

Μάθημα 3

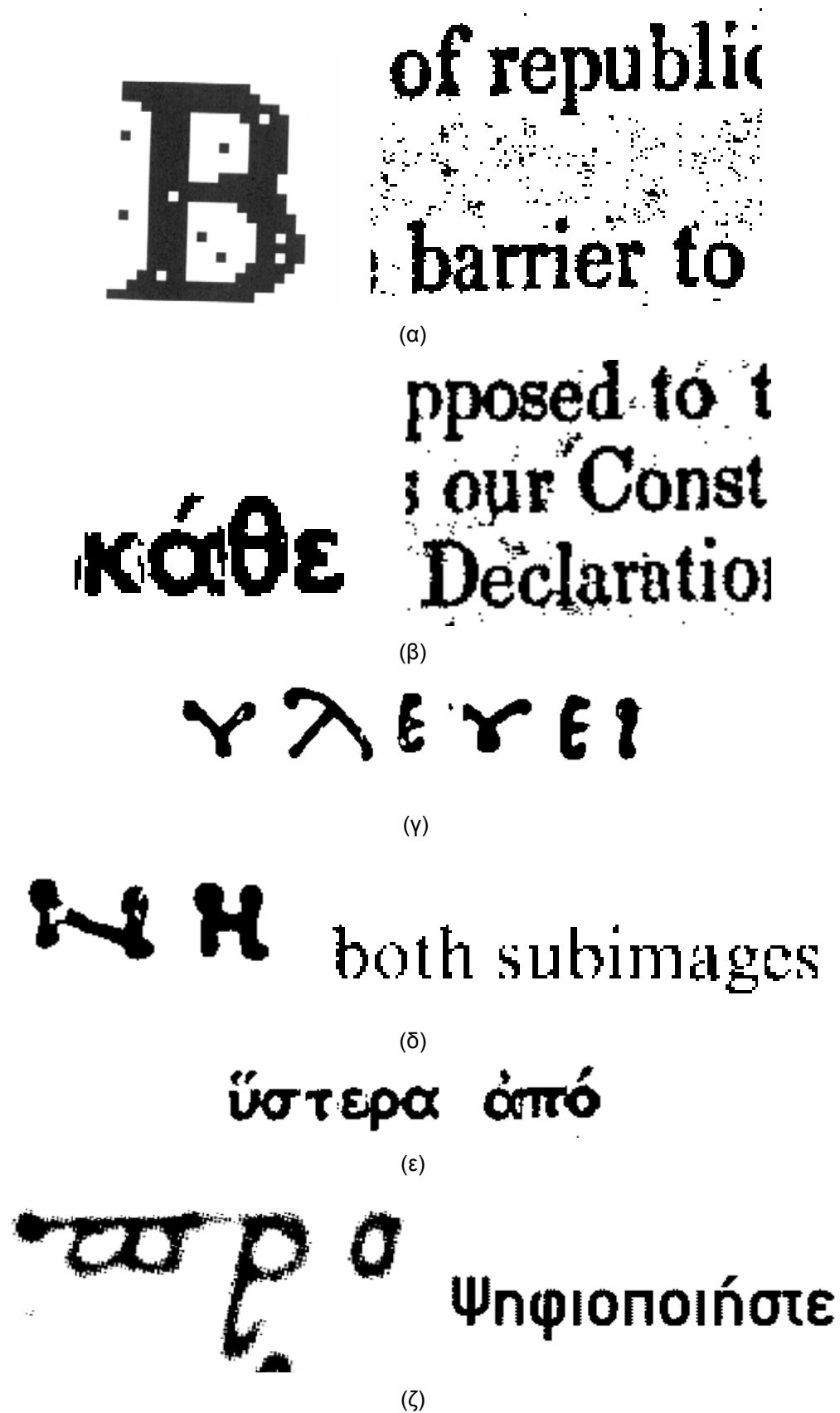
Βελτίωση ποιότητας

3.1 Προβλήματα στην ποιότητα των δυαδικών εικόνων

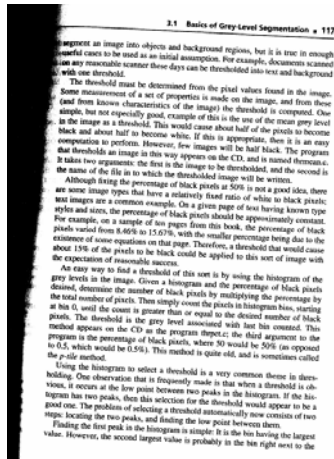
Οι δυαδικές εικόνες των εγγράφων συνήθως χρειάζονται ένα στάδιο προεπεξεργασίας για την βελτίωση της ποιότητάς τους. Στο στάδιο αυτό αντιμετωπίζονται διάφορα προβλήματα στην ποιότητα της εικόνας λόγω χαμηλής ποιότητας αρχικού εγγράφου ή λόγω κακής ψηφιοποίησης ή λόγω παλαιότητας του εγγράφου κ.λ.π. Το στάδιο αυτό είναι ιδιαίτερα σημαντικό γιατί βοηθάει στην επιτυχία των επόμενων σταδίων επεξεργασίας του εγγράφου που αφορούν την κατάτμηση και αναγνώριση των χαρακτήρων. Τα βασικά προβλήματα στην ποιότητα της εικόνας που πρέπει να αντιμετωπιστούν είναι τα ακόλουθα:

- Θόρυβος μορφής αλατοπίπερου (salt and pepper noise):** Είναι η παρουσία μεμονωμένων μαύρων pixels σε λευκές περιοχές ή μεμονωμένων λευκών pixels σε μαύρες περιοχές. Ο θόρυβος αυτός μπορεί να προέλθει λόγω ψηφιοποίησης ή δυαδικής μετατροπής (σχήμα 3.1α).
- Θόρυβος αποτελούμενος από τμήματα εικόνας μικρού μεγέθους:** Τμήματα της εικόνας με μικρό εμβαδό τα οποία έχουν προκύψει λόγω θορύβου ή λόγω του επιπέδου περιγραφής του εγγράφου κατά την ψηφιοποίηση (σχήμα 3.1β).
- Οπές στο σώμα των χαρακτήρων:** Το σώμα των χαρακτήρων αντί να είναι συμπαγές, παρουσιάζει οπές. Αυτό μπορεί να δημιουργήσει ιδιαίτερο πρόβλημα στην αναπαράσταση και αναγνώριση των χαρακτήρων (σχήμα 3.1γ).
- Ασυνέχειες στο σώμα των χαρακτήρων:** Δεν υπάρχει συνεκτικότητα των χαρακτήρων. Σπασίματα ή κενά εμφανίζονται στο σώμα των χαρακτήρων. Ιδιαίτερα σημαντικό πρόβλημα για το στάδιο της κατάτμησης των χαρακτήρων (σχήμα 3.1δ).
- Ενωμένα γράμματα:** Λόγω κακής ψηφιοποίησης ή ύπαρξης θορύβου ανάμεσα στους χαρακτήρες, γειτονικοί χαρακτήρες ενώνονται μεταξύ τους (σχήμα 3.1ε).
- Μη ομαλή περίμετρος των χαρακτήρων:** Η μη ομαλή περίμετρος των χαρακτήρων είναι πρόβλημα που εισάγεται κατά την ψηφιοποίηση και μπορεί να έχει επίπτωση στο στάδιο της αναγνώρισης των χαρακτήρων (σχήμα 3.1ζ).
- Εμφάνιση μαύρου περιθωρίου στα όρια της εικόνας:** Λόγω κακής ψηφιοποίησης του εγγράφου (υπερφωτισμός στα όρια του εγγράφου) ή λόγω δυαδικής μετατροπής εμφανίζεται ένα μαύρο περιθώριο στα όρια της εικόνας του εγγράφου (σχήμα 3.2).
- Επικάλυψη συμβόλων του υποβάθρου με το κείμενο:** Σύμβολα του υποβάθρου (γραμμές ή περιοδικά εμφανιζόμενα σύμβολα) επικαλύπτουν τις περιοχές κειμένου (σχήμα 3.3).

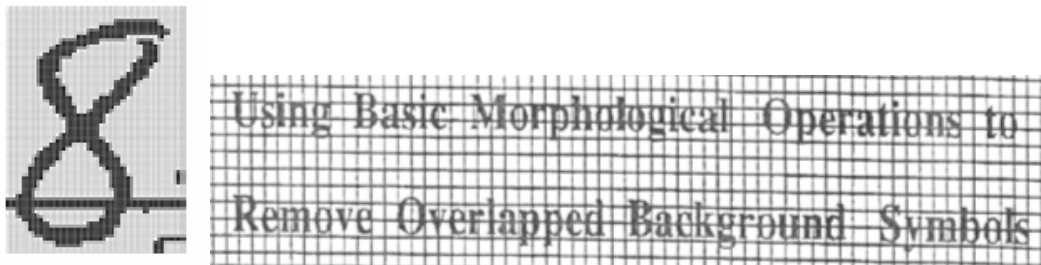
Για την αντιμετώπιση των παραπάνω προβλημάτων χρησιμοποιούνται μέθοδοι που βασίζονται κυρίως στην χρήση μασκών, στην μαθηματική μορφολογία, σε φίλτρα συρρίκνωσης και διόγκωσης καθώς και σε ειδικούς αλγορίθμους που εντοπίζουν και αφαιρούν τα μαύρα περιθώρια στα όρια της εικόνας.



