ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

# Διαχείριση Εικόνων Εγγράφων σε Πολυμεσικά Αρχειακά Συστήματα

Θεόδωρος Χ. Γκαραβέλος

Μεταπτυχιακή Εργασία

Ηράκλειο, Φεβρουάριος 1999

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

# Διαχείριση Εικόνων Εγγράφων σε Πολυμεσικά Αρχειακά Συστήματα

Εργασία που υποβλήθηκε από τον
Θεόδωρο Χ. Γκαραβέλο
ως μερική εκπλήρωση των απαιτήσεων
για την απόκτηση
ΜΕΤΑΠΤΥΧΙΑΚΟΥ ΔΙΠΛΩΜΑΤΟΣ ΕΙΔΙΚΕΥΣΗΣ

Συγγραφέας:

Θεόδωρος Χ. Γκαραβέλος
Τμήμα Επιστήμης Υπολογιστών

Εισηγητική Επιτροπή:

Πάνος Τραχανιάς
Αναπληρωτής Καθηγητής, Επόπτης

Απόστολος Τραγανίτης
Αναπληρωτής Καθηγητής, Μέλος

Πάνος Κωνσταντόπουλος
Καθηγητής, Μέλος

Δεκτή:

Πάνος Κωνσταντόπουλος
Καθηγητής
Πρόεδρος Επιτροπής Μεταπτυχιακών Σπουδών

Ηράκλειο, Φεβρουάριος 1999

# Ευχαριστίες

Στο σημείο αυτό θα ήθελα να ευχαριστήσω όλους όσους, με τον ένα ή τον άλλο τρόπο, με στήριξαν κατά τη διάρκεια της προσπάθειάς μου για την πραγματοποίηση αυτής της εργασίας.

Καταρχήν οφείλω να ευχαριστήσω τον επόπτη μου, τον Πάνο Τραχανιά, υπό την καθοδήγηση του οποίου πραγματοποιήθηκε αυτή η εργασία. Η συνεργασία μας ξεκίνησε από το τέταρτο έτος των σπουδών μου, στα πλαίσια της διπλωματικής μου εργασίας. Δυόμισι χρόνια ήταν αρκετά για να τον εκτιμήσω τόσο σαν επιστήμονα όσο και σαν άνθρωπο. Θέλω να τον ευχαριστήσω για την επιστημονική του καθοδήγηση, τις συμβουλές του και το χρόνο που μου αφιέρωσε, συζητώντας μαζί μου τα θέματα που με απασχόλησαν στην εργασία μου.

Σημαντική συμπαράσταση είχα και από το φίλο και συνεργάτη μου, Κώστα Χανδρινό. Η συνεργασία μας στα πλαίσια του προγράμματος, του οποίου τμήμα αποτέλεσε και αυτή η εργασία, ήταν ιδιαίτερα εποικοδομητική σε γνώσεις και εμπειρία. Με τη δική του βοήθεια και τις καίριες παρατηρήσεις του, μπόρεσα να ξεπεράσω μερικά από τα δύσκολα σημεία της εργασίας αυτής.

Θέλω να ευχαριστήσω όλους τους φίλους και συναδέλφους, οι οποίοι με υπομονή και όρεξη με βοήθησαν με τα σχόλια και τις παρατηρήσεις τους σε πολλά στάδια της εργασίας. Θέλω ακόμη να ευχαριστήσω τον Jonh Immerkær για τις πολύτιμες συμβουλές του στο σύντομο διάστημα παραμονής του στην Ελλάδα.

Επίσης πρέπει να ευχαριστήσω το Τμήμα Επιστήμης Υπολογιστών του Πανεπιστημίου Κρήτης, καθώς και το Ινστιτούτο Πληροφορικής του Ιδρύματος Τεχνολογίας και Έρευνας για την οικονομική ενίσχυση και υλικοτεχνική υποστήριξη που μου παρείχαν κατά τη διάρκεια των μεταπτυχιακών μου σπουδών.

Περισσότερο από όλους όμως θέλω να ευχαριστήσω τους ανθρώπους που με στήριξαν και με βοήθησαν ώστε να φτάσω ως εδώ. Τις αδερφές μου, Γεωργία και Μαρία, για την αγάπη τους που μου έδινε δύναμη, και τους δύο ανθρώπους που βρίσκονται πάντα κοντά μου, με στηρίζουν, με ενθαρρύνουν και μου δίνουν τη δύναμη να αγωνίζομαι.

Σε αυτούς αφιερώνω αυτήν την εργασία.

Στους γονείς μου

# Διαχείρηση Εικόνων Εγγράφων σε Πολυμεσικά Αρχειακά Συστήματα

Θεόδωρος Γκαραβέλος

Μεταπτυχιακή Εργασία

Τμήμα Επιστήμης Υπολογιστών

Πανεπιστήμιο Κρήτης

## ΠΕΡΙΛΗΨΗ

Η διαφύλαξη και μελέτη των ιστορικών στοιχείων και της πολιτιστικής κληρονομιάς αποτελούν πρώτιστο μέλημα, όχι μόνο των ειδικών, αλλά και του κοινωνικού συνόλου εν γένει. Ειδικότερα, ιστορικά στοιχεία υπό μορφή εγγράφων αποτελούν αδιάψευστα ιστορικά ντοκουμέντα και συνεπώς αναπόσπαστο στοιχείο και μαρτυρία της ιστορίας κοινωνικών ή και εθνικών ομάδων.

Η πιο γνωστή προσέγγιση στη διαφύλαξη των εγγράφων, ήταν -και συνεχίζει να είναι- η φύλαξη και η χρήση τους στους χώρους μιας βιβλιοθήκης. Η παλαιότητα των εγγράφων, η ανάγκη πρόσβασης σε αυτά από μεγάλο αριθμό ερευνητών (πολλές φορές και ταυτόχρονη), η ανάγκη της φυσικής ύπαρξης κάποιου στο χώρο της βιβλιοθήκης για τη μελέτη των εγγράφων, είναι μερικά από τα προβλήματα που καθιστούν την προσέγγιση αυτή μη αποδοτική και πολλές φορές δύσκολη ή και ακατόρθωτη.

Για τους λόγους αυτούς, και με δεδομένη την τρέχουσα τάση για ηλεκτρονική πρόσβαση σε πληροφορίες, τα τελευταία χρόνια άρχισε να υπάρχει μια στροφή προς τη χρήση ψηφιακών βιβλιοθηκών για τη παροχή πρόσβασης σε ιστορικά έγγραφα. Με τον τρόπο αυτό δίνεται λύση στο πρόβλημα της διαφύλαξης των εγγράφων αλλά και σε αυτό της εύκολης και γρήγορης ανάκτησης υλικού. Ακόμη περισσότερο δίνεται η δυνατότητα της ενσωμάτωσης επιπλέον υλικού (όπως περίληψης, μεταγραφής, μετάφρασης), κάτι που προσφέρει μεγαλύτερη ευχρηστία και καλύτερη εκμετάλλευση του υπάρχοντος υλικού.

Μια προσπάθεια ανάπτυξης μιας τετοιας ψηφιακής βιβλιοθήκης είναι και αυτή που εκπονείται στα πλαίσια του προγράμματος ΑΡΧΟΝ, με γνώμονα συγκεκριμένες

ανάγκες της Βικελαίας Δημοτικής Βιβλιοθήκης Ηρακλείου. Στόχοι του ΑΡΧΟΝ είναι η διαφύλαξη και η όσο το δυνατόν αρτιότερη εκμετάλλευση μερους του αρχειακού υλικού που υπάρχει στη Βικελαία Δημοτική Βιβλιοθήκη με τη μορφή εγγράφων.

Το σύστημα ΑΡΧΟΝ αποτελείται από τρία επιμέρους υποσυστήματα:

- υποσύστημα διαχείρισης εικόνας
- υποσύστημα βάσης δεδομένων
- υποσύστημα επικοινωνίας και διασύνδεσης

Η παρούσα εργασία έχει σαν στόχο την ανάπτυξη του υποσυστήματος διαχείρισης εικόνων του πολύμεσου συστήματος ψηφιακής βιβλιοθήκης ΑΡΧΟΝ. Συγκεκριμένα ασχολείται με τη μελέτη και ανάπτυξη όλων εκείνων των τμημάτων που χρειάζονται προκειμένου η ψηφιακή βιβλιοθήκη να αποκτήσει υλικό - σε μορφή εικόνων για τη δική μας περίπτωση - και να μπορεί κατόπιν να διαχειρίζεται το υλικό αυτό (με την εμφάνισή του στην οθόνη του χρήστη καθώς και με την παροχή διαφόρων εργαλείων). Για τον σκοπό αυτό έχουν αναπτυχθεί όλες οι διαδικασίες που είναι απαραίτητες για να φτάσουμε από το πρωτότυπο κείμενο σε μία εικόνα που θα έχει όσο το δυνατόν καλύτερη ποιότητα εμφάνισης και που θα είναι όσο γίνεται πιο κοντά στο πρωτότυπο. Δόθηκε μεγάλη έμφαση στους αλγόριθμους επεξεργασίας εικόνας που αναπτύχθηκαν ώστε να διασφαλιστεί η διατήρηση της πληροφορίας που εμπεριέχεται στα έγγραφα. Επίσης έγινε σύγκριση και έλεγχος των διαδικασιών συμπίεσης που υπάρχουν, ώστε να επιλεχθεί αυτή με τα καλύτερα συγκριτικά πλεονεκτήματα για τη δική μας εφαρμογή. Τέλος αναπτύχθηκε η επιφάνεια επαφής χρήστη που λειτουργεί σαν συνδετικός κρίκος ανάμεσα στα υποσυστήματα ανάκτησης των εγγράφων και σε αυτό της εμφάνισης των εικόνων στην οθόνη.

Το αποτέλεσμα αυτής της εργασίας είναι ένα ολοκληρωμένο υποσύστημα διαχείρισης εικόνων που, σε συνδυασμό με τα άλλα υποσυστήματα, αποτελούν ένα ενιαίο πολυμεσικό αρχειακό σύστημα. Το υποσύστημα διαχείρησης εικόνων έχει ήδη δοκιμαστεί σε συνθήκες εργασίας κοντά στις πραγματικές και τα αποτελέσματα είναι άκρως ενθαρρυντικά.

Επόπτης: Πάνος Τραχανιάς
Αναπληρωτής Καθηγητής Επιστήμης Υπολογιστών
Πανεπιστήμιο Κρήτης

# Document Image Management in Multimedia Archival Systems

Theodoros Garavelos

Master of Computer Science

University of Crete

## ABSTRACT

The preservation and study of historical records and of cultural inheritance are of utmost importance, not only to the experts, but also to the society. In particular, historical documents constitute indisputable historical evidence, and are therefore an integral part of societal and/or ethnic groups.

The most known approach in document preservation has been - and remains to be - their safekeeping and usage in the premises of a library. However, there are several drawbacks in it; the documents age, the need for access by a large number of researchers (many times simultaneously), and the need of physical presence at the library in order to study the documents, are some of the problems that make this approach ineffective, and in many cases difficult, or even impossible, to implement.

For these reasons, and considering the current trend for electronic access to information, we witness in recent years a gradual turn towards the adoption and usage of digital libraries for providing access to historical documents. Such an approach effectively addresses the problem of document preservation and also provides for efficient and fast retrieval mechanisms. Furthermore, we are given the means for integrating with the documents additional information (such as summary, comments, translation) which offers added value to the existing material.

In line with the above, the development of a digital library is pursued within the ARHON project; the requirements for the design and subsequent development of all relevant sub-systems have been set to meet the special needs of the Vikelaia Municipality Library of Heraklion. ARHON is aiming at the preservation and the best possible usage of archival material that exists in Vikelaia Municipality Library in the form of documents.

