Vis. Graphics Image Process. 18 (1982) 236-258
- (Pratt 1991)** Pratt, W.K.: Digital Image Processing, 2nd edn. Wiley, New York (1991)
- (Trier 1996)** Trier, O.D., Jain, A.K., Taxt, T.: Feature Extraction Methods for Character Recognition – A Survey. Pattern Recognition, Vol. 29, No. 4 (1996) 641-662
- (Tubbs 1989)** Tubbs, J.D.: A note on binary template matching. Pattern Recognition, Vol. 22, No. 4 (1989) 359-365