Σχήμα 3.1. Προβλήματα στην ποιότητα των δυαδικών εικόνων. (α) Θόρυβος μορφής αλατοπίπερου. (β) Θόρυβος αποτελούμενος από τμήματα εικόνας μικρού μεγέθους. (γ) Οπές στο σώμα των χαρακτήρων. (δ) Ασυνέχειες στο σώμα των χαρακτήρων. (ε) Ενωμένα γράμματα. (ζ) Μη ομαλή περίμετρος των χαρακτήρων.



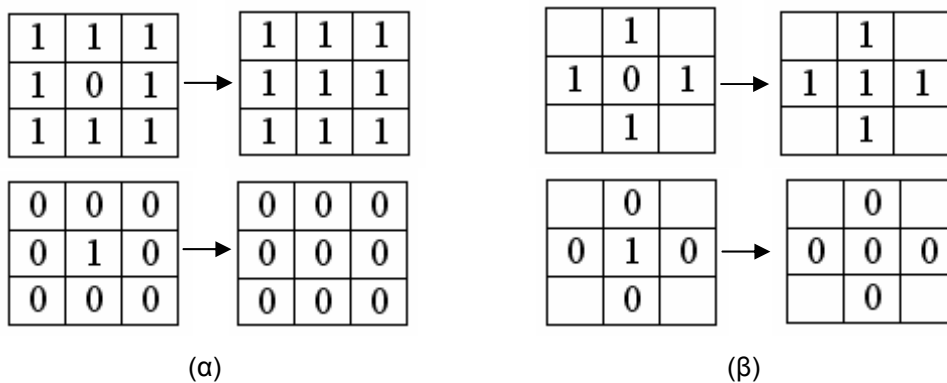
Σχήμα 3.2. Εμφάνιση μαύρου περιθωρίου στα όρια της εικόνας.



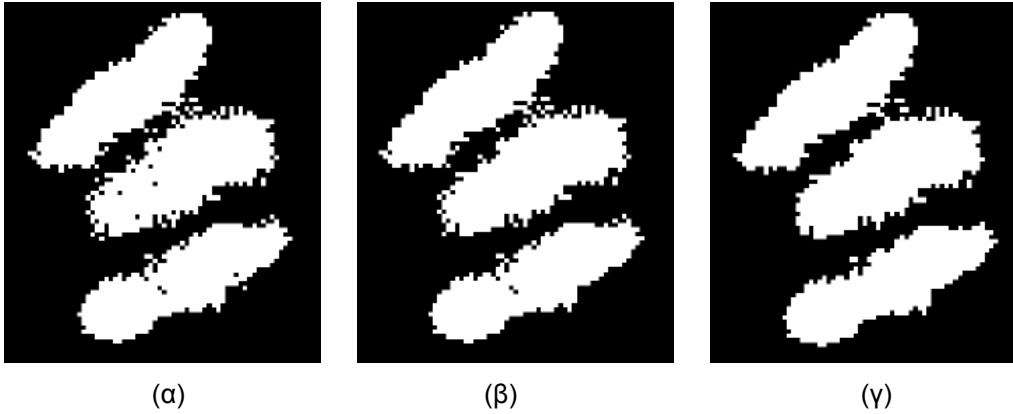
Σχήμα 3.3. Σύμβολα του υποβάθρου (γραμμές ή περιοδικά εμφανιζόμενα σύμβολα) επικαλύπτουν τις περιοχές κειμένου.

3.2 Χρήση μασκών

Εφαρμόζονται διάφορες μάσκες σε κάθε σημείο της επιφάνειας της εικόνας. Οι πιο απλές μάσκες είναι οι 3x3 μάσκες βάσει των 8 ή των 4 γειτονικών pixels (σχήμα 3.4) και χρησιμοποιούνται για την αφαίρεση του θορύβου μορφής αλατοπίπερου (salt and pepper noise). Αν η μάσκα ταιριάζει στην εικόνα τότε η εικόνα μετατρέπεται όπως στο σχήμα 3.4. Στο σχήμα 3.5 φαίνεται το αποτέλεσμα της χρήσης μασκών για αφαίρεση θορύβου μορφής αλατοπίπερου.

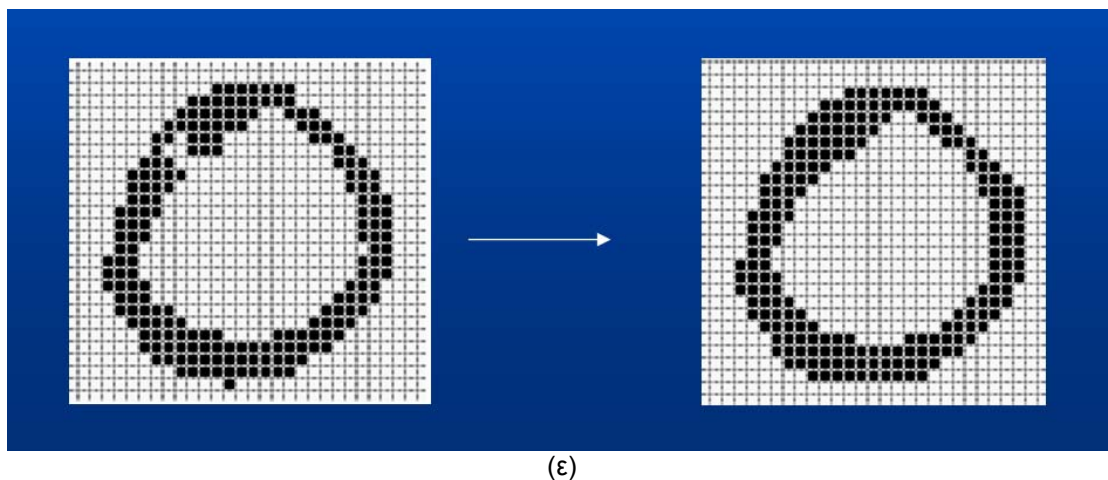
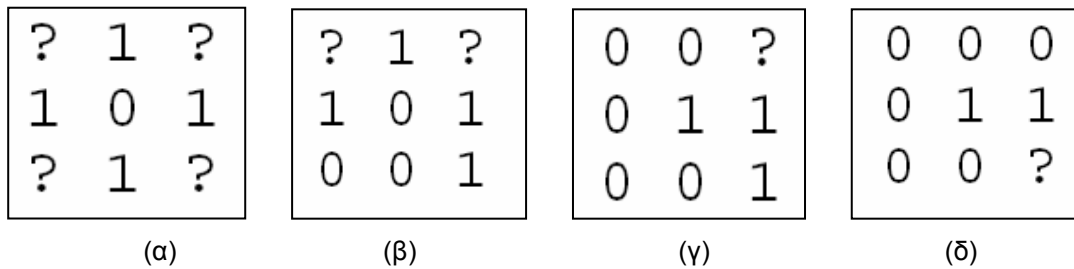


Σχήμα 3.4. Χρήση μασκών για αφαίρεση θορύβου μορφής αλατοπίπερου (salt and pepper noise). (α) Βάσει των 8 γειτονικών pixels. (β) Βάσει των 4 γειτονικών pixels.



Σχήμα 3.5. Αποτέλεσμα της χρήσης масκών για αφαίρεση θορύβου μορφής αλατοπίπερου (salt and pepper noise). (α) Αρχική εικόνα. (β) Αφαίρεση θορύβου χρησιμοποιώντας την μάσκα των 8 γειτονικών pixels. (γ) Αφαίρεση θορύβου χρησιμοποιώντας την μάσκα των 4 γειτονικών pixels.

Οι μάσκες χρησιμοποιούνται και για την εξομάλυνση της επιφάνειας και βελτίωση της ποιότητας των χαρακτήρων. Σύμφωνα με την προσέγγιση (Suen 1992), αν η εικόνα ταιριάζει με την μάσκα του σχήματος 3.6α, τότε το κεντρικό pixel μετατρέπεται σε pixel εικόνας. Ίδια μετατροπή πραγματοποιείται όταν η εικόνα ταιριάζει με την μάσκα του σχήματος 3.6β ή τις μάσκες που προκύπτουν με στροφή κατά 90° , 180° και 270° . Αντίστοιχα, διαγραφή του κεντρικού pixel πραγματοποιείται όταν ταιριάζει στην εικόνα η μάσκα του σχήματος 3.6γ ή η μάσκα του σχήματος 3.6δ ή οποιαδήποτε από τις έξι μάσκες προκύπτουν με στροφή κατά 90° , 180° και 270° των δύο αυτών масκών. Στο σχήμα 3.6ε δίδεται ένα παράδειγμα εξομάλυνσης της επιφάνειας και βελτίωσης της ποιότητας των χαρακτήρων με χρήση масκών.



Σχήμα 3.6. Εξομάλυνση της επιφάνειας και βελτίωση της ποιότητας των χαρακτήρων με χρήση масκών. (α)-(δ) Μάσκες που χρησιμοποιούνται. (ε) Τελικό αποτέλεσμα.