The ARHON system consists of three subsystems:

- image management subsystem

- data base subsystem
- communication and interconnection subsystem

In this work we have addressed the development of the first subsystem of ARHON, i.e. the image management subsystem. More specifically, we have studied, developed and tested all modules that are needed for document acquisition (as images) and storage in electronic form, and the subsequent management and usage of the document images (visualization, image manipulation and related services). To that end, we have developed image digitization procedures that are adapted to the documents at hand. Special emphasis has been given in the design of the image processing algorithms in order to preserve all useful information in the document images. An extensive comparison of existing compression methods has also been performed, in order to select and adopt the most appropriate for our purposes. Finally a user interface has been developed, acting as a link between the image retrieval subsystem and the image display module.

The result of this work is a complete image management subsystem that, together with the other subsystems, consist an integrated multimedia archival system. The image management subsystem has already been extensively tested in realistic conditions and the results are very encouraging.

Supervisor: Panos Trahanias
Associate Professor of Computer Science
University of Crete

# Περιεχόμενα

**Ευχαριστίες** **i**

**Περίληψη** **iv**

**Abstract** **vi**

**Περιεχόμενα** **1**

**Κατάλογος Σχημάτων** **3**

**1 Εισαγωγή** **5**

1.1 Γενική εισαγωγή . . . . . 5

1.2 Σκοπός της παρούσας εργασίας . . . . . 7

1.3 Προηγούμενες εργασίες . . . . . 9

1.3.1 Δομή εγγράφων . . . . . 10

1.3.1.1 Χρήση αρχείων κειμένου . . . . . 10

1.3.1.2 Χρήση εικόνων και αρχείων πληροφοριών . . . . . 11

1.3.1.3 Χρήση εικόνων και μεταγραφών . . . . . 12

1.3.1.4 Συνδυασμοί όλων των παραπάνω μεθόδων . . . . . 12

1.3.2 Απόκτηση και ποιότητα εικόνων . . . . . 13

1.3.3 Ασφάλεια εικόνων . . . . . 14

1.3.4 Χρήση πάνω από το διαδίκτυο . . . . . 14

1.4 Οργάνωση του υλικού . . . . . 15

**2 Σύστημα ΑΡΧΟΝ** **17**

2.1 Γενική περιγραφή . . . . . 17

2.2 Υποσύστημα εισαγωγής εγγράφων και επεξεργασίας εικόνας . . . . . 18

2.3 Υποσύστημα βάσης δεδομένων . . . . . 19

2.4 Υποσύστημα εμφάνισης εγγράφων και παροχής υπηρεσιών . . . . . 19

2.5 Υποσύστημα επικοινωνίας και διασύνδεσης . . . . . 20

2.6 Αναμενόμενα οφέλη . . . . . 20

**3 Διαχείριση Εικόνων Εγγράφων** **23**
3.1 Στάδιο εισαγωγής . . . 23
3.1.1 Σάρωση χειρογράφων . . . 23
3.1.2 Μετατροπή της εικόνας σε ασπρόμαυρη. . . . 28
3.1.2.1 Ενιαίο κατώφλι . . . 29
3.1.2.2 Προσαρμοζόμενο κατώφλι . . . 29
3.1.2.3 Μέθοδος ελαχίστου-μεγίστου . . . 32
3.1.2.4 Μέθοδος υποβιβασμού κατωφλίου . . . 34
3.1.3 Πειραματική αποτίμηση μεθόδου . . . 39
3.1.4 Καθαρισμός εικόνας . . . 40
3.1.5 Διαδικασία κλεισίματος . . . 45
3.1.6 Συμπίεση Εικόνων . . . 49
3.1.6.1 Συμπίεση με τον αλγόριθμο PackBits . . . 50
3.1.6.2 Συμπίεση LZW (Lempel, Ziv και Welch) . . . 50
3.1.6.3 Συμπίεση με τους αλγορίθμους CCITT Group3 και Group4 51
3.1.6.4 Συμπίεση με τον αλγόριθμο JBIG . . . 52
3.1.7 Αποφάσεις για τη δική μας εργασία . . . 53
3.2 Στάδιο ανάκτησης και εμφάνισης . . . 56
3.2.1 Γλώσσα προγραμματισμού Java . . . 57
3.2.2 Επικοινωνία χρήστη . . . 58
3.2.3 Εμφάνιση εγγράφων . . . 60
3.2.4 Σύνοψη λειτουργιών . . . 63

**4 Μελλοντικές κατευθύνσεις** **65**

**Βιβλιογραφία** **68**

# Κατάλογος Σχημάτων


2.1 Γενική δομή . . . . . 18

3.1 Γενική δομή . . . . . 24
3.2 Παραδείγματα εγγράφων κακής ποιότητας από α) πότισμα μελανιού, β) χρήση βουλοκεριού, γ) τρύπες στο χαρτί, δ) τσάκισμα του χαρτιού, ε) σφραγίδα και τσαλακωμένο χαρτί και στ) μουντζούρες στο χαρτί . . . . . 27
3.3 Αρχική εικόνα . . . . . 30
3.4 Αποτέλεσμα μεθόδου ενιαίου κατωφλίου . . . . . 31
3.5 Αποτέλεσμα μεθόδου προσαρμοζόμενου κατωφλίου . . . . . 33
3.6 Αποτέλεσμα μεθόδου ελαχίστου-μεγίστου . . . . . 35
3.7 Αποτέλεσμα μεθόδου υποβιβασμού κατωφλίου . . . . . 37
3.8 Ένα τμήμα εικόνας σε επίπεδα γκρίζου . . . . . 38
3.9 Αποτελέσματα στο τμήμα εικόνας των μεθόδων α) ενιαίου κατωφλίου, β) προσαρμοζόμενου κατωφλίου, γ) ελαχίστου-μεγίστου και δ) υποβιβασμού κατωφλίου . . . . . 38
3.10 Γραφική παράσταση της διαφοράς των δύο μεθόδων . . . . . 40
3.11 Εικόνα που προέρχεται από τη μέθοδο υποβιβασμού κατωφλίου . . . . . 41
3.12 Εικόνα που προέρχεται από τη μέθοδο ενιαίου κατωφλίου με διαφορά 6%, από αυτήν της μεθόδου υποβιβασμού κατωφλίου . . . . . 42
3.13 Επέκταση περιοχής . . . . . 44
3.14 Διαστάσεις λευκής ζώνης για την αφαίρεση μιας περιοχής . . . . . 45
3.15 Αποτέλεσμα της διαδικασίας καθαρισμού για την εικόνα του Σχ. 3.7 . . . 46
3.16 Αποτέλεσμα της διαδικασίας καθαρισμού για το τμήμα εικόνας του Σχ. 3.9δ . . . . . 47
3.17 Ανάγκη επαναληπτικότητας της διαδικασίας καθαρισμού . . . . . 47
3.18 Αποτέλεσμα της διαδικασίας κλεισίματος . . . . . 48
3.19 Παράδειγμα κωδικοποίησης για τον αλγόριθμο CCITT Group3 . . . . . 52
3.20 Σύγκριση αλγορίθμων κωδικοποίησης . . . . . 55
3.21 Επιδόσεις στη συμπίεση εικόνων, στις οποίες δεν έχει προηγηθεί η διαδικασία καθαρισμού . . . . . 56
3.22 Μοντέλο επικοινωνίας του συστήματος . . . . . 57
3.23 Εμφάνιση αποτελεσμάτων επερώτησης στη βάση . . . . . 59

3.24 Η αρχική εικόνα ενός εγγράφου και μια μεγέθυνση περιοχής . . . . . 61
3.25 Αλλαγή στη γεωμετρία του παραθύρου εμφάνισης . . . . . 62
3.26 Εφαρμογή των φίλτρων διόρθωσης φωτεινότητας και διόρθωσης γάμμα αντίστοιχα . . . . . 63
3.27 Εκτύπωση μέσα από το πρόγραμμα εμφάνισης . . . . . 64

# Κεφάλαιο 1

# Εισαγωγή

## 1.1 Γενική εισαγωγή

Η πολιτιστική κληρονομιά και η προσπάθεια του ανθρώπου να την διασφαλίσει και να την προάγει είναι δύο από τα χαρακτηριστικά που ξεχωρίζουν τον άνθρωπο από το ζωικό βασίλειο. Ίχνη αυτής της προσπάθειας, που συνεχίζεται μέχρι και σήμερα με διάφορους τρόπους, έχουμε από την αρχή της ύπαρξης του ανθρώπινου είδους. Οι τοιχογραφίες σε σπηλιές, τα διάφορα αγαλματίδια, οι μνημειακές κατασκευές και οι προφορικές διαδόσεις μύθων και παραδόσεων είναι μερικές εκφράσεις αυτής της προσπάθειας.

Με την εμφάνιση και τη διάδοση του γραπτού λόγου αρχίζουμε να έχουμε και τα πρώτα γραπτά κείμενα. Κάποια από αυτά θα διασωθούν και θα περάσουν σαν πολύτιμος θησαυρός στα χέρια μελλοντικών γενιών.

Το πόσο πολύτιμος είναι αυτός ο θησαυρός δεν μπορούμε να το εκτιμήσουμε άμεσα. Αρκεί όμως να αναφέρουμε ότι μέσα από τα γραπτά αυτά κείμενα μπορούμε να αναπλάσουμε μια άλλη εποχή, μια εποχή με διαφορετικά ήθη και έθιμα, με διαφορετική νοοτροπία και ιδέες. Μπορούμε να σχηματίσουμε στο μυαλό μας μια εικόνα των επιδιώξεων των ανθρώπων εκείνων και των σκοπών τους. Των συνθηκών που επικρατούσαν σε μια κοινωνία ή ένα κράτος και το οδήγησαν να πάρει την όποια απόφαση σε κάποια δεδομένη χρονική στιγμή. Πολλά από αυτά τα έγγραφα είναι ανεκτίμητης ιστορικής αξίας. Περιέχουν πληροφορίες που κάποιος ιστορικός θα μπορούσε να χρησιμοποιήσει για να επιλύσει διάφορα εθνικά θέματα που προκύπτουν. Δεν είναι λίγες εξάλλου οι περιπτώσεις που χρειάζεται αναδρομή σε κάποια τέτοιου είδους έγγραφα για να λυθούν εδαφικές διεκδικήσεις, αξιώσεις ή ισχυρισμοί. Το να είναι κάποιος σε θέση χρησιμοποιώντας τέτοια αδιάψευστα ντοκουμέντα να αποδείξει το δίκιο των ισχυρισμών του είναι πολύ σημαντικό και επιθυμητό. Η εθνική λοιπόν και ιστορική σημασία των εγγράφων μπορεί να είναι πολύ μεγάλη. Μεγάλη όμως αξία μπορούν να έχουν και για θρησκευτικούς λόγους. Σε παλαιότερες κοινωνίες και περιόδους η θρησκεία έπαιζε έναν ιδιαίτερα σημαντικό ρόλο. Το να μπορεί να αποφανθεί κανείς για τη θρησκευτική κατάσταση που επικρατούσε σε κάποιο μέρος, κάποια χρονική περίοδο,

σημαίνει πως μπορεί να εξάγει πολλές και χρήσιμες πληροφορίες όπως για το ποιοι λαοί αποτελούσαν εχθρικά κράτη (που αρκετές φορές είναι και αλλόθρησκα), ποιοι ήταν οι σύμμαχοί τους, πως επηρεάζονταν από τις τυχόν γενικότερες εξελίξεις που τους αφορούσαν έμμεσα, εξελίξεις που πολύ συχνά είχαν θρησκευτική βάση (αποσχίσεις, εμφάνιση νέων θρησκευτικών ρευμάτων και ιδεών κ.α. ). Πολύ σημαντικές πληροφορίες μπορεί κάποιος να εξάγει και για τη κοινωνική κατάσταση που επικρατούσε κάποια περίοδο. Τα έγγραφα μπορεί να αναφέρονται σε κάποιες πτυχές της καθημερινής ζωής (π.χ. δωρεές, γάμους και συνοικέσια, κηδείες κ.α. ), και επομένως μας βοηθούν να αντλήσουμε πληροφορίες για την οργάνωση και τη δομή της τότε κοινωνίας.