3.3 Μαθηματική μορφολογία

Η μαθηματική μορφολογία, η οποία αναπτύχθηκε στα τέλη της δεκαετίας του 1960, είναι ένα σχετικά ανεξάρτητο κομμάτι της επεξεργασίας εικόνας. Οι δύο πρωτοπόροι της μαθηματικής μορφολογίας είναι ο Matheron (Matheron 1975) και ο Serra (Serra 1982) οι οποίοι έχουν γράψει σχετικά βιβλία με έμφαση στα μαθηματικά. Η μαθηματική μορφολογία αναφέρεται στην μελέτη της γεωμετρίας, της τοπολογίας και της μορφής των αντικειμένων. Αφορά συγκεκριμένες πράξεις όπου μία εικόνα αλληλεπιδρά με ένα δομικό στοιχείο (structuring element) και μετατρέπεται σε μια πιο απλοποιημένη και λειτουργική μορφή κρατώντας συγχρόνως τα βασικά χαρακτηριστικά της μορφής της.

Οι μορφολογικές πράξεις χρησιμοποιούνται κυρίως για τους παρακάτω σκοπούς:

- Προεπεξεργασία της εικόνας και βελτίωση της δομής των αντικειμένων (φιλτράρισμα θορύβου, βελτίωση της ποιότητας και απλοποίηση των αντικειμένων, σκελετός, λέπτυνση, πάχυνση της εικόνας).
- Ποσοτική περιγραφή των αντικειμένων (χαρακτηριστικά εμβαδού, περιμέτρου, προβολών).

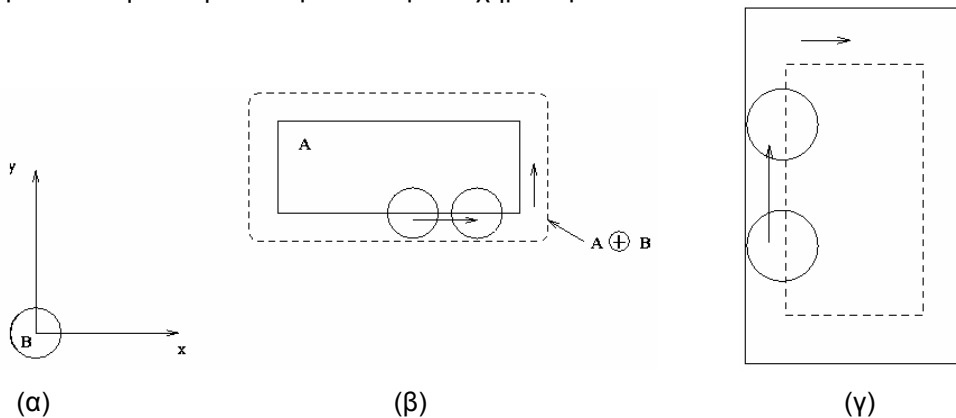
Οι περισσότερες μορφολογικές πράξεις μπορούν να οριστούν χρησιμοποιώντας τις δύο βασικές μορφολογικές πράξεις οι οποίες είναι το erosion (διάβρωση) και το dilation (διαστολή). Έστω το αντικείμενο A και το δομικό στοιχείο B εικόνες του δισδιάστατου χώρου. B_x είναι η μετατροπή του B ώστε η αρχή του είναι το x . Το erosion της εικόνας A με δομικό στοιχείο B ορίζεται σαν το σύνολο των σημείων x τέτοια ώστε το B_x περιλαμβάνεται στην εικόνα A :

$$A \ominus B = \{x : B_x \subseteq A\}. \quad (3.1)$$

Αντίστοιχα, το dilation της εικόνας A με δομικό στοιχείο B ορίζεται σαν το σύνολο των σημείων x τέτοια ώστε το B_x και η εικόνα A έχουν μη κενή τομή:

$$A \oplus B = \{x : \hat{B}_x \cap A \neq \emptyset\}. \quad (3.2)$$

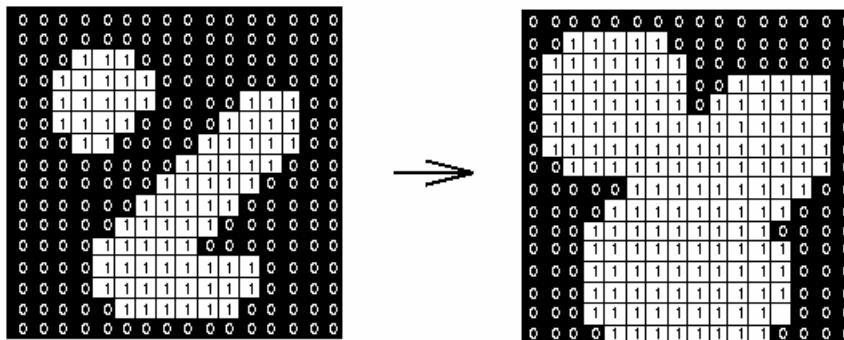
Στο σχήμα 3.7 φαίνεται μία αναπαράσταση του dilation και erosion της εικόνας A με δομικό στοιχείο B . Η αρχική εικόνα A έχει συνεχόμενη γραμμή ενώ με διακεκομμένη γραμμή αναπαράσσεται η τελική εικόνα μετά τον μετασχηματισμό.



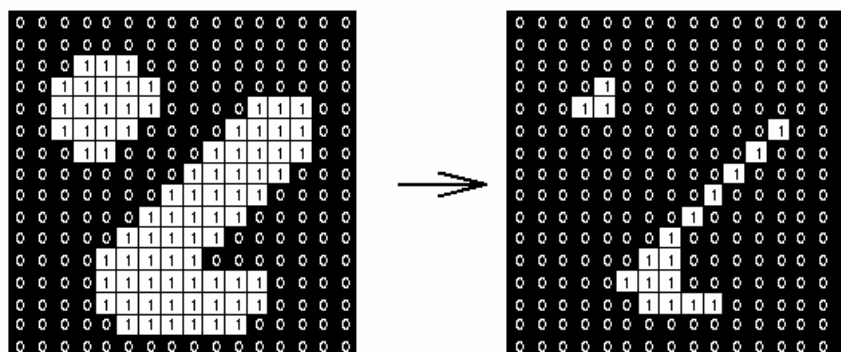
Σχήμα 3.7. Αναπαράσταση του dilation και erosion της εικόνας A με δομικό στοιχείο B . (α) Το δομικό στοιχείο B . (β) Αναπαράσταση του dilation. (γ) Αναπαράσταση του erosion.

1	1	1
1	1	1
1	1	1

(α)



(β)



(γ)

Σχήμα 3.8. Εφαρμογή των πράξεων dilation και erosion σε δυαδική εικόνα. (α) Το δομικό στοιχείο. (β) Εφαρμογή του dilation. (γ) Εφαρμογή του erosion.

Η εφαρμογή του dilation σε μία ασπρόμαυρη εικόνα ισοδυναμεί με εφαρμογή του κεντρικού σημείου του δομικού στοιχείου σε κάθε σημείο του υποβάθρου και αν υπάρχει τομή με την εικόνα, τότε το σημείο αυτό μετατρέπεται σε σημείο εικόνας (σχήμα 3.8β). Δηλαδή το dilation ισοδυναμεί με διόγκωση-διαστολή του φόντου της εικόνας. Αντίστοιχα, η εφαρμογή του erosion ισοδυναμεί με εφαρμογή του κεντρικού σημείου του δομικού στοιχείου σε κάθε σημείο της εικόνας και αν όλα τα σημεία του δομικού στοιχείου δεν περιλαμβάνονται στην εικόνα, τότε το σημείο αυτό μετατρέπεται σε σημείο υποβάθρου (σχήμα 3.8γ). Το erosion ισοδυναμεί με διάβρωση-συρρίκνωση της εικόνας.

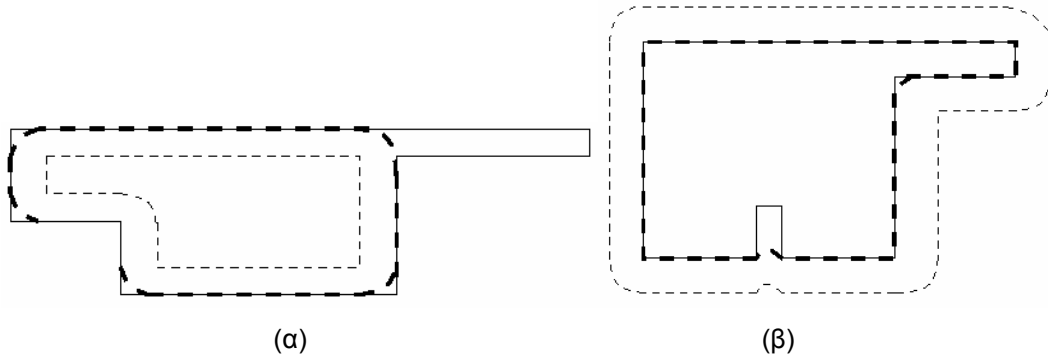
Άλλες δύο σημαντικές πράξεις της μαθηματικής μορφολογίας είναι το opening και το closing. Το opening είναι ένα erosion το οποίο ακολουθείται από ένα dilation:

$$A \circ B = (A \ominus B) \oplus B. \tag{3.3}$$

ενώ το closing είναι ένα dilation το οποίο ακολουθείται από ένα erosion:

$$A \bullet B = (A \oplus B) \ominus B. \tag{3.4}$$

Το opening και το closing δεν έχουν κάποια επίδραση στο πάχος της εικόνας (διόγκωση ή συρρίκνωση) όμως το opening βοηθάει στην μείωση του θορύβου, στην εξομάλυνση της περιμέτρου των αντικειμένων, στην απομόνωση αντικειμένων που έχουν ενωθεί λόγω κακής ψηφιοποίησης ή θορύβου, ενώ το closing βοηθάει στην κάλυψη οπών ή ασυνεχειών στο σώμα των αντικειμένων καθώς και στην εξομάλυνση της περιμέτρου των αντικειμένων. Στο σχήμα 3.9 φαίνεται μία αναπαράσταση του opening και του closing. Η αρχική εικόνα έχει συνεχόμενη γραμμή ενώ με έντονη διακεκομμένη γραμμή αναπαράσσεται η τελική εικόνα μετά τον μετασχηματισμό. Με λεπτή διακεκομμένη γραμμή αναπαράσσεται ο ενδιάμεσος μετασχηματισμός (erosion για το opening και dilation για το closing).