Από όλες αυτές τις πληροφορίες που μπορεί να αντλήσει κανείς, μπορεί να σχηματίσει μια ολοκληρωμένη εικόνα για την κοινωνία στην οποία αναφέρονται τα έγγραφα. Σε συνδυασμό με κάποιες άλλες ιστορικές πηγές θα μπορούσε να αποκτήσει μια ακόμη πιο πλήρη εικόνα και να κατανοήσει τις εξελίξεις της κοινωνίας, τις αντιδράσεις και τις επιδιώξεις της, γνωρίζοντας πολλά πράγματα από τον τρόπο σκέψης και τις συνθήκες που επικρατούσαν την περίοδο εκείνη.

Ένας ακόμη λόγος που τα γραπτά έγγραφα έχουν τόση μεγάλη σημασία είναι το πλεονέκτημα του μη-μεταβλητού περιεχομένου. Δεν βασίζεται, δηλαδή, το περιεχόμενό τους στη μνήμη κάποιου ούτε στη θέση την οποία παίρνει απέναντι σε κάποιο γεγονός. Το περιεχόμενό τους είναι αμετάβλητο και περνά σε κάθε μελετητή το ίδιο. Το πώς θα το χρησιμοποιήσει ο κάθε μελετητής και το πώς θα καταφέρει να εξάγει όσο πιο πολλές, ακριβείς και αληθείς πληροφορίες εναπόκειται στις γνώσεις και τις ικανότητες του κάθε μελετητή.

Καταλαβαίνουμε λοιπόν από τα παραπάνω πόσο τεράστια είναι η σημασία της διαφύλαξης αυτών των κειμένων. Ένα πρώτο βήμα είναι η φύλαξή τους στο χώρο μιας βιβλιοθήκης. Αλλά αυτό από μόνο του δεν είναι αρκετό μερικές φορές. Η παλαιότητα των ίδιων των εγγράφων, η έλλειψη πόρων και μέσων για τη σωστή διαφύλαξη των εγγράφων από μέρους της βιβλιοθήκης και η ανάγκη για πρόσβαση σε αυτά (ίσως και ταυτόχρονα) από μεγάλο αριθμό μελετητών μας οδηγούν στη διαπίστωση πως το κλασικό μοντέλο της βιβλιοθήκης δεν αρκεί.

Έτσι τα τελευταία χρόνια πολλές βιβλιοθήκες άρχισαν να κατευθύνονται προς την έννοια της ψηφιακής βιβλιοθήκης [9], μιας βιβλιοθήκης δηλαδή που το υλικό της υπάρχει και σε ψηφιακή μορφή. Με τον τρόπο αυτό δίνεται βιώσιμη λύση στο πρόβλημα της διατήρησης/διάσωσης των εγγράφων αλλά και εύκολη πρόσβαση στο υλικό. Επίσης δίνεται η δυνατότητα του μετά την ψηφιοποίηση σχολιασμού των εγγράφων. Σημαντικά στοιχεία για την όσο το δυνατόν αρτιότερη εκμετάλλευση του συνόλου του αρχειακού υλικού όπως περίληψη, μεταγραφή, μετάφραση, ο χώρος και ο χρόνος σύνταξης, ο συντάκτης κ.α. είναι δυνατόν να ενσωματωθούν σταδιακά. Ακόμη περισσότερο οι ψηφιακές βιβλιοθήκες επιτρέπουν γρήγορη και εύκολη ανάκτηση των εγγράφων. Για την επίτευξη του σκοπού αυτού κάνουν χρήση διαφόρων μεθόδων όπως

- ανάκτηση με τη χρήση λέξεων αναζήτησης (search words). Οι λέξεις αναζήτησης αναζητούνται μέσα στο κείμενο του εγγράφου και μπορούν να είναι οποιοσδήποτε

όρος μας ενδιαφέρει. Με την εφαρμογή πιο σύνθετων κανόνων ψαξίματος, όπως η χρήση του ΚΑΙ για την ταυτόχρονη αναζήτηση πάνω από ένα όρο (πχ. Επανάσταση ΚΑΙ 1821), του ΚΑΙ ΟΧΙ για την αναζήτηση ενός όρου και τον αποκλεισμό κάποιου άλλου (πχ. δωρεά ΚΑΙ ΟΧΙ χρήματα), και αρκετών ακόμη κανόνων, μπορούμε να αυξήσουμε τη χρηστικότητα της μεθόδου αναζήτησης. Η χρήση έξυπνων και γρήγορων μεθόδων αναζήτησης είναι από τα πολύ σημαντικά ζητήματα που αυξάνουν την απόδοση τέτοιων τρόπων αναζήτησης.

- χρήση σημασιολογικής (semantic) πληροφορίας που επιτρέπει συσχετισμό πολλών εγγράφων μεταξύ τους με βάση έναν ή περισσότερους όρους. Οι όροι μπορούν να αφορούν το θέμα προς ανάκτηση, το περιεχόμενο του εγγράφου ή την κατηγορία στην οποία έχει καταταγεί/ταξινομηθεί το ίδιο το έγγραφο. Πολλές φορές επίσης μπορούν να χρησιμοποιούνται όροι που αφορούν άλλα χαρακτηριστικά του εγγράφου όπως το όνομα του συγγραφέα και οποιεσδήποτε άλλες πληροφορίες υπάρχουν και πλαισιώνουν το έγγραφο.

- χρήση έξυπνων προγραμμάτων αναζήτησης σε βάσεις που περιέχουν πληροφορίες με την μορφή εικόνων [29]. Η αναζήτηση εδώ γίνεται με βάση μια περιγραφή του περιεχομένου των εικόνων που επιθυμούμε να βρούμε. Αφού το πρόγραμμα αναζήτησης συγκρίνει την περιγραφή που του έχουμε δώσει με τις εικόνες της βάσης, μας επιστρέφει πίσω αυτές που μοιάζουν περισσότερο με την περιγραφή μας. Φυσικά στη περίπτωση που η βάση μας αποτελείται από έγγραφα τα οποία έχουν παρόμοια μορφή, η αναζήτηση είναι αρκετά πιο δύσκολη. Παραμένει όμως μια δυνατότητα που απλοποιεί αρκετά την ανεύρεση των επιθυμητών εγγράφων, ιδίως αν σκεφτεί κανείς πως χωρίς την χρήση υπολογιστών θα έπρεπε να ανατρέξει σε ένα αρκετά μεγάλο όγκο εγγράφων για να εντοπίσει τα ζητούμενα.

Ακόμη πιο πρόσφατα οι νέες τεχνολογίες και η ανάπτυξη του Internet κάνουν το σωστό σχεδιασμό συστημάτων αρχειοθέτησης μια πραγματική πρόκληση. Θέματα όπως η αρχιτεκτονική ενός τέτοιου συστήματος, η οργάνωση του υλικού αλλά και η απόκτηση του σε όσο το δυνατόν καλύτερη ποιότητα είναι από τα πιο ουσιαστικά και η απάντησή τους παίζει τον πρώτο ρόλο στη δύναμη του συστήματος.

## 1.2 Σκοπός της παρούσας εργασίας

Η παρούσα εργασία εκπονήθηκε στα πλαίσια του προγράμματος **ΑΡΧΟΝ** (πολύμεσο σύστημα **ΑΡ**χειοθέτησης σ**ΧΟ**λιασμού και α**Ν**άκτησης ιστορικών εγγράφων [17]), το οποίο στοχεύει στην ανάπτυξη ενός πολύμεσου συστήματος ψηφιακής βιβλιοθήκης. Το σύστημα πρόκειται να χρησιμοποιηθεί στη Βικελαία Δημοτική Βιβλιοθήκη του Δήμου Ηρακλείου. Ανάμεσα στους κυριότερους στόχους του είναι η διατήρηση και η παροχή εύκολης πρόσβασης στα ιστορικά έγγραφα της βιβλιοθήκης, η δυνατότητα σχολιασμού των εγγράφων με επιπρόσθετες πληροφορίες, η ανάκτηση και συσχέτιση

των εγγράφων με χρήση σημασιολογικής πληροφορίας. Η ευκολία χρήσης αναφέρεται τόσο στην εύκολη αναζήτηση και εύρεση ενός εγγράφου όσο και στις δυνατότητες λειτουργίας πάνω από ένα δίκτυο που θα παρέχει το σύστημα. Το τελευταίο είναι αρκετά σημαντικό αφού επιτρέπει τη χρήση της βιβλιοθήκης από απόσταση καθώς και λειτουργίες όπως ηλεκτρονική ανταλλαγή δεδομένων ανάμεσα σε βιβλιοθήκες.

Τα έγγραφα που θα καταχωρηθούν αποτελούν αντικείμενα μεγάλης πολιτισμικής και ιστορικής αξίας και η κατασκευή του εν λόγω συστήματος θα επιτρέψει την εύκολη χρήση και μελέτη τους χωρίς να εμπλέκονται τα πρωτότυπα κείμενα. Με την αντιγραφή των εγγράφων και την καταχώρησή τους στο σύστημα παύει να υφίσταται ο κίνδυνος καταστροφής των πρωτότυπων κειμένων. Δεδομένου δε ότι ο αριθμός των κειμένων αυτών είναι πολύ μεγάλος, καθίσταται φανερό ότι ένα αυτοματοποιημένο σύστημα για τη διαχείρισή τους είναι απαραίτητο.

Η προσπάθεια αυτή είναι πολύ σημαντική για τη διατήρηση, διαφύλαξη και μελέτη της πολιτισμικής κληρονομιάς. Σημειώνεται ότι τα προαναφερθέντα αρχεία περιέχουν ιστορικούς θησαυρούς μεγάλης αξίας για τις Κρητολογικές μελέτες. Τα αρχεία στη παρούσα τους μορφή (πρωτότυπα) είναι δύσχρηστα και πολύ επιρρεπή σε φθορές. Επιπλέον, είναι πολύ δύσκολο να διασφαλιστεί το περιεχόμενό τους εφ'όσον αυτό δεν υπάρχει φυλαγμένο σε καμία άλλη μορφή.

Το σύστημα ΑΡΧΟΝ αποτελείται από τρία κυρίως υποσυστήματα:

**υποσύστημα διαχείρισης εικόνας** Περιλαμβάνει όλα τα απαραίτητα εκείνα τμήματα που έχουν σχέση με την βελτίωση της ποιότητας των εικόνων, με την συμπίεσή/αποσυμπίεσή τους καθώς και το υποσύστημα επαφής χρήστη. Γενικότερα μπορούμε να πούμε πως εδώ εντάσσονται όλες οι λειτουργίες που αφορούν την απόκτηση υλικού και την διαχείρισή του.

**υποσύστημα βάσης δεδομένων** Βασίζεται στο Σύστημα Σημασιολογικού Ευρετηριασμού [37] που έχει αναπτυχθεί στο Ινστιτούτο Τεχνολογίας και Έρευνας (ΙΤΕ), ένα σύστημα που επιτρέπει την οργάνωση, αποτίμηση και αρχειοθέτηση του υλικού και κατόπιν την εύκολη ανεύρεση και χρήση του.

**υποσύστημα επικοινωνίας και διασύνδεσης** Αφορά τα τμήματα εκείνα που χρειάζονται προκειμένου να μπορούν να επικοινωνούν όλα τα υποσυστήματα μεταξύ τους. Περιέχει επίσης τα υποσυστήματα εκείνα που υποστηρίζουν τη δικτυακή λειτουργία του όλου συστήματος.

Η παρούσα εργασία έχει σαν στόχο την ανάπτυξη του υποσυστήματος διαχείρισης εικόνων του πολύμεσου συστήματος ψηφιακής βιβλιοθήκης ΑΡΧΟΝ. Συγκεκριμένα ασχολείται με την ανάπτυξη και τη μελέτη όλων εκείνων των τμημάτων που χρειάζονται προκειμένου η ψηφιακή βιβλιοθήκη να αποκτήσει υλικό - σε μορφή εικόνων για τη δική μας περίπτωση - και να μπορεί κατόπιν να διαχειρίζεται το υλικό αυτό (με την εμφάνισή του στην οθόνη του χρήστη καθώς και με την παροχή διαφόρων εργαλείων). Για τον σκοπό αυτό έχουν αναπτυχθεί όλες οι διαδικασίες που είναι απαραίτητες για

να φτάσουμε από το πρωτότυπο κείμενο σε μία εικόνα που θα έχει όσο το δυνατόν καλύτερη ποιότητα εμφάνισης και που θα είναι όσο γίνεται πιο κοντά στο πρωτότυπο. Επίσης έγινε σύγκριση και έλεγχος των διαδικασιών συμπίεσης που υπάρχουν, ώστε να επιλεχθεί αυτή με τα καλύτερα συγκριτικά πλεονεκτήματα για τη δική μας εφαρμογή. Τέλος αναπτύχθηκε η επιφάνεια επαφής χρήστη που λειτουργεί σαν συνδετικός κρίκος ανάμεσα στα υποσυστήματα ανάκτησης των εγγράφων και σε αυτό της εμφάνισης των εικόνων στην οθόνη.

Πολλές είναι οι κρίσιμες εκείνες αποφάσεις που οδήγησαν στην επιλογή μεθόδων, εργαλείων και τεχνολογίας προκειμένου να έχουμε ένα σύστημα ικανό να αποκριθεί με επιτυχία τόσο στις απαιτήσεις των χρηστών όσο και στα προβλήματα που το αντικείμενο εμπεριέχει από την φύση του.

Ο μεγάλος όγκος των δεδομένων προς ψηφιοποίηση, που συνήθως υπάρχει σε τέτοιες περιπτώσεις, μας οδήγησε σε μεθόδους με τις οποίες πετύχαμε μείωση του όγκου αυτού. Αλγόριθμοι επεξεργασίας εικόνων χρησιμοποιήθηκαν ώστε να εξασφαλίσουμε όσο γίνεται περισσότερο την ποιότητα των εικόνων. Έχοντας κατά νου τη γοργή εξάπλωση του Internet και την τάση για διασύνδεση όλων των πηγών πληροφοριών στον κορμό του, έγινε επιλογή εργαλείων ανάπτυξης (γλώσσα προγραμματισμού Java) που εγγενώς υποστηρίζουν τη δικτυακή λειτουργία και επικοινωνία. Τέλος σκεπτόμενοι πως οι τελικοί χρήστες θα είναι, ως επί το πλείστον, εκτός του χώρου της τεχνολογίας υπολογιστών (ιστορικοί μελετητές, κοινωνιολόγοι, οικονομολόγοι, γλωσσολόγοι, σπουδαστές κ.α.) έγινε προσπάθεια ώστε να αυτοματοποιηθούν, στο μεγαλύτερο βαθμό που αυτό είναι εφικτό, όλες οι διαδικασίες καθώς και να αναπτυχθεί ένα περιβάλλον επικοινωνίας απλό και κατανοητό από όλους.

Το αποτέλεσμα αυτής της εργασίας είναι το ολοκληρωμένο υποσύστημα διαχείρισης εικόνων, το οποίο μαζί με το Σύστημα Σημασιολογικού Ευρετηριασμού καθώς και το υποσύστημα επικοινωνίας και διασύνδεσης θα αποτελέσει ένα τελικό προϊόν το οποίο και θα λειτουργήσει στη Βικελαία Δημοτική Βιβλιοθήκη του Ηρακλείου. Το υποσύστημα διαχείρισης εικόνων έχει ήδη δοκιμαστεί σε συνθήκες εργασίας κοντά στις πραγματικές και τα αποτελέσματα είναι άκρως ενθαρρυντικά.

## 1.3 Προηγούμενες εργασίες

Είναι φυσικό για ένα τόσο σημαντικό πρόβλημα όπως είναι αυτό της διαφύλαξης των εγγράφων μιας βιβλιοθήκης να έχουν γίνει αρκετές ερευνητικές και αναπτυξιακές προσπάθειες. Από τα προηγούμενα έχει γίνει ίσως φανερό, το πλήθος των προβλημάτων που αντιμετωπίζει κάποιος που προσπαθεί να υλοποιήσει ένα σύστημα ψηφιακής βιβλιοθήκης. Ο τρόπος με τον οποίο μπορεί κάποιος να συλλέξει το υλικό για την βιβλιοθήκη του έχει άμεσο αντίκτυπο στην ποιότητα των υπηρεσιών που μπορεί να παρέχει. Κακής ποιότητας εικόνες αποδυναμώνουν τη ψηφιακή βιβλιοθήκη, δεν προσφέρουν κανένα συγκριτικό πλεονέκτημα ούτε διορθώνουν τις αδυναμίες της κλασικής βιβλιοθήκης. Ένα υποσύστημα αρχειοθέτησης και ανάκτησης θα πρέπει να

είναι ικανό να χειρίζεται μεγάλους όγκους δεδομένων χωρίς προβλήματα απόδοσης ή ταχύτητας και να εξασφαλίζει την απρόσκοπτη καλή λειτουργία του όλου συστήματος ανεξάρτητα από τις ποικίλες ενέργειες των χρηστών (αναζητήσεις, καταχωρήσεις, διαγραφές, προσθήκες, μετατροπές κα.). Η ανάκτηση θα πρέπει να γίνεται με όσο το δυνατό πιο έξυπνο και αποδοτικό τρόπο ώστε να μπορεί κάποιος να εντοπίζει τα κείμενα εκείνα, και μόνο, που τον ενδιαφέρουν και τον αφορούν. Μια κακή επιλογή τέτοιου υποσυστήματος ή η επιλογή ενός αδύναμου υποσυστήματος μπορεί να προκαλέσει τη μείωση της απόδοσης όλης της ψηφιακής βιβλιοθήκης. Η διεπιφάνεια εργασίας του χρήστη θα πρέπει να προσφέρει όλες εκείνες τις λειτουργίες που είναι απαραίτητες για την εύκολη και ολοκληρωμένη διαχείριση των εικόνων στην οθόνη. Θα πρέπει επίσης να είναι κατανοητά από ανθρώπους με μικρή εμπειρία και επαφή με υπολογιστές. Το υποσύστημα επικοινωνίας και διασύνδεσης πρέπει να λειτουργεί με αδιαφάνεια ως προς τον χρήστη ώστε να μην χρειάζεται να γνωρίζει αυτός λεπτομέρειες που δεν τον ενδιαφέρουν, αλλά ταυτόχρονα θα πρέπει να προσφέρει υψηλής ποιότητας επικοινωνία ώστε να έχουμε μεγάλη απόδοση του συστήματος.

Ακριβώς επειδή υπάρχουν τόσα πολλά προβλήματα, οι απαντήσεις που μπορεί να δώσει κάποιος είναι εξίσου πολυάριθμες. Ο υπεύθυνος σχεδιασμού και ανάπτυξης του συστήματος, σε στενή συνεργασία με τους φορείς για τους οποίους θα γίνει η ανάπτυξη, θα πρέπει προσεκτικά να συγκρίνει και να αξιολογήσει τα υπέρ και τα κατά της κάθε προσέγγισης και κατόπιν να προτείνει όποια θεωρεί σαν την καλύτερη λύση. Ας εξετάσουμε κάποιες από τις λύσεις που έχουν ήδη δοκιμαστεί σε διάφορα συστήματα ψηφιακής βιβλιοθήκης.

### 1.3.1 Δομή εγγράφων

Ένα από τα πιο σημαντικά προβλήματα είναι αυτό της οργάνωσης της πληροφορίας που περιέχει η βάση δεδομένων μας. Η οργάνωση των δεδομένων καθορίζει το είδος και την ποιότητα υπηρεσιών που μπορούμε να παρέχουμε. Έχουν προταθεί και χρησιμοποιηθεί αρκετές μέθοδοι [10] που συνοψίζονται στις επόμενες παραγράφους.

#### 1.3.1.1 Χρήση αρχείων κειμένου

Η πιο απλή περίπτωση είναι αυτή κατά την οποία μια βιβλιοθήκη επιλέγει να οργανωθεί γύρω από αρχεία κειμένου μόνο. Από πλευράς τεχνολογίας μια τέτοια επιλογή έχει τις λιγότερες δυνατές απαιτήσεις, και στην περίπτωση που η απόκτηση υλικού γίνεται μέσω δακτυλογράφησης κειμένων η χρήση της τεχνολογίας περιορίζεται στο μέγιστο βαθμό.

Τα προβλήματα που αντιμετωπίζει η μέθοδος αυτή είναι αρκετά και μερικά πολύ σημαντικά. Ένα από τα πιο σοβαρά είναι και αυτή της αργής μετατροπής του αρχείου σε ψηφιακή μορφή. Από τη στιγμή που βασιζόμαστε (στη περίπτωση αυτή) σχεδόν αποκλειστικά στην ανθρώπινη εργασία, το μοντέλο αυτό οργάνωσης έχει εξαιρετικά αργούς ρυθμούς δημιουργίας της ψηφιακής βιβλιοθήκης και συνεπακόλουθα μεγάλο

χρηματικό κόστος. Ένα άλλο πρόβλημα είναι η αδυναμία υποστήριξης εγγράφων τα οποία περιέχουν σχέδια ή εικόνες. Σε μερικές περιπτώσεις- είδη εγγράφων αυτό δεν είναι ιδιαίτερα κακό. Στη γενική όμως περίπτωση δημιουργεί σημαντικά προβλήματα αφού πολλές φορές τα σχήματα και οι εικόνες αποτελούν βοηθήματα για την καλύτερη κατανόηση του κειμένου, ενώ άλλες φορές περιέχουν τα ίδια την μεγαλύτερη ποσότητα πληροφορίας. Τέλος ένα ακόμη πρόβλημα είναι ότι με την δημιουργία αρχείων κειμένου έχουμε απώλεια της πραγματικής μορφής του εγγράφου. Χάνουμε την ιδιαίτερη αίσθηση που μας δίνει το έγγραφο όταν το βλέπουμε, χάνουμε τις πληροφορίες που μπορούμε να έχουμε από το γραφικό χαρακτήρα του συγγραφέα, χάνουμε την αυθεντική μορφή του κειμένου (όταν π.χ. υπάρχουν σημειώσεις γραμμένες στα πλάγια ή όταν γίνεται χρήση διαφορετικών μεγεθών γραμμάτων, μελανιών κ.α.), απώλειες οι οποίες μπορούν να οδηγήσουν ακόμη και σε λανθασμένη εξαγωγή πληροφοριών από το έγγραφο.

Δεν πρέπει βέβαια να αποσιωπήσουμε τα πλεονεκτήματα μιας τέτοιας μορφής οργάνωσης. Και αυτά είναι το μικρότερο μέγεθος που απαιτείται για την οργάνωση της βιβλιοθήκης καθώς και το γεγονός ότι με τον τρόπο αυτό υπάρχει η δυνατότητα για αναζήτηση πληροφοριών σε ολόκληρο το κείμενο, κάτι που αυξάνει πολύ την χρηστικότητα ενός τέτοιου συστήματος.

Αυτή η μέθοδος οργάνωσης είναι η πρώτη που εμφανίστηκε και εξακολουθεί να είναι ευρέως διαδεδομένη. Ένας τεράστιος όγκος πληροφοριών είναι διαθέσιμος στη μορφή αυτή και πάρα πολλές ψηφιακές βιβλιοθήκες εξακολουθούν να είναι οργανωμένες με τον τρόπο αυτό [28], [24], [18].

Η χρήση οπτικής αναγνώρισης χαρακτήρων (OCR) είναι ένας τρόπος να βελτιωθεί και να αυτοματοποιηθεί η μετατροπή του υλικού σε ψηφιακή μορφή. Τα προβλήματα της οπτικής αναγνώρισης χαρακτήρων είναι πως ακόμη και σήμερα εξακολουθεί να είναι μη ακριβής έχοντας μεγάλα ποσοστά σφάλματος ενώ είναι πρακτικά αδύνατη η χρήση της σε χειρόγραφα λόγω των εξαιρετικά ιδιόμορφων γραφικών χαρακτήρων που μπορεί κάποιος να συναντήσει.