Σχήμα 3.9. Αναπαράσταση του opening και του closing της εικόνας. (α) Αναπαράσταση του opening. (β) Αναπαράσταση του closing.

Στο σχήμα 3.10 δίδονται παραδείγματα εφαρμογής των βασικών μορφολογικών πράξεων για προεπεξεργασία δυαδικής εικόνας εγγράφου. Συγκεκριμένα, χρησιμοποιείται το opening για εξομάλυνση της περιμέτρου των χαρακτήρων, το erosion για απομόνωση των χαρακτήρων και το closing για την κάλυψη οπών ή ασυνεχειών στο σώμα των χαρακτήρων.

**Dr. P.N. C
Mining Sur
Holroyd Ro
Reading,
Berks.**



**Dr. P.N. C
Mining Sur
Holroyd Ro
Reading,
Berks.**

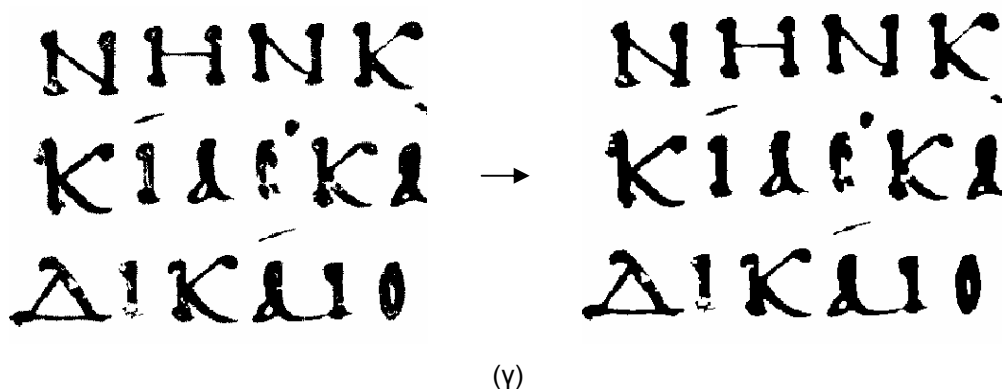
(α)

**Permit
transmission**



**Permit
transmission**

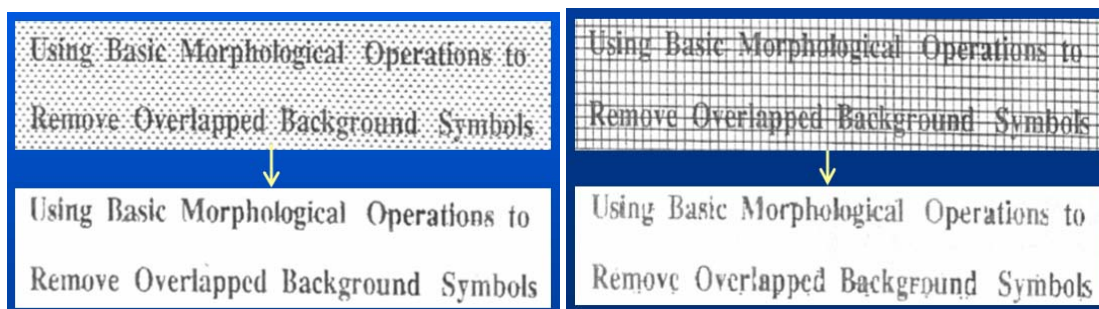
(β)



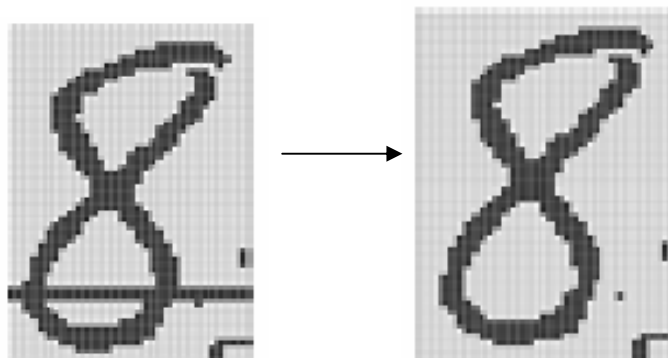
Σχήμα 3.10. Εφαρμογή μορφολογικών πράξεων για την προεπεξεργασία εικόνας εγγράφου. (α) Εφαρμογή opening για εξομάλυνση της περιμέτρου των χαρακτήρων. (β) Εφαρμογή erosion για απομόνωση των χαρακτήρων. (γ) Εφαρμογή closing για την κάλυψη οπών ή ασυνεχειών στο σώμα των χαρακτήρων.

Απομάκρυνση συμβόλων του υποβάθρου

Η μαθηματική μορφολογία μπορεί να χρησιμοποιηθεί για την απομάκρυνση συμβόλων του υποβάθρου τα οποία επικαλύπτουν τους χαρακτήρες σε εικόνα εγγράφου. Τα σύμβολα αυτά μπορεί να είναι περιοδικά επαναλαμβανόμενα σχήματα (σχήμα 3.11) ή ευθύγραμμα τμήματα (σχήμα 3.12).



Σχήμα 3.11. Παραδείγματα εφαρμογής της μαθηματικής μορφολογία για την απομάκρυνση περιοδικά επαναλαμβανόμενων συμβόλων του υποβάθρου που επικαλύπτουν τους χαρακτήρες.



Σχήμα 3.12. Παράδειγμα εφαρμογής της μαθηματικής για την απομάκρυνση ευθύγραμμων τμημάτων που επικαλύπτουν τους χαρακτήρες.

Σύμφωνα με την προσέγγιση του (Liang 1994), για να εντοπίσουμε την οριζόντια και κατακόρυφη απόσταση PDH και PDV μεταξύ των συμβόλων του υποβάθρου (σχήμα 3.13) χρησιμοποιούμε τους παρακάτω τύπους:

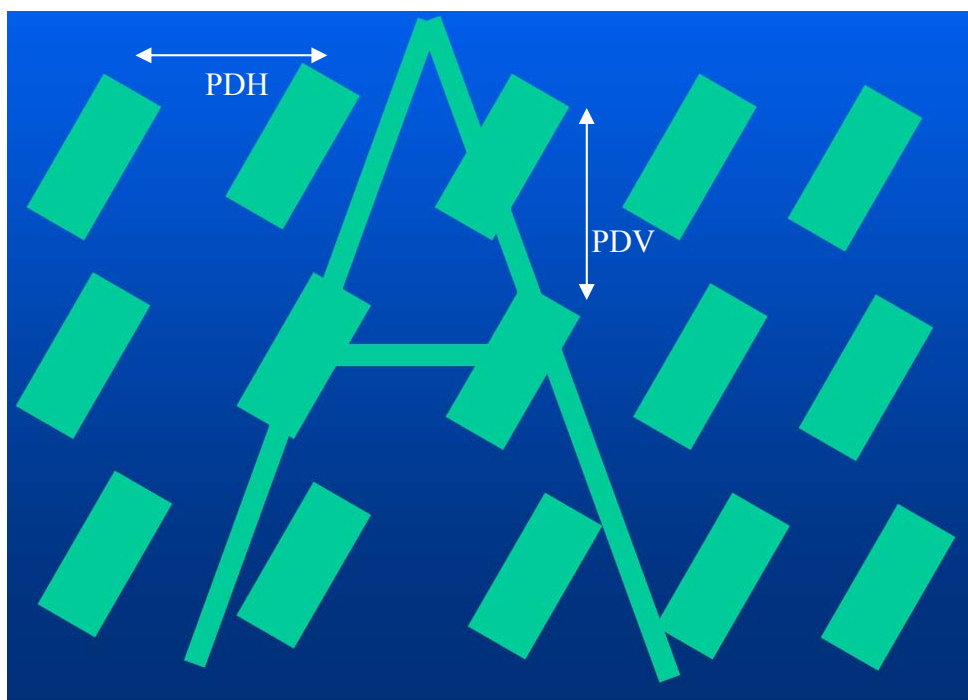
$$CL((X \ominus T1) \ominus B_{PDH}) > CL((X \ominus T1) \ominus B_i) \quad (3.5)$$

$i = 1, 2, \dots, M; i \neq PDH$

$$CL((X \ominus T2) \ominus B_{PDV}) > CL((X \ominus T2) \ominus B_j) \quad (3.6)$$

$j = 1, 2, \dots, M; j \neq PDV$

όπου X η αρχική ασπρόμαυρη εικόνα, $CL(x)$ συνάρτηση που μετράει τα σημεία της εικόνας x και $T1, T2, B_i, B_j$ δομικά στοιχεία που δίδονται στο σχήμα 3.14. Το erosion της εικόνας με το δομικό στοιχείο $T1$ μας δίνει το από αριστερά περίγραμμα της εικόνας (σχήματα 3.15 και 3.16) ενώ με το $T2$ μας δίνει το από πάνω περίγραμμα της εικόνας. Σύμφωνα με τους παραπάνω τύπους, το τελικό erosion με τα δομικά στοιχεία B_i και B_j μας δίνει τον μέγιστο αριθμό σημείων όταν $B_i = B_{PDH}$ και $B_j = B_{PDV}$.