#### 1.3.1.2 Χρήση εικόνων και αρχείων πληροφοριών

Μια άλλη μέθοδος οργάνωσης είναι αυτή με τη χρήση εικόνων και αρχείων πληροφοριών [21], [25]. Εδώ τα έγγραφα υπάρχουν στη ψηφιακή βιβλιοθήκη με την μορφή εικόνων. Με τον τρόπο αυτό όλα τα προβλήματα που σχετίζονται με την αργή ψηφιοποίηση των εγγράφων, την αδυναμία υποστήριξης εικόνων και σχεδίων και την απώλεια της πραγματικής μορφής των εγγράφων, που αναφέραμε παραπάνω δεν υπάρχουν. Όλες οι πληροφορίες που μπορεί να πάρει κανείς από το έγγραφο εξακολουθούν να υπάρχουν και τώρα. Απαραίτητη όμως προϋπόθεση είναι να γίνει η απόκτηση των εικόνων με μεγάλη προσοχή ώστε να διασφαλιστεί η όσο το δυνατόν καλύτερη ποιότητα τους. Αυτό επιβραδύνει τη διαδικασία αφού απαιτούνται συνεχείς ποιοτικοί έλεγχοι, αλλά εξακολουθεί να παραμένει αρκετά πιο γρήγορη από την μεταγραφή σε

αρχεία κειμένου.

Τα αρχεία πληροφοριών είναι εξωτερικά των εικόνων αρχεία, τα οποία παρέχει ο υπεύθυνος δημιουργίας του ψηφιακού αρχείου (στις πιο πολλές περιπτώσεις η ίδια η βιβλιοθήκη) προκειμένου να κάνει τη χρήση των εγγράφων πιο εύκολη και το περιεχόμενό τους πιο κατανοητό. Μπορούν να περιέχουν διαφόρων ειδών πληροφορίες (βιογραφικά στοιχεία του συγγραφέα, ιστορικές πληροφορίες της περιόδου στην οποία αναφέρεται το έγγραφο, αναφορές σε άλλες πηγές πληροφοριών κ.α. ). Με τον τρόπο αυτό η χρήση των εγγράφων καθώς και η αναζήτηση πληροφοριών γίνεται πιο εύκολη αφού μπορεί κάποιος να αντλεί πληροφορίες που να σχετίζονται με κάποιο θέμα πολύ απλά, χωρίς να είναι αναγκασμένος να τις ψάχνει ο ίδιος ελέγχοντας μεγάλο πλήθος εικόνων.

#### 1.3.1.3 Χρήση εικόνων και μεταγραφών

Προσφέροντας ένα επίπεδο καλύτερων υπηρεσιών μια ψηφιακή βιβλιοθήκη μπορεί να οργανωθεί χρησιμοποιώντας και εικόνες και μεταγραφές των εγγράφων [35], [22]. Εδώ οι μεταγραφές παρέχονται αφ'ενός για να διευκολύνουν την ανάγνωση (ιδίως στις περιπτώσεις που τα πρωτότυπα είναι σε κακή κατάσταση ή με δυσνόητο γραφικό χαρακτήρα), και αφ'ετέρου για να επιτρέπουν το ψάξιμο λέξεων-κλειδιών σε ολόκληρο το κείμενο.

Επειδή οι μεταγραφές μπορούν να γίνουν μόνο από ειδικευμένους ερευνητές και απαιτούν προσεκτικούς ελέγχους ώστε να διασφαλιστεί η σωστή μεταγραφή των κειμένων, είναι φανερό πως η ταυτόχρονη δημιουργία εικόνων και μεταγραφών δεν είναι δυνατή. Χρειαζόμαστε λοιπόν μια διαδικασία κατά την οποία να μπορούμε εκ των υστέρων να προσθέτουμε την πληροφορία που έχουμε διαθέσιμη πια. Η δυνατότητα αυτή αυξάνει τη πολυπλοκότητα ενός τέτοιου συστήματος αλλά και την δύναμή του. Στη δική μας προσέγγιση έχει ενσωματωθεί μια τέτοια δυνατότητα.

Η μέθοδος εξακολουθεί να εμπεριέχει το πρόβλημα της αργής ψηφιοποίησης του αρχείου αλλά προσφέρει πολύ μεγαλύτερη λειτουργικότητα από τις προηγούμενες μεθόδους.

#### 1.3.1.4 Συνδυασμοί όλων των παραπάνω μεθόδων

Τα τελευταία χρόνια γίνεται μια προσπάθεια από όλο και μεγαλύτερο αριθμό φορέων (και η δική μας προσέγγιση εντάσσεται στη προσπάθεια αυτή) για να αξιοποιηθούν όλες οι δυνατότητες που παρέχει η νέα τεχνολογία στον τομέα των ψηφιακών βιβλιοθηκών. Για τον σκοπό αυτό, όπου είναι δυνατό, γίνεται χρήση όλων των επιμέρους συστατικών οργάνωσης που αναφέρθηκαν παραπάνω [36], [1]. Το αποτέλεσμα είναι μια ψηφιακή βιβλιοθήκη με ακόμη μεγαλύτερες δυνατότητες. Η χρήση εικόνων επιτρέπει την ύπαρξη όλης της πληροφορίας που μπορεί να αντλήσει κανείς από το πρωτότυπο έγγραφο. Οι μεταγραφές/μεταφράσεις συμπληρώνουν την εικόνα δίνοντας δυνατότητες αναζήτησης πάνω από αυτές που μπορεί να παρέχει μια κλασική βιβλιοθήκη, αυξάνοντας ακόμη

περισσότερο την ευχρηστία. Η ύπαρξη αρχείων πληροφοριών προσφέρει επιπλέον στην κατανόηση του περιεχομένου και εμπλουτίζει με πληροφορίες το κάθε έγγραφο, επιτρέποντας ομαδοποιήσεις και συσχετισμούς μεταξύ των εγγράφων.

Η χρήση των ψηφιακών βιβλιοθηκών μέσα από το διαδίκτυο ήταν μια αναμενόμενη εξέλιξη μετά την ραγδαία εξάπλωσή του και την απαίτηση μεγάλης μερίδας κόσμου για παροχή πληροφοριών μέσα και από αυτό το φορέα. Πηγαίνοντας ένα βήμα ακόμη πιο πέρα η δική μας προσέγγιση προσφέρει τη δυνατότητα τόσο για την αναζήτηση των εγγράφων μέσω επερωτήσεων στη βάση δεδομένων, όσο και την προσθήκη νέων πληροφοριών με τη μορφή μεταγραφών και μεταφράσεων και νέων αρχείων πληροφοριών πάνω από το διαδίκτυο.

### 1.3.2 Απόκτηση και ποιότητα εικόνων

Δεχόμενοι ότι η ύπαρξη των πρωτότυπων εγγράφων στη ψηφιακή βιβλιοθήκη υπό μορφή εικόνων είναι άκρως επιθυμητή, η απόκτηση των εικόνων είναι μία άλλη σημαντική παράμετρος στο συνολικό εγχείρημα της ανάπτυξης της ψηφιακής βιβλιοθήκης. Επειδή οι εικόνες των εγγράφων είναι το κομμάτι εκείνο με το οποίο έρχονται οι χρήστες σε άμεση επαφή, η ποιότητά τους καθορίζει σε μεγάλο βαθμό την επιτυχία ή όχι του συστήματος. Καλής ποιότητας εικόνες σημαίνει πως η ψηφιακή βιβλιοθήκη μπορεί να εκπληρώσει με ικανοποιητικό τρόπο την αποστολή της.

Η ποιότητα των εικόνων σχετίζεται άμεσα με τον τρόπο σάρωσης που θα επιλεχθεί. Υπάρχουν τρεις τρόποι να σαρώθουν οι εικόνες και αυτοί έχουν να κάνουν με τον αριθμό χρωμάτων που θα θέλαμε να υπάρχουν στο τελικό αρχείο εικόνας: δύο χρώματα, άσπρο και μαύρο, στην περίπτωση που επιλέξουμε σαν τελική μορφή την ασπρόμαυρη εικόνα, διακόσιες πενήντα έξι αποχρώσεις γκρίζου στη περίπτωση που μιλάμε για εικόνα με αποχρώσεις γκρίζου, και τέλος υπάρχει και η έγχρωμη σάρωση που διατηρεί όλα τα χρώματα της εικόνας. Αυτή η τελευταία δεν είναι του άμεσου ενδιαφέροντός μας, αφού τα έγγραφα, συνήθως, δεν περιέχουν άλλα χρώματα πέρα από αποχρώσεις του γκρίζου. Παρ'όλα αυτά υπάρχει δουλειά και σε αυτή τη κατεύθυνση, της έγχρωμης σάρωσης [1], και στα [21], [25], αναφέρονται κάποια πρώτα αποτελέσματα για ένα σχετικά μικρό αριθμό εγγράφων (της τάξης των 20.000).

Και οι δύο άλλες μέθοδοι σάρωσης, σε 2 και σε 256 χρώματα, έχουν τα πλεονεκτήματά τους και τα μειονεκτήματά τους. Εν συντομία θα αναφέρουμε εδώ μερικά και θα παραπέμψουμε τον αναγνώστη στη παράγραφο 3.1.1 για περισσότερες πληροφορίες. Γενικά, η σάρωση εικόνων με 2 χρώματα δημιουργεί εικόνες με δραστικά μικρότερο μέγεθος, άρα και εικόνες που είναι πιο εύκολο να τις αποθηκεύσει κανείς, να τις χειριστεί και να τις μεταδώσει πάνω από ένα δίκτυο. Οι εικόνες αποχρώσεων γκρίζου, αν και πολύ μεγαλύτερες σε μέγεθος αρχείου, κρατούν όλη την πληροφορία που περιέχει ένα έγγραφο, δίνουν πολύ καλύτερα την συνολική εικόνα και υφή του εγγράφου και απαιτούν λιγότερη προσοχή, άρα και χρόνο, στη διαδικασία της σάρωσης [5], [32].

Η ποιότητα των εικόνων μπορεί να βελτιωθεί χρησιμοποιώντας λειτουργίες επεξερ-

γασίας εικόνων. Οι λειτουργίες αυτές προσπαθούν σε γενικές γραμμές να αφαιρέσουν στοιχεία από την εικόνα που δεν θεωρούνται στοιχεία πληροφορίας (μαύρες περιοχές που προέρχονται από κακή σάρωση, μεμονωμένες κουκίδες μαύρων εικονοστοιχείων που μπορεί να προέκυψαν από ανωμαλίες στην επιφάνεια του χαρτιού κα.) ενώ ταυτόχρονα προσπαθούν να τονίσουν και/ή να διορθώσουν ακόμη, τις περιοχές που περιέχουν πληροφορία με την μορφή γραμμάτων, σφραγίδων, σχεδίων κα. Σε αρκετές περιπτώσεις τέτοιου είδους επεμβάσεις δεν είναι καθόλου εύκολο να γίνουν. Αν η αρχική ποιότητα των εγγράφων είναι αρκετά κακή, τότε τα αποτελέσματα που παίρνει κανείς δεν είναι καθόλου ικανοποιητικά. Πειραματισμός και δοκιμές πάνω σε πολλές μεθόδους είναι απαραίτητα στοιχεία για την δημιουργία καλών και αποτελεσματικών προγραμμάτων.