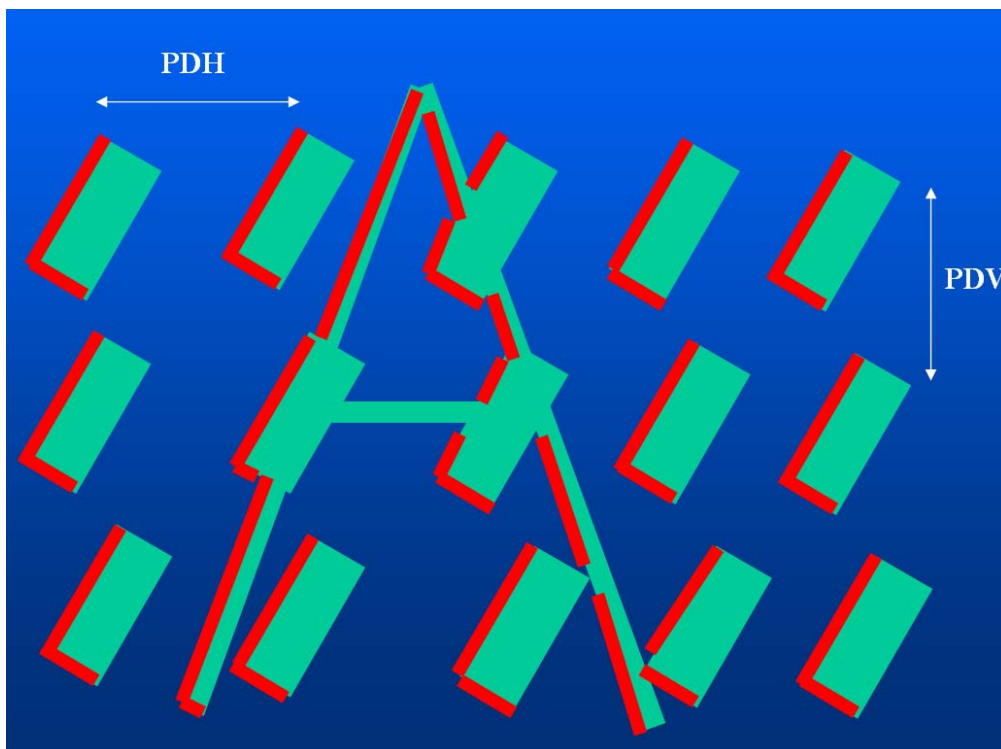
Σχήμα 3.13. Παραδείγματα επαναλαμβανόμενων συμβόλων του υποβάθρου που επικαλύπτουν τους χαρακτήρες. PDH και PDV είναι η οριζόντια και κατακόρυφη απόσταση μεταξύ των συμβόλων του υποβάθρου.

$$T1 = \{ \circ \downarrow \} \quad T2 = \{ \circ \leftarrow \}$$

$$B_i = \{ \downarrow \bullet \bullet \bullet \bullet \bullet \} \quad B_j = \{ \leftarrow \bullet \bullet \bullet \bullet \bullet \} \quad j$$

i

Σχήμα 3.14. Δομικά στοιχεία που χρησιμοποιούνται για τον εντοπισμό των επαναλαμβανόμενων συμβόλων του υποβάθρου.



Σχήμα 3.15. Αποτέλεσμα του erosion της εικόνας του σχήματος 3.13 χρησιμοποιώντας το δομικό στοιχείο T1 (οι κόκκινες γραμμές).



Σχήμα 3.16. Αποτέλεσμα του erosion της εικόνας χρησιμοποιώντας το δομικό στοιχείο T1. (α) Η αρχική εικόνα. (β) Η εικόνα μετά το erosion με το δομικό στοιχείο T1.

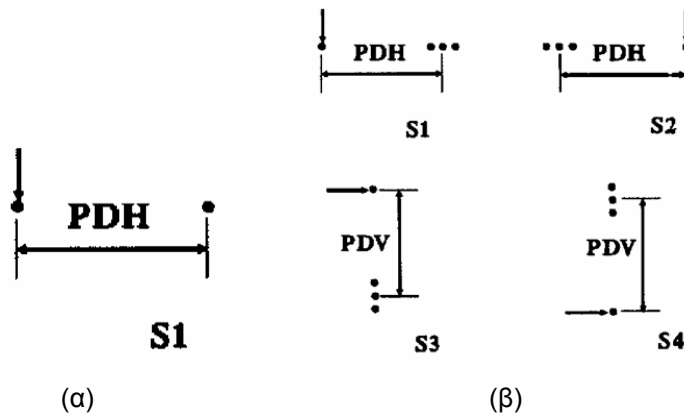
Αφού έχουμε εντοπίσει την οριζόντια και κατακόρυφη απόσταση PDH και PDV μεταξύ των συμβόλων του υποβάθρου, μπορούμε να εξαγάγουμε το υπόβαθρο κάνοντας ένα erosion με δομικά στοιχεία τα οποία είναι της μορφής του σχήματος 3.17α, δηλαδή έχουν άσους σε απόσταση PDH ή PDV. Για την καλύτερη εφαρμογή της μεθόδου, χρησιμοποιούνται τα δομικά στοιχεία του σχήματος 3.17β όπου από τα τρία συνεχόμενα pixels αρκεί ένα να είναι άσπος για να ανήκει το κεντρικό pixel (με το βελάκι) στην νέα εικόνα. Χρησιμοποιούνται δύο οριζόντια (S1: από αριστερά και S2: από δεξιά) και δύο κάθετα (S3: από πάνω και S4: από κάτω) δομικά στοιχεία για να μπορεί να γίνει έλεγχος ακόμα και των ακραίων σημείων της εικόνας. Η εικόνα του υποβάθρου προκύπτει με διαδοχική και επαναληπτική εφαρμογή του erosion της εικόνας με τα δομικά στοιχεία S1, S2, S3 και S4:

$$RE_1(X)_{\{S1|S2|S3|S4\}} = (((((X \ominus_{(LR)} S1) \ominus_{(TB)} S3) \ominus_{(RL)} S2) \ominus_{(BT)} S4))), \quad (3.7)$$

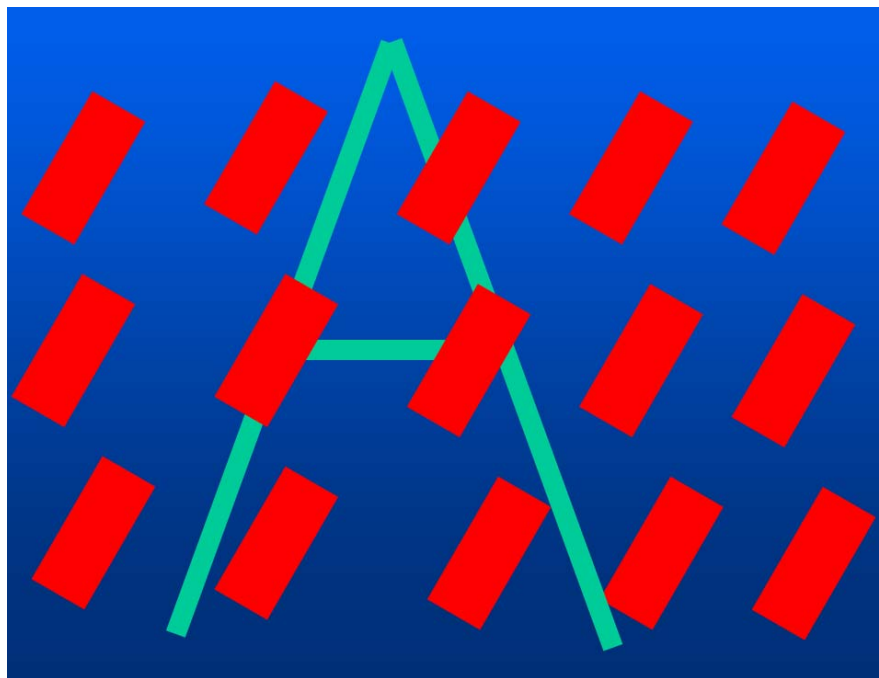
Η επαναληπτική εφαρμογή των παραπάνω erosions σταματάει όταν:

$$X_i = X_{(i-1)} \ominus S_j, \quad j = 1, 2, 3 \text{ and } 4. \quad (3.8)$$

Το τελικό αποτέλεσμα εντοπισμού των επαναλαμβανόμενων συμβόλων του υποβάθρου φαίνεται στα σχήματα 3.18 και 3.19.



Σχήμα 3.17. Δομικά στοιχεία που χρησιμοποιούνται για τον εντοπισμό των επαναλαμβανόμενων συμβόλων του υποβάθρου.



Σχήμα 3.18. Το αποτέλεσμα της διαδικασίας εντοπισμού των διαδοχικών συμβόλων του υποβάθρου για την εικόνα 3.13.



Σχήμα 3.19. Το αποτέλεσμα της διαδικασίας εντοπισμού των διαδοχικών συμβόλων του υποβάθρου σε εικόνα κειμένου. (α)-(δ) Τα αποτελέσματα των διαδοχικών επαναλήψεων.

Β. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

Αν X η αρχική εικόνα και Y η εικόνα του υποβάθρου που έχουμε εντοπίσει με την παραπάνω διαδικασία, τότε η τελική εικόνα Z που περιέχει μόνο τα pixels του κειμένου δίδεται από τον τύπο:

$$Z = \text{XOR} (X, Y) \quad (3.9)$$

Για να βελτιώσουμε την ποιότητα του τελικού αποτελέσματος, μπορούμε να εφαρμόσουμε επιπλέον και την παρακάτω πράξη:

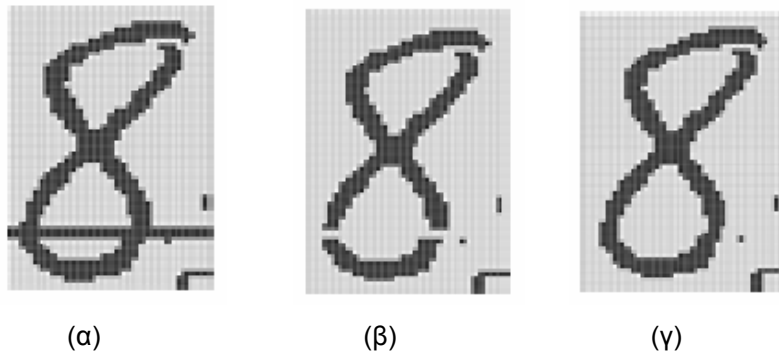
$$W = ((Z \oplus B1) \cap Y) \bullet B2 \quad (3.10)$$

όπου $B1$ και $B2$ δύο 5×5 τετραγωνικά δομικά στοιχεία. Το τελικό αποτέλεσμα της διαδικασίας φαίνεται στο σχήμα 3.20.

Για την περίπτωση που το κείμενο επικαλύπτεται με οριζόντιες ή κάθετες γραμμές, για τον εντοπισμό των γραμμών αρκεί ένα erosion οριζόντιο ή κάθετο με δομικό στοιχείο μήκους όσο το ελάχιστο αναμενόμενο μήκος των οριζόντιων ή κάθετων γραμμών (ΥΕ 2000). Το αποτέλεσμα αυτής της διαδικασίας δίδεται στο σχήμα 3.21.



Σχήμα 3.20. Το αποτέλεσμα της διαδικασίας αφαίρεσης των διαδοχικών συμβόλων του υποβάθρου από εικόνα κειμένου. (α) Η αρχική εικόνα. (β) Το τελικό αποτέλεσμα.



Σχήμα 3.21. Το αποτέλεσμα της διαδικασίας αφαίρεσης οριζόντιων και κάθετων γραμμών από εικόνα κειμένου. (α) Η αρχική εικόνα. (β) Η εικόνα μετά την αφαίρεση των γραμμών. (γ) Η τελική εικόνα.

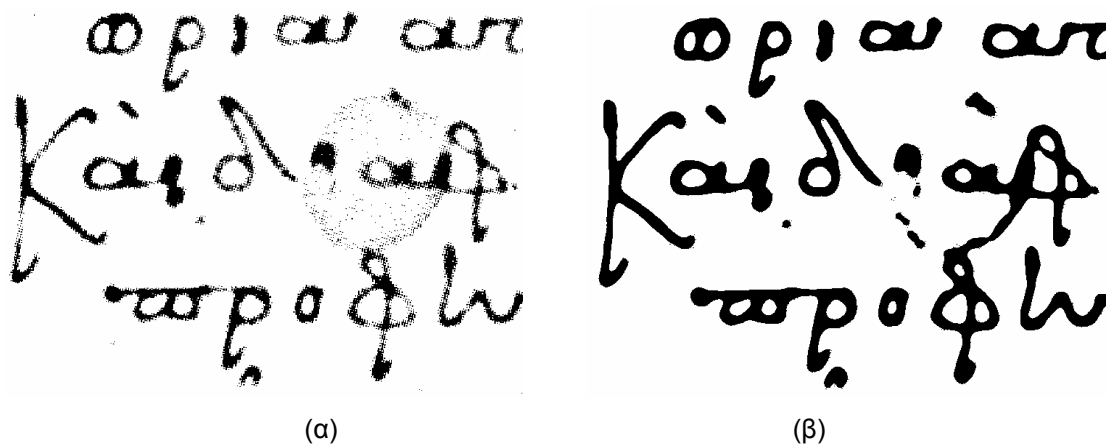
3.4 Φίλτρα συρρίκνωσης και διόγκωσης (shrink & swell)

Τα φίλτρα συρρίκνωσης (shrink) και διόγκωσης (swell) είναι μία παραλλαγή των μορφολογικών πράξεων erosion και dilation (Schilling 2000). Ένα $n \times n$ παράθυρο εφαρμόζεται σε κάθε σημείο εικόνας ή υποβάθρου και αφού υπολογίζεται το άθροισμα των σημείων υποβάθρου ή εικόνας μέσα στο παράθυρο αποφασίζεται η μετατροπή ή όχι του συγκεκριμένου σημείου. Τα φίλτρα συρρίκνωσης και διόγκωσης εφαρμόζονται στις δυαδικές εικόνες εγγράφων ώστε να μειωθεί ο θόρυβος, να βελτιωθεί η ποιότητα των περιοχών κειμένου καθώς και η συνεκτικότητα των χαρακτήρων αφαιρώντας απομονωμένα pixels και γεμίζοντας τυχόν σπασίματα, κενά ή τρύπες στο σώμα των χαρακτήρων. Πιο συγκεκριμένα:

Το φίλτρο συρρίκνωσης (shrink filter) χρησιμοποιείται για την αφαίρεση θορύβου από το υπόβαθρο της εικόνας. Σαρώνεται ολόκληρη η ασπρόμαυρη εικόνα και εξετάζεται κάθε σημείο κειμένου. Αν P_{sh} είναι ο αριθμός των pixels του υποβάθρου σε ένα κινούμενο παράθυρο $n \times n$, το οποίο έχει το σημείο του κειμένου σαν κεντρικό, τότε το σημείο μετατρέπεται σε σημείο υποβάθρου της εικόνας αν $P_{sh} > k_{sh}$ όπου k_{sh} ορίζεται πειραματικά.

Το φίλτρο διόγκωσης (swell filter) χρησιμοποιείται για το γέμισμα πιθανών ασυνχειών, κενών ή οπών στην επιφάνεια των χαρακτήρων. Σαρώνεται ολόκληρη η ασπρόμαυρη εικόνα και εξετάζεται κάθε σημείο υποβάθρου. Αν P_{sw} είναι ο αριθμός των pixels του κειμένου σε ένα κινούμενο παράθυρο $n \times n$, το οποίο έχει το σημείο (x,y) του υποβάθρου σαν κεντρικό, τότε το σημείο μετατρέπεται σε σημείο κειμένου της εικόνας αν $P_{sw} > k_{sw}$ όπου k_{sw} ορίζεται πειραματικά. Για να αποφύγουμε την πάχυνση των χαρακτήρων μπορούμε εκτός από την συνθήκη $P_{sw} > k_{sw}$ να ελέγχουμε και τις συνθήκες $|x-x_a| < dx$ και $|y-y_a| < dy$ για την μετατροπή ενός σημείου υποβάθρου σε κείμενο, όπου x_a, y_a οι μέσες τιμές των συντεταγμένων x, y στο παράθυρο $n \times n$. Με αυτό τον τρόπο ελέγχουμε μόνο τα σημεία υποβάθρου που βρίσκονται ανάμεσα σε ομοιόμορφα κατανεμημένα σημεία κειμένου.

Η διαδοχική εφαρμογή των παραπάνω φίλτρων σε δυαδικές εικόνες εγγράφων έχει σημαντικά αποτελέσματα ως προς την βελτίωση της ποιότητας (σχήμα 3.22).