### 1.3.3 Ασφάλεια εικόνων

Με τον όρο ασφάλεια εικόνων αναφερόμαστε στην προστασία του υλικού της βιβλιοθήκης, και δη των εικόνων, από κακόβουλη χρήση. Έχουμε επανειλημμένα τονίσει την μεγάλη αξία και σημασία που μπορούν να έχουν τα έγγραφα μιας βιβλιοθήκης. Το να υπάρχει ένας τρόπος να διασφαλιστεί η πνευματική ιδιοκτησία της βιβλιοθήκης είναι ένας απόλυτα κατανοητός στόχος και μια πολύ θεμιτή επιδίωξη. Σε ένα σύστημα που επιδιώκουμε να προσφέρουμε τη δυνατότητα να το χρησιμοποιούν μεγάλος αριθμός χρηστών και με πολλούς διαφορετικούς τρόπους (άλλοι σαν ερευνητές/σχολιαστές, άλλοι σαν απλοί χρήστες, άλλοι σαν χειριστές του συστήματος, από απόσταση μέσω δικτύου κλπ.) δεν μπορούμε να εμπιστευτούμε την καλόβουλη χρήση του από όλους. Εξάλλου κανείς δεν θα ήθελε να δει υλικό που του ανήκει να χρησιμοποιείται κάπου αλλού χωρίς καν να γίνεται αναφορά στη πηγή από την οποία το προμηθεύτηκε. Ούτε κανείς θα ήθελε κάποιος άλλους να οικιοποιηθεί δικά του πράγματα ή ακόμη και να αποκομίσει χρηματικό όφελος από ξένη ιδιοκτησία. Για τους λόγους αυτούς θα πρέπει με κάποιο τρόπο να τους αποτρέψουμε από την κακόβουλη χρήση χωρίς κάτι τέτοιο να γίνει σε βάρος της ευχρηστίας ή της πολυπλοκότητας του συστήματος.

Αρκετή δουλειά έχει γίνει σε ολόκληρο τον κόσμο προς αυτή την κατεύθυνση, κυρίως με τη χρήση του υδατογραφήματος [4], [20], [39]. Η υδατογραφία είναι μια μέθοδος κατά την οποία μια εικόνα, το υδατογράφημα, που συνήθως είναι κάποιο χαρακτηριστικό έμβλημα του ιδιοκτήτη, ενσωματώνεται στην αρχική εικόνα. Η ενσωμάτωση γίνεται με τέτοιο τρόπο, ώστε αφ'ενός η εικόνα να παραμένει λειτουργική και αφ'ετέρου να μην υπάρχει αμφιβολία για την ιδιοκτησία της εικόνας.

### 1.3.4 Χρήση πάνω από το διαδίκτυο

Το διαδίκτυο (Internet) είναι ένα δίκτυο επικοινωνίας υπολογιστών που έχει συνενώσει τα περισσότερα σημεία του πλανήτη μεταξύ τους. Με τον τρόπο αυτό υπολογιστές μπορούν να επικοινωνούν μεταξύ τους από πολύ μακρινές αποστάσεις διευκολύνοντας

την επαφή των ανθρώπων και κάνοντας την ανταλλαγή απόψεων και πληροφοριών μια εύκολη υπόθεση.

Τα τελευταία χρόνια το διαδίκτυο έχει γνωρίσει τρομερή ανάπτυξη και έχει γίνει ευρέως γνωστό ακόμη και σε ανθρώπους που δεν είχαν προηγουμένως σχέση με υπολογιστές. Η ανάπτυξή του συνεχίζεται ακόμη και σήμερα με πολύ γοργούς ρυθμούς ενώ ταυτόχρονα συνεχίζεται η προσπάθεια για την ποικιλόμορφη και ποικιλότροπη εκμετάλευσή του.

Ανάμεσα στα άλλα το διαδίκτυο έχει χρησιμοποιηθεί για την εύκολη πρόσβαση σε κάθε είδους πηγές πληροφορίας, ξεκινώντας από ανώνυμους ιδιώτες που με τη μορφή ενασχόλησης στο ελεύθερό τους χρόνο έχουν φτιάξει μικρές βάσεις δεδομένων και φτάνοντας στις τεράστιες βιβλιοθήκες ιδρυμάτων, κυβερνητικών οργανισμών και ιδιωτικών επιχειρήσεων.

Γίνεται επομένως φανερό ότι αν δεν θέλει κανείς να αποκλειστεί από όλες αυτές τις δυνατότητες και τις προοπτικές που ανοίγονται μέσα από το διαδίκτυο θα πρέπει να λάβει σοβαρά υπ'όψιν του την χρήση του μέσα από οποιαδήποτε εφαρμογή προσπαθεί να αναπτύξει. Όχι μόνο προσφέρει καλύτερες υπηρεσίες στους χρήστες διευκολύνοντάς τους στο να μπορούν ακόμη και από μακρινές αποστάσεις να έχουν πρόσβαση στο υλικό της βιβλιοθήκης, όχι μόνο προβάλλεται η ίδια η βιβλιοθήκη και το πρόγραμμα προς τα έξω, αλλά είναι και ένας εύκολος τρόπος για να εμπλουτιστεί η βιβλιοθήκη με καινούργιο υλικό μέσω της ηλεκτρονικής ανταλλαγής δεδομένων σε συνεργασία με άλλες βιβλιοθήκες οπουδήποτε στο κόσμο.

Οι περισσότερες από τις σύγχρονες ψηφιακές βιβλιοθήκες προβάλλουν το υλικό τους και μέσω του διαδικτύου [25], [1], [22], [35], [36].

## 1.4 Οργάνωση του υλικού

Η αναφορά αυτή έχει οργανωθεί ως εξής: στο κεφάλαιο 2 γίνεται αναλυτικά η παρουσίαση του προγράμματος ΑΡΧΟΝ. Ακολουθούν στο κεφάλαιο 3 όλες οι λεπτομέρειες που αφορούν τη δική μας εργασία πάνω στα διάφορα υποσυστήματα του ΑΡΧΟΝ καθώς και όλα τα συμπεράσματα και τα αποτελέσματα που προέκυψαν από την εργασία αυτή. Η αναφορά κλείνει με το κεφάλαιο 4 και τον απολογισμό της εργασίας καθώς και την παρουσίαση των πιθανών μελλοντικών επεκτάσεων και προοπτικών.

# Κεφάλαιο 2

# Σύστημα ΑΡΧΟΝ

## 2.1 Γενική περιγραφή

Το πρόγραμμα ΑΡΧΟΝ αποβλέπει στην ανάπτυξη ενός πολύμεσου υπολογιστικού συστήματος για την αρχειοθέτηση, σχολιασμό και ανάκτηση ιστορικών εγγράφων. Τα έγγραφα αυτά βρίσκονται στη Βικελαία Δημοτική Βιβλιοθήκη του Δήμου Ηρακλείου και αποτελούν αντικείμενα μεγάλης πολιτισμικής και ιστορικής αξίας. Το πλήθος τους είναι πολύ μεγάλο και ανέρχεται σε περίπου 1.500.000 έγγραφα[1]. Τα έγγραφα καλύπτουν ευρύ φάσμα ως προς την κατάσταση στην οποία βρίσκονται, και χρονολογούνται από την αρχή του 17$^{ου}$ ως την αρχή του 20$^{ου}$ αιώνα. Προς το παρόν, δεν υπάρχει κάποιος εύκολος και αποδοτικός τρόπος χρήσης των εν λόγω εγγράφων, μιας και απαιτείται η φυσική πρόσβαση στα πρωτότυπα για οποιαδήποτε μελέτη. Η ανάπτυξη του συστήματος ΑΡΧΟΝ δίνει βιώσιμη λύση σ'αυτό το πρόβλημα και επιπλέον συμβάλλει στην πιο αποτελεσματική χρησιμοποίησή τους, αφού καταστεί δυνατή την εύκολη πρόσβαση, μελέτη αλλά και σχολιασμό των εγγράφων με διάφορα στοιχεία.

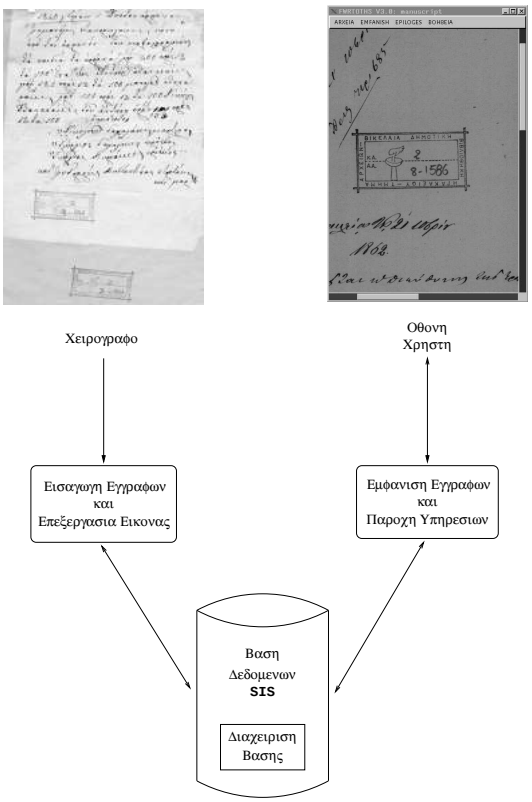
Η ανάπτυξη άρχισε την 1η Ιανουαρίου 1997 με κύριο ανάδοχο το Ίδρυμα Τεχνολογίας και Έρευνας και φορέα χρήστη τη Βικελαία Δημοτική Βιβλιοθήκη Ηρακλείου. Η χρηματοδότηση του έργου εξασφαλίστηκε από το Περιφερειακό Επιχειρησιακό Πρόγραμμα (ΠΕΠ) Κρήτης 1994-1999.

Η ανάπτυξη του συστήματος ΑΡΧΟΝ βασίζεται στο Σύστημα Σημασιολογικού Ευρετηριασμού (ΣΣΕ) που έχει σχεδιαστεί και υλοποιηθεί από τον Τομέα Πληροφοριακών Συστημάτων και Τεχνολογίας Λογισμικού του Ινστιτούτου Πληροφορικής του ΙΤΕ. Το Σύστημα Σημασιολογικού Ευρετηριασμού χρησιμοποιείται για την καταχώρηση και ανάκτηση των εγγράφων και η διασύνδεσή του με τα άλλα υποσυστήματα (αποτίμησης και αρχειοθέτησης εικόνων εγγράφων, βελτίωσης ποιότητας των εικόνων, συμπίεσης/αποσυμπίεσης, υποσύστημα επαφής χρήστη) έχει ως αποτέλεσμα ένα ενιαίο σύστημα, κατάλληλο για την εν λόγω εφαρμογή.

Η δομή του ΑΡΧΟΝ φαίνεται στο Σχ. 2.1.

Μπορούμε να δούμε ότι το σύστημα χωρίζεται σε τρία επιμέρους υποσυστήματα

[1] ο ακριβής αριθμός των εγγράφων δεν είναι γνωστός.




Σχήμα 2.1: Γενική δομή

Το πρώτο αφορά τις διαδικασίες εκείνες με τις οποίες μπορούμε από τα έγγραφα στη πρωτότυπή τους μορφή να καταλήξουμε σε εικόνες οι οποίες θα αποθηκευτούν στη βάση δεδομένων. Το τμήμα αυτό επικοινωνεί με τη βάση δεδομένων για να της προσφέρει τις πληροφορίες που χρειάζονται για μια σωστή (ολοκληρωμένη από άποψη απαραίτητων πληροφοριών) εισαγωγή. Η βάση δεδομένων στηρίζεται στο Σύστημα Σημασιολογικού Ευρετηριασμού για την οργάνωση και διαχείριση των δεδομένων. Οι πληροφορίες για την κάθε εικόνα οργανώνονται με τέτοιο τρόπο ώστε να είναι εύκολη η ανάκτησή τους και η συσχέτισή τους με τις πληροφορίες άλλων εικόνων. Με τη βάση δεδομένων επικοινωνεί και το υποσύστημα επαφής χρήστη. Παίρνει από αυτή τα αποτελέσματα των εντολών αναζήτησης του χρήστη και με το πρόγραμμα εμφάνισης εικόνων προβάλλει στην οθόνη την αντίστοιχη εικόνα.

Τα υποσυστήματα του ΑΡΧΟΝ περιγράφονται αναλυτικά παρακάτω.

## 2.2 Υποσύστημα εισαγωγής εγγράφων και επεξεργασίας εικόνας

Το υποσύστημα εισαγωγής εγγράφων και επεξεργασίας εικόνας αποτελεί τον συνδετικό κρίκο μεταξύ των πρωτότυπων κειμένων και της ψηφιακής βιβλιοθήκης. Οι λειτουργίες που το συνθέτουν αποτελούν και τον κύριο όγκο της εργασίας αυτής. Για το λόγο αυτό αναβάλλουμε τη λεπτομερή περιγραφή του υποσυστήματος αυτού για το κεφάλαιο

3. Μας αρκεί να γνωρίζουμε, για την ώρα, πως στόχος του υποσυστήματος αυτού είναι η μετατροπή του αρχείου εγγράφων σε ηλεκτρονική ψηφιακή μορφή, για την αποθήκευσή του σε μια βάση δεδομένων.

## 2.3 Υποσύστημα βάσης δεδομένων

Όπως ήδη αναφέραμε το υποσύστημα αυτό βασίζεται στο Σύστημα Σημασιολογικού Ευρετηριασμού (ΣΣΕ) [37]. Το ΣΣΕ είναι ένα εργαλείο για την περιγραφή και τεκμηρίωση μεγάλων ποσοτήτων δεδομένων [6]. Στη περίπτωση που υπάρχουν πολλά δεδομένα με πολύπλοκες σχέσεις ανάμεσα τους, τότε το ΣΣΕ έχει περιγραφική δύναμη σαφώς ανώτερη από εκείνη των παραδοσιακών βάσεων δεδομένων.

Το ΣΣΕ αποτελείται από ένα συνεπή[2] μηχανισμό αποθήκευσης ο οποίος βασίζεται σε ένα οντοκεντρικό μοντέλο δεδομένων σημασιολογικών δικτύων [19]. Το μοντέλο αυτό των δεδομένων οργανώνει την πληροφορία των αντικειμένων που περιγράφει χρησιμοποιώντας σημασιολογικά δίκτυα. Τα σημασιολογικά δίκτυα αποτελούνται από αντικείμενα και διαφόρων ειδών συσχετίσεις ανάμεσά τους.

Τα δεδομένα που μπορεί να οργανώσει δεν περιορίζονται σε έναν τύπο αλλά μπορούν είναι εικόνες, κείμενα, ήχοι και/ή βίντεο. Μπορεί να χειριστεί μια συνεχώς αυξανόμενη ποσότητα δομικής πληροφορίας καθώς προκύπτουν καινούργιες σχέσεις που συνδέουν τα δεδομένα μεταξύ τους. Μπορεί να χρησιμοποιήσει ελάχιστη, αρχικά, πληροφορία για τα αντικείμενα που οργανώνει και σταδιακά να εμπλουτίζεται με νέες ποσότητες πληροφοριών. Υποστηρίζει την έννοια της δοσοληψίας (μια σειρά ενεργειών που τις θεωρούμε, λογικά, σαν μια αδιαίρετη ενέργεια, transaction) καθώς και την ταυτόχρονη πρόσβαση πολλών χρηστών.

Η επιλογή του βασίστηκε στη πλούσια περιγραφική του δύναμη όσον αφορά το σημασιολογικό περιεχόμενο των εγγράφων [8], στη μεγάλη ταχύτητα απόκρισης που έχει, δεδομένου του μεγάλου όγκου πληροφοριών που θα έχουμε στη βάση μας, και στη δυνατότητα προγραμματισμού των επερωτήσεων για ανεύρεση δεδομένων. Η δυνατότητα προγραμματισμού των επερωτήσεων είναι ουσιώδης αφού με τον τρόπο αυτό μπορεί ένα εξωτερικό, του ΣΣΕ, εργαλείο να έχει πρόσβαση στις πληροφορίες του συστήματος.

## 2.4 Υποσύστημα εμφάνισης εγγράφων και παροχής υπηρεσιών

Το δεύτερο σημαντικό τμήμα αυτής της εργασίας αναφέρεται στο υποσύστημα εμφάνισης εγγράφων και παροχής υπηρεσιών. Το υποσύστημα αυτό ουσιαστικά αποτελεί το υποσύστημα επαφής του χρήστη. Είναι δηλαδή, το κομμάτι εκείνο του ΑΡΧΟΝ που είναι υπεύθυνο για την χρησιμοποίηση του συστήματος από τον χρήστη, και την

[2] κατά την εισαγωγή πληροφορίας στη βάση, γίνεται αυτόματος έλεγχος για την σημασιολογική πληρότητα αυτής της πληροφορίας.

παρουσίαση των αποτελεσμάτων των ερωτήσεων του χρήστη. Οι δυνατότητές του καθορίζουν σε μεγάλο βαθμό την ευκολία χρήσης του συστήματος. Η αναλυτική περιγραφή και αυτού του υποσυστήματος θα γίνει στο κεφάλαιο 3.

## 2.5 Υποσύστημα επικοινωνίας και διασύνδεσης

Στα προηγούμενα υποσυστήματα θα πρέπει να προσθέσουμε και ένα άλλο, λιγότερο εμφανές, αλλά εξίσου σημαντικό: το υποσύστημα επικοινωνίας και διασύνδεσης. Σκοπός του είναι να επιτρέψει στα διάφορα υποσυστήματα να επικοινωνούν μεταξύ τους.

Μεσολαβεί μεταξύ του υποσυστήματος της επεξεργασίας εικόνας και της βάσης δεδομένων για να εφοδιάσει τη βάση με όλες εκείνες της πληροφορίες που απαιτούνται για τη σωστή καταχώρηση μιας νέας εικόνας. Για το σκοπό αυτό μετατρέπει όλες τις πληροφορίες που λαμβάνει, σε μορφή που η βάση δεδομένων μπορεί να κατανοήσει και κάνει τις κατάλληλες ενέργειες για να προκαλέσει την εισαγωγή των δεδομένων.

Μεσολαβεί επίσης, μεταξύ των υποσυστημάτων της βάσεως δεδομένων και του προγράμματος εμφάνισης για να επιτρέψει τόσο την αποστολή των επερωτήσεων στη βάση δεδομένων όσο και τη λήψη των απαντήσεων στις επερωτήσεις για την μετέπειτα εμφάνισή τους στην οθόνη.

Δεδομένης της απαίτησης για λειτουργία του συστήματος πάνω από δίκτυο (για παροχή δυνατότητας χρήσης του διαδικτύου) έπρεπε να γίνει επιλογή εργαλείων και μεθόδων ανάπτυξης που να επιτρέπουν τέτοιας μορφής λειτουργικότητα. Η Java [26] είναι μια γλώσσα προγραμματισμού που δίνει τέτοιες δυνατότητες και υποστηρίζει την λειτουργία προγραμμάτων πάνω από δίκτυο. Η χρήση της μαζί με τη χρήση Servlet (προγράμματα που παρέχουν διάφορες υπηρεσίες μέσα από το περιβάλλον ενός φυλομετρητή - browser) μας έδωσε τη δυνατότητα να δημιουργήσουμε ένα αμιγώς δικτυακό υποσύστημα επικοινωνίας και να συνδέσουμε έτσι τα διάφορα υποσυστήματα μεταξύ τους.

## 2.6 Αναμενόμενα οφέλη

Τα αποτελέσματα του προγράμματος ΑΡΧΟΝ αναμένεται να συμβάλλουν αποφασιστικά στη διατήρηση, διαφύλαξη και μελέτη της πολιτισμικής κληρονομιάς της Βικελαίας Δημοτικής Βιβλιοθήκης Ηρακλείου, και συγκεκριμένα των Αρχείων Δημογεροντίας, Αρχείων Βενετοκρατίας και του Τουρκικού Αρχείου, που βρίσκονται στο Τμήμα Μη-Δανειζομένων Βιβλίων και Αρχείων της βιβλιοθήκης.

Τα προαναφερθέντα αρχεία περιέχουν ιστορικούς θησαυρούς μεγάλης αξίας. Τα αρχεία στην παρούσα τους μορφή είναι δύσχρηστα και επιρρεπή σε φθορές. Το περιεχόμενό τους είναι πολύ δύσκολο να διασφαλιστεί εφ'όσον αυτό δεν υπάρχει φυλαγμένο σε καμμία άλλη μορφή. Πέραν αυτών, η ίδια η χρήση τους είναι δύσκολη αφού η αναζήτηση συγκεκριμένων πληροφοριών είναι ένα επίπονο, και πολλές φορές

άκαρπο, έργο. Η ολοκλήρωση του έργου θα δώσει λύσεις στα παραπάνω αφού θα καταστήσει δυνατή τη διατήρηση των αρχείων και θα παρέχει μεγάλη ευκολία στη μελέτη τους. Επίσης, η καταχώρηση των αρχείων θα επιτρέψει και την ανταλλαγή με άλλες βιβλιοθήκες (ηλεκτρονική ανταλλαγή δεδομένων) αφού η προσέγγιση που προτάθηκε για την καταχώρηση των εγγράφων είναι αυτή που ακολουθείται διεθνώς.

# Κεφάλαιο 3

# Διαχείριση Εικόνων Εγγράφων

## 3.1 Στάδιο εισαγωγής

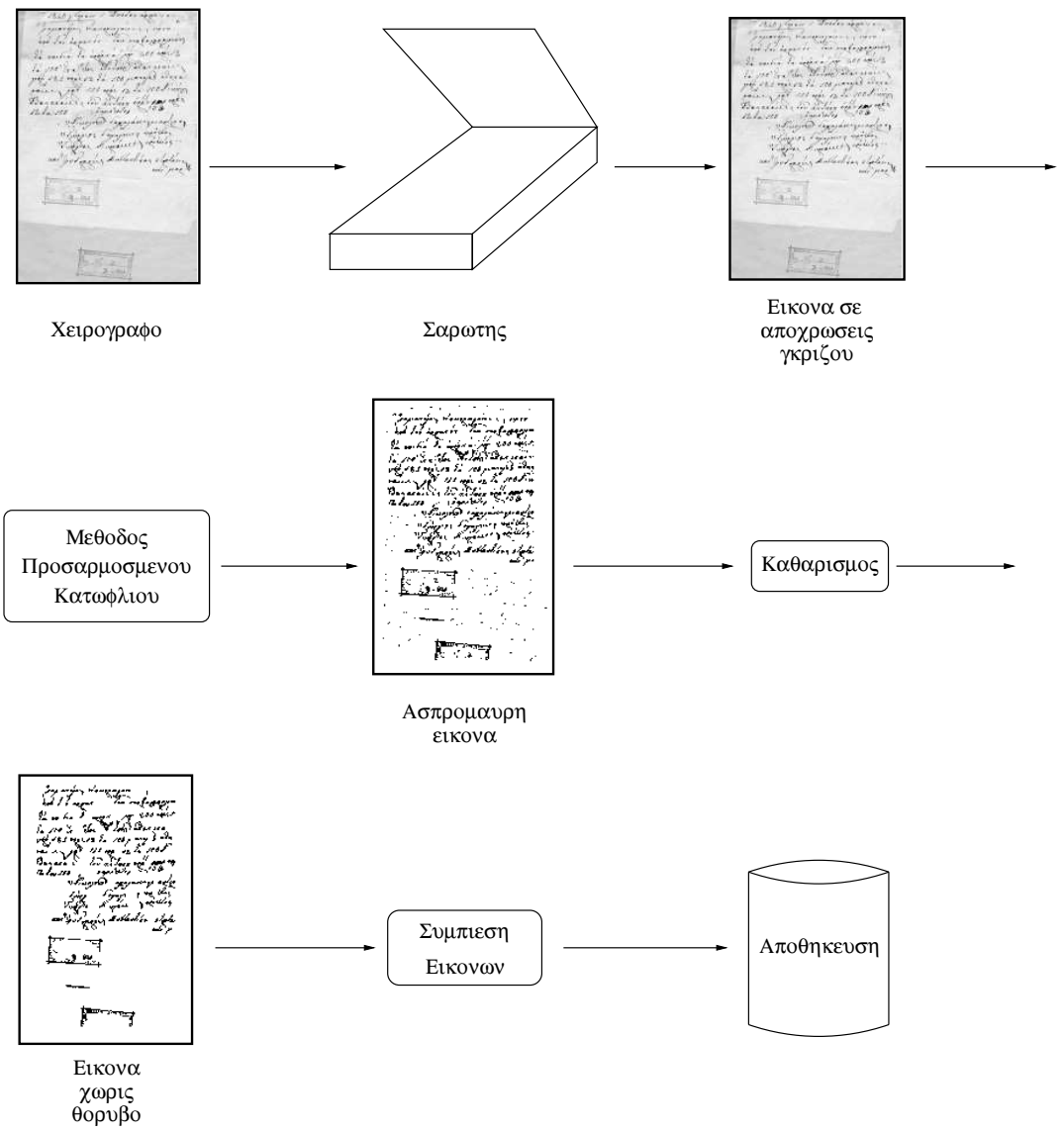
Στο στάδιο της εισαγωγής σκοπός μας είναι να αποκτήσουμε μια όσο το δυνατόν πιο πιστή (σε όλες της τις λεπτομέρειες) εικόνα του πραγματικού χειρόγραφου. Αυτός όμως ο σκοπός περιέχει πολλούς αντικρουόμενους παράγοντες [32], [31], [24] όπως θα δούμε αναλυτικά και πιο κάτω. Στο Σχ. 3.1 φαίνεται η όλη ακολουθία των λειτουργικών διεργασιών που απαιτούνται προκειμένου να είναι έτοιμη μια εικόνα για εισαγωγή στη βάση δεδομένων.