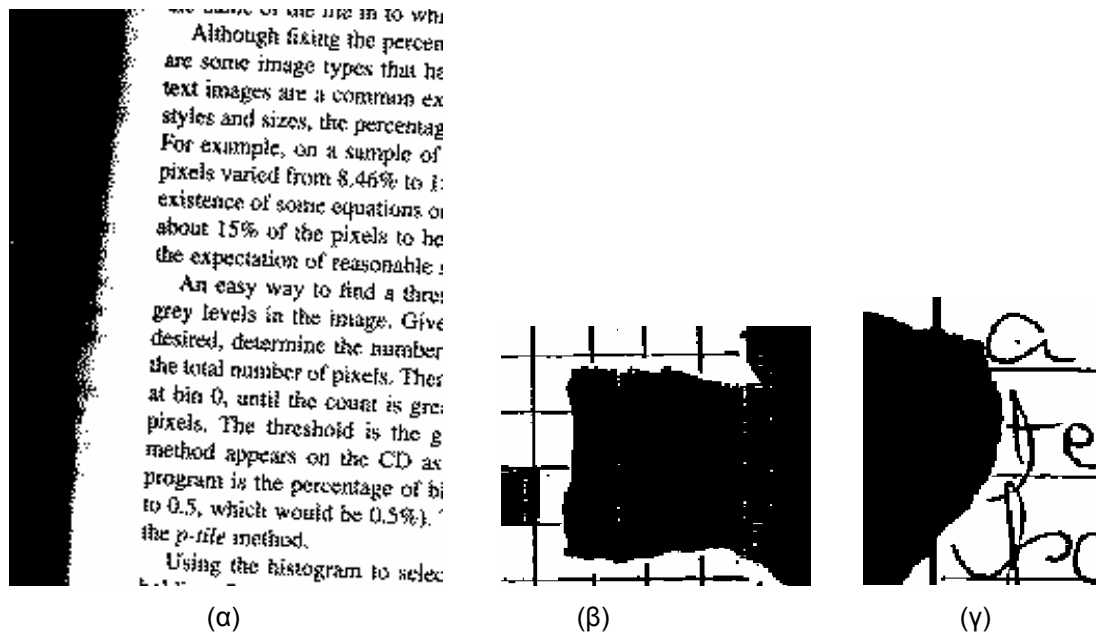


Σχήμα 3.22. Παράδειγμα διαδοχικής εφαρμογής φίλτρων συρρίκνωσης (shrink) και διόγκωσης (swell) σε δυαδική εικόνα εγγράφου. (α) Αρχική ασπρόμαυρη εικόνα. (β) Τελική εικόνα μετά την εφαρμογή των φίλτρων.

3.5 Αλγόριθμος αφαίρεσης περιθωρίου (Avila)

Λόγω κακής ψηφιοποίησης του εγγράφου (υπερφωτισμός στα όρια του εγγράφου) ή λόγω δυαδικής μετατροπής είναι αρκετά συνηθισμένο να εμφανίζεται ένα μαύρο περιθώριο στα όρια της εικόνας του εγγράφου (σχήμα 3.2). Το περιθώριο αυτό αλλοιώνει την όψη του εγγράφου, δυσκολεύει την παραπέρα επεξεργασία, κατάτμηση και αναγνώριση του ενώ ταυτόχρονα μεγαλώνει το μέγεθός του. Μερικά από τα χαρακτηριστικά του μαύρου περιθωρίου είναι τα ακόλουθα:

- Βρίσκεται στα όρια της εικόνας (σχήμα 3.23α).
- Συνήθως είναι συμπαγές αλλά υπάρχουν και περιθώρια με θόρυβο (σχήμα 3.23β).
- Μπορεί να εισχωρεί αρκετά στο έγγραφο (περιπτώσεις όπου λείπει κομμάτι του εγγράφου) (σχήμα 3.23γ).
- Μπορεί να επικαλύπτει τις περιοχές κειμένου του εγγράφου (σχήμα 3.23γ).



Σχήμα 3.23. Παράδειγμα εμφάνισης μαύρου περιθωρίου σε εικόνα εγγράφου.

Σύμφωνα με την προσέγγιση των Avila & Lins (Avila 2004) προτείνονται δύο αλγόριθμοι για τον εντοπισμό και αφαίρεση του μαύρου περιθωρίου των εγγράφων. Ο πρώτος αλγόριθμος (εισβολής) προϋποθέτει ότι το μαύρο περιθώριο δεν επικαλύπτει περιοχές κειμένου. Ο δεύτερος αλγόριθμος (μη εισβολής) εξασφαλίζει ότι αν το μαύρο περιθώριο επικαλύπτει περιοχές κειμένου τότε αυτές παραμένουν στην εικόνα και δεν αφαιρούνται.

Αλγόριθμος εισβολής

Σύμφωνα με τον αλγόριθμο αυτό εισβάλουμε από τα σημεία του μαύρου περιθωρίου που βρίσκονται στα όρια της εικόνας προς το εσωτερικό της εικόνας. Τα βήματα του αλγορίθμου έχουν ως εξής:

Βήμα 1: Δημιουργία λίστας με τα σημεία που γειτονεύουν με τα όρια της εικόνας.

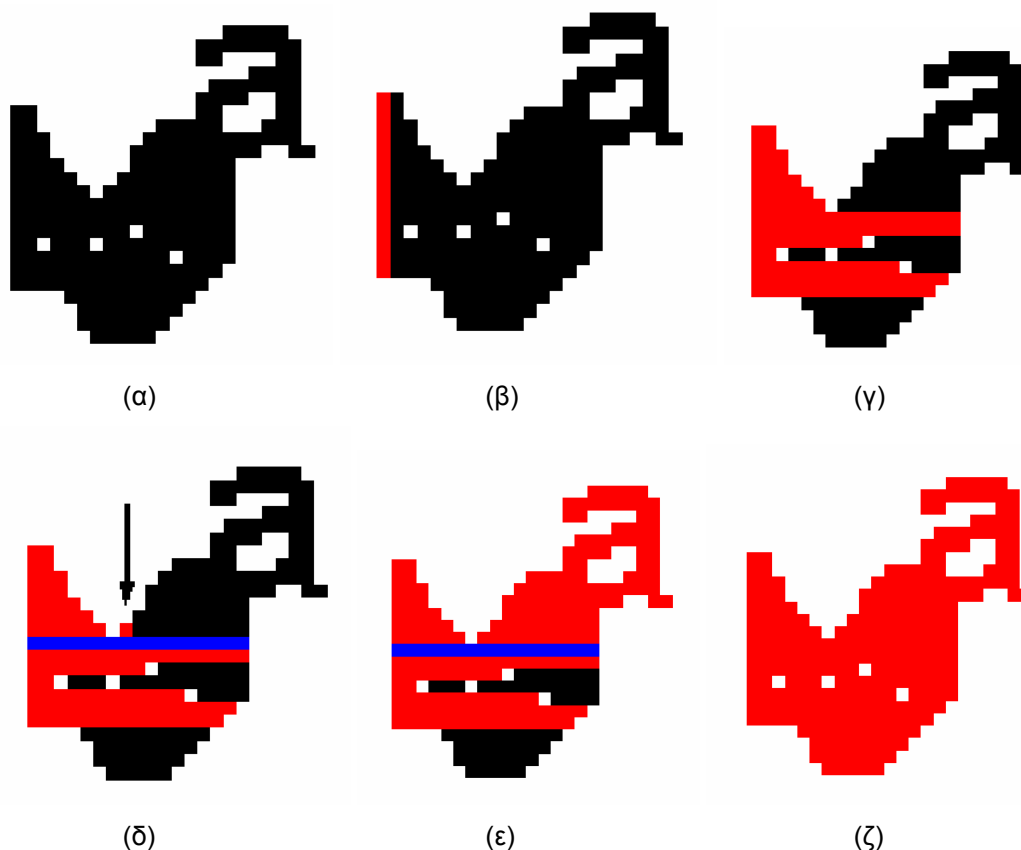
Βήμα 2: Μετακίνηση από κάθε σημείο της λίστας του βήματος 1 προς τα δεξιά και προς τα αριστερά μέχρι να συναντήσουμε σημείο φόντου. Όλα τα νέα σημεία προστίθενται στην αρχική λίστα.

Βήμα 3: Μαρκάρισμα όλων των σημείων της λίστας σαν *border_0*.

Βήμα 4: Από κάθε σημείο της λίστας μετακινούμαστε προς τα πάνω και προς τα κάτω μέχρι να συναντήσουμε σημείο φόντου. Τα νέα σημεία προστίθενται στην λίστα και για κάθε νέο σημείο επαναλαμβάνουμε τα βήματα 2 μέχρι 4.

Βήμα 5: Όλα τα μαρκαρισμένα σημεία ως *border_0* μετατρέπονται σε φόντο.

Στο παράδειγμα 3.24 δίδεται ένα παράδειγμα απλής εικόνας όπου εντοπίζεται το μαύρο περιθώριο της εικόνας με χρήση του αλγορίθμου εισβολής. Στο παράδειγμα αυτό, το γράμμα που επικαλύπτεται με το μαύρο περιθώριο αφαιρείται.



Σχήμα 3.24. Παράδειγμα αλγορίθμου εισβολής. (α) Αρχική εικόνα. (β) Εντοπισμός των σημείων που γειτονεύουν με τα όρια της εικόνας – Βήμα 1. (γ) Μετακίνηση προς τα δεξιά και αριστερά – Βήμα 2. (δ) Μετακίνηση προς τα πάνω και προς τα κάτω, εύρεση νέου σημείου – Βήμα 4. (ε) Επέκταση της λίστας των σημείων λόγω του νέου σημείου - επαναλήψεις βημάτων 2-4. (ζ) Τελική λίστα μαρκαρισμένων σημείων για διαγραφή.

Αλγόριθμος μη εισβολής

Σύμφωνα με τον αλγόριθμο αυτό η εισβολή των σημείων του μαύρου περιθωρίου που βρίσκονται στα όρια της εικόνας προς το εσωτερικό της εικόνας δεν επεκτείνεται σε pixels που ανήκουν στο κείμενο των εγγράφων. Για τον σκοπό αυτό εισάγονται δύο παράμετροι οι οποίες αφορούν την φύση του κειμένου που βρίσκεται στα έγγραφα. Η παράμετρος SEGMENT η οποία είναι το μέγιστο πάχος των γραμμών που ανήκουν σε χαρακτήρες και η παράμετρος LINE η οποία είναι το ελάχιστο μήκος των γραμμών που ανήκουν σε χαρακτήρες.