Η διαδικασία ξεκινά από το έγγραφο στη μορφή του πρωτότυπου. Από αυτό με τη χρήση ενός σαρωτή αποκτούμε την εικόνα του κειμένου σε ψηφιακή μορφή. Αυτή η εικόνα είναι κακής ποιότητας εικόνα αποχρώσεων γκρίζου (grayscale image). Οι εικόνες με αποχρώσεις γκρίζου είναι γενικά μεγάλες σε μέγεθος αρχείου. Χρειαζόμαστε λοιπόν μια μέθοδο για να βελτιώσουμε την ποιότητά τους και για να περάσουμε σε ασπρόμαυρες εικόνες (binary ή black & white images) που είναι αρκετά μικρότερες σε μέγεθος. Η ασπρόμαυρη εικόνα που προκύπτει έχει γενικά αρκετό θόρυβο. Ακολουθεί μια μέθοδος αφαίρεσης του θορύβου που βελτιώνει περεταίρω την ποιότητα της εικόνας. Παρότι στο σημείο αυτό οι εικόνες μας έχουν μικρότερο μέγεθος από το αρχικό των εικόνων αποχρώσεων γκρίζου, το μέγεθος τους παραμένει μεγάλο. Ένα επιπλέον στάδιο ακολουθεί, αυτό της συμπίεσης της εικόνας. Μετά από το στάδιο αυτό η εικόνα είναι έτοιμη για να εισαχθεί, μαζί με τις υπόλοιπες απαραίτητες πληροφορίες, στη βάση δεδομένων.

### 3.1.1 Σάρωση χειρογράφων

Επειδή ο πρωταρχικός μας στόχος είναι η όσο το δυνατόν καλύτερη ποιότητα εικόνας [24], στο στάδιο αυτό έπρεπε να γίνει επιλογή σαρωτή και μεθόδου σάρωσης που να δώσουν τα βέλτιστα αποτελέσματα. Υπάρχουν τρεις τρόποι να σαρώσει κανείς ένα αντικείμενο[1] σε ψηφιακή μορφή:

[1] έγγραφο, φωτογραφία, βιβλίο, περιοδικό.




Σχήμα 3.1: Γενική δομή

| Είδος χαρτιού | Διαστάσεις εικόνας[2] | Μέγεθος εικόνας αποχρώσεων γκρίζου | Μέγεθος ασπρόμαυρης εικόνας |
|---|---|---|---|
| A3 | 4961 x 3508 | 17403248 | 2177892 |
| A4 | 3508 x 2480 | 8699900 | 1087579 |
| B5 | 2953 x 2074 | 6124582 | 767793 |
| Τηλεομοιότυπου[3] | 2376 x 1728 | 4105788 | 513229 |
| Letter | 3300 x 2550 | 8415060 | 1052713 |
| Executive | 3150 x 2175 | 6851310 | 856813 |

Πίνακας 3.1: Τυπικά μεγέθη αρχείων εικόνων

**Έγχρωμη σάρωση** Εδώ γίνεται σάρωση με τέτοιο τρόπο ώστε να διατηρήσουμε όλη τη χρωματική πληροφορία. Είναι απαραίτητη για φωτογραφίες και άλλο έγχρωμο υλικό αλλά δημιουργεί πολύ μεγάλα αρχεία εικόνας.

**Σάρωση σε 256 επίπεδα γκρίζου** Για φωτογραφίες ή έντυπο υλικό που δεν έχει άλλα χρώματα παρά μόνο αποχρώσεις γκρίζου είναι η ενδεδειγμένη λύση. Διατηρεί αρκετή χρωματική πληροφορία και δίνει μεσαίας τάξης μέγεθος αρχείων.

**Ασπρόμαυρη σάρωση** Στη περίπτωση που το υλικό μας δεν έχει παρά μόνο δύο χρώματα μπορούμε να το σαρώσουμε με ασπρόμαυρη σάρωση. Είναι η πιο οικονομική σε μέγεθος αρχείων μέθοδος.

Η τάση που επικρατεί στην αγορά όσον αφορά τη σάρωση εγγράφων, με την προϋπόθεση τα έγγραφα να μην έχουν έγχρωμη πληροφορία, είναι αυτά να σαρώνονται σε δυαδική (binary) μορφή, να έχουν δηλαδή μόνο δυο χρώματα, άσπρο και μαύρο, στην τελική τους μορφή. Ο κύριος λόγος που έχει οδηγήσει σε μια τέτοια απόφαση είναι το ότι με αυτό τον τρόπο έχουμε πολύ μικρότερες απαιτήσεις σε μέγεθος αρχείων (βλέπε πίνακα 3.1).

Ένας άλλος λόγος είναι πως η ασπρόμαυρη σάρωση είναι "λιγότερο απαιτητική". Ακριβώς εξαιτίας του μεγάλου μεγέθους αρχείων που δημιουργούνται από τη σάρωση σε 256 επίπεδα γκρίζου, ο εξοπλισμός που χρειάζεται πρέπει να είναι αρκετά ισχυρός (μεγάλη κεντρική μνήμη, ισχυρός επεξεργαστής, μεγάλοι αποθηκευτικοί δίσκοι, περιφερειακά υψηλών αποδόσεων σε μεταφορά δεδομένων) και συνεπακόλουθα αρκετά ακριβός. Με την ασπρόμαυρη σάρωση μπορούν να χρησιμοποιηθούν και "μικρότερες" μηχανές που είναι πιο προσιτές στο ευρύ αγοραστικό κοινό.

Η ασπρόμαυρη σάρωση είναι αρκετά διαδεδομένη και για ένα ακόμη λόγο: είναι αρκετά αποτελεσματική σε πάρα πολλές περιπτώσεις. Η πλειονότητα των εφημερίδων, των δακτυλογραφημένων κειμένων, καθώς και ένα αρκετά μεγάλο μέρος των βιβλίων

[2] σε πλήθος εικονοστοιχείων.

[3] σελίδα fax.

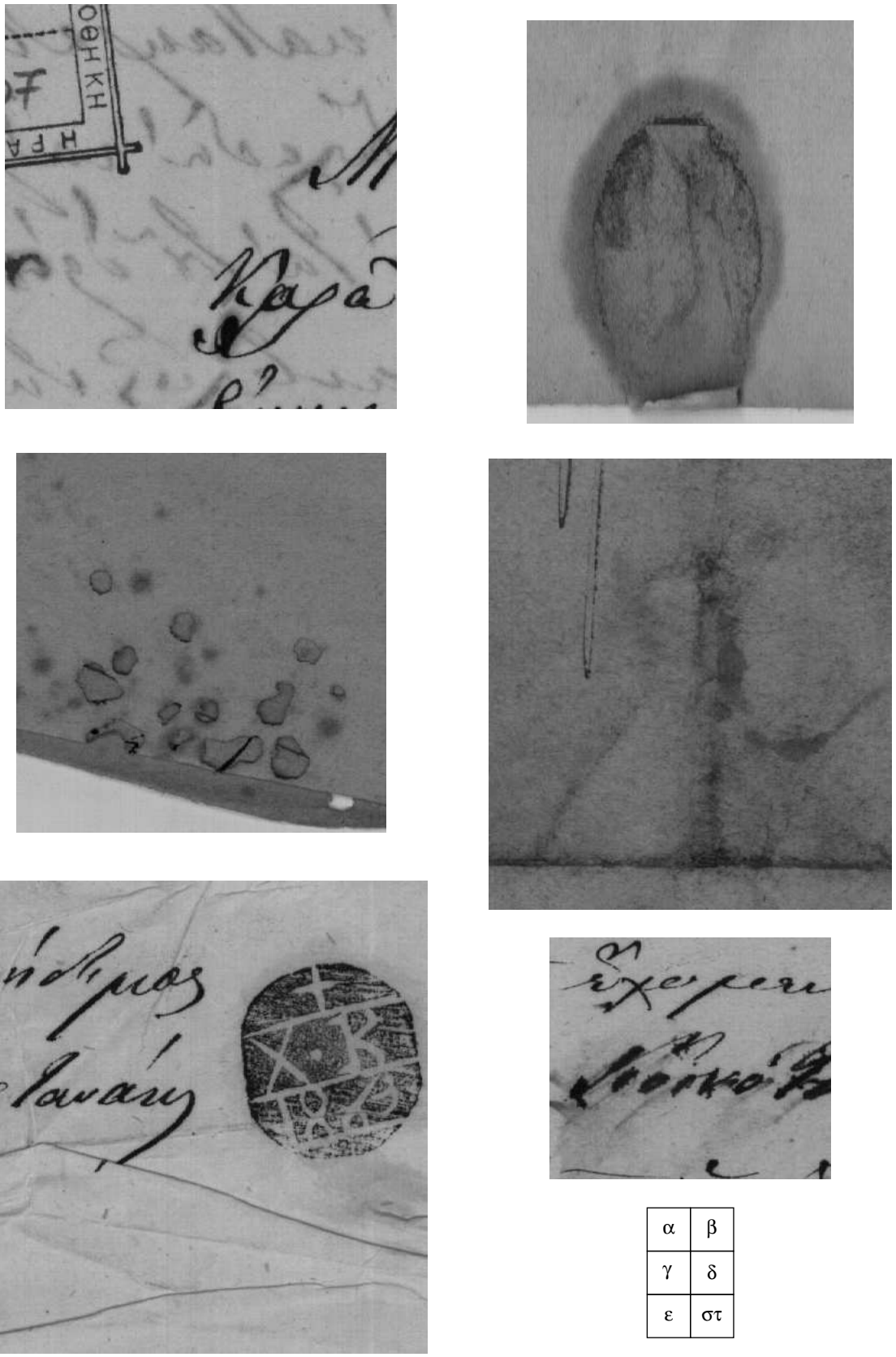
είναι τυπωμένα με μαύρο μελάνι μόνο. Σε αυτή την περίπτωση εξάλλου δεν υπάρχει φόβος απώλειας κάποιου είδους πληροφορίας, χρωματικής ή σε επίπεδο λεπτομέρειας, μιας και όλες οι μέθοδοι σάρωσης θα αποφέρουν την ίδια ποσότητα πληροφορίας. Ως εκ τούτου η ασπρόμαυρη σάρωση είναι ο φυσικός τρόπος να ψηφιοποιήσει κανείς το