Ο αλγόριθμος μη εισβολής ακολουθεί τα βήματα 1 μέχρι 4 του αλγορίθμου εισβολής. Ανάμεσα στα βήματα 2 και 3 εισάγονται τα βήματα 2.1, 2.2, 2.3 και 2.4. Επίσης το βήμα 5 αντικαθίσταται από τα βήματα 6 μέχρι 9. Τα βήματα του αλγορίθμου έχουν ως εξής:

Βήμα 1: Δημιουργία λίστας με τα σημεία που γειτονεύουν με τα όρια της εικόνας.

Βήμα 2: Μετακίνηση από κάθε σημείο της λίστας του βήματος 1 προς τα δεξιά και προς τα αριστερά μέχρι να συναντήσουμε σημείο φόντου. Όλα τα νέα σημεία προστίθενται στην αρχική λίστα.

Βήμα 2.1: Ελέγχουμε το πάχος της οριζόντιας γραμμής που δημιουργείται και αν σε κάποιο σημείο της έχει πάχος μικρότερο από SEGMENT τότε ελέγχουμε την περιοχή αυτή (ίσως ανήκει σε κείμενο).

Βήμα 2.2: Αν το μήκος της κατακόρυφης προβολής (προς το εσωτερικό της εικόνας) της περιοχής που έχει εντοπιστεί στο βήμα 2.1 είναι μεγαλύτερο από LINE, τότε η οριζόντια

Β. Γάτος, Ψηφιακή Επεξεργασία και Αναγνώριση Εγγράφων

γραμμή είναι στα όρια του μαύρου περιθωρίου και του κειμένου. Τα pixels ανάμεσα στο περιθώριο και στο κείμενο ονομάζονται «σημεία αναφοράς».

Βήμα 2.3: Εξετάζουμε και τα κατακόρυφα τμήματα πάνω και κάτω από τα «σημεία αναφοράς» τα οποία έχουν πάχος μικρότερο από SEGMENT.

Βήμα 2.4: Για κάθε κατακόρυφο τμήμα του βήματος 2.3 εξετάζουμε αν το μήκος της κατακόρυφης προβολής (προς το εσωτερικό της εικόνας) είναι μεγαλύτερο από LINE. Αν είναι, τότε το κατακόρυφο τμήμα είναι το όριο του περιθωρίου με το κείμενο.

Βήμα 3: Μαρκάρισμα όλων των σημείων της λίστας σαν *border_0*.

Βήμα 4: Από κάθε σημείο της λίστας μετακινούμαστε προς τα πάνω και προς τα κάτω μέχρι να συναντήσουμε σημείο φόντου. Τα νέα σημεία προστίθενται στην λίστα και για κάθε νέο σημείο επαναλαμβάνουμε τα βήματα 2 μέχρι 4.

Βήμα 5: Δημιουργία λίστας *border_0* σημείων τα οποία συνορεύουν με τα όρια της εικόνας.

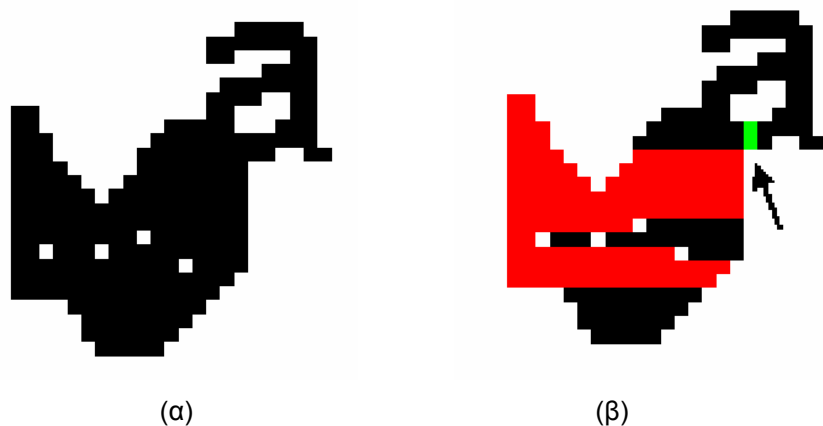
Βήμα 6: Μετακίνηση από τα σημεία του βήματος 5 προς τα δεξιά και αριστερά μέχρι να συναντήσουμε σημείο φόντου ή «σημείο αναφοράς». Όλα τα νέα σημεία προστίθενται στην λίστα του βήματος 5.

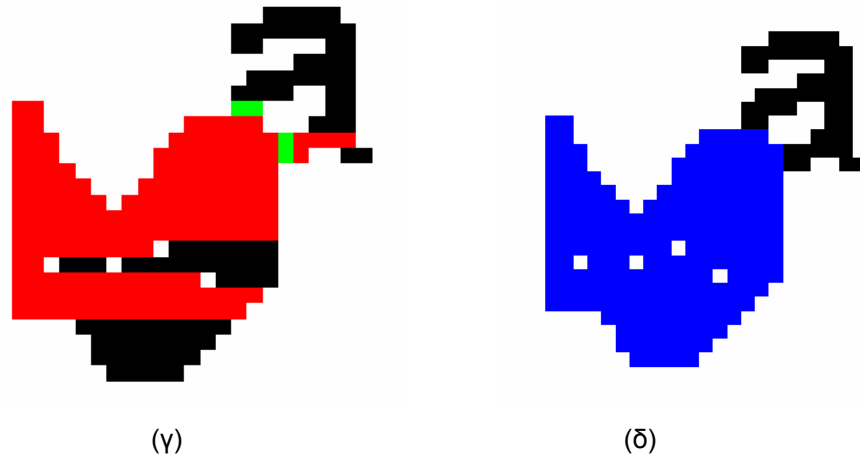
Βήμα 7: Όλα τα σημεία της νέας λίστας μαρκάρονται ως *border_1*.

Βήμα 8: Για κάθε σημείο της νέας λίστας ελέγχουμε τα προς τα πάνω και προς τα κάτω γειτονικά pixels αν είναι μαρκαρισμένα ως *border_0*. Τα νέα σημεία προστίθενται στην λίστα και επαναλαμβάνουμε τα βήματα 6 μέχρι 8.

Βήμα 9: Όλα τα μαρκαρισμένα σημεία ως *border_1* μετατρέπονται σε φόντο.

Στο σχήμα 3.25 δίδεται ένα παράδειγμα εφαρμογής του αλγορίθμου σε μία απλή εικόνα όπου το μαύρο περιθώριο επικαλύπτεται με το κείμενο. Τα δύο πρώτα βήματα του αλγορίθμου είναι ίδια με τον αλγόριθμο εισβολής μέχρι να συναντήσουμε την περίπτωση που περιγράφεται στα σχήμα 3.25β. Έχουμε ένα κατακόρυφο τμήμα με πάχος μικρότερο από SEGMENT. Θέλουμε να ελέγξουμε αν το τμήμα αυτό ανήκει στο περιθώριο ή στο κείμενο. Επεκτείνουμε το τμήμα προς τα δεξιά (σχήμα 3.25γ) και παρατηρούμε μήκος μεγαλύτερο από LINE. Άρα το κατακόρυφο τμήμα είναι στα όρια του μαύρου περιθωρίου και του κειμένου και τα σημεία του γίνονται «σημεία αναφοράς». Στο σχήμα 3.25δ φαίνεται το αποτέλεσμα της διαδικασίας μετά από την ολοκλήρωση του αλγορίθμου.





Σχήμα 3.25. Παράδειγμα εφαρμογής του αλγορίθμου μη εισβολής. (α) Η αρχική εικόνα. (β) Τα «σημεία αναφοράς» τα οποία ανήκουν σε τμήματα με πάχος μικρότερο από SEGMENT. (γ) Επέκταση των «σημείων αναφοράς» προς τα δεξιά και έλεγχος του μήκους αν είναι μεγαλύτερο από LINE. (δ) Τελικό αποτέλεσμα αλγορίθμου.

Βιβλιογραφία

- (Avila 2004)** Avila, B. T., Lins, R. D.: A new algorithm for removing noisy border from monochromatic documents. 2004 ACM Symposium on Applied Computing (2004)
- (Liang 1994)** Liang, Su, Ahmadi, M., Shridhar M.: A morphological approach to text string extraction from regular periodic overlapping text/background images. Proc. of IEEE Int. Conf. on Image Processing, (ICIP-94) (1994) 144-148
- (Matheron 1975)** Matheron, G.: Random Sets and Integral Geometry. New York, Wiley (1975)
- (Schilling 2000)** Schilling, R. J.: Fundamental of Robotics Analysis and Control. Prentice-Hall, Englewood Cliffs, NJ (1990)
- (Serra 1982)** Serra, J.: Image Analysis and Mathematical Morphology. New York, Academic Press (1982)
- (Suen 1992)** Suen, C.Y., Nadal, C., Legault, R., Mai, T.A., Lam, L.: Computer recognition of unconstrained handwritten numerals. Proc. IEEE, 80(7) (1992) 1162-1180
- (YE 2000)** Ye, X., Cheriet, M., Suen, C. Y.: A Generic System to extract and clean handwritten data from business forms. Proc. of the Seventh Int. Workshop on Frontiers in Handwriting Recognition (2000) 63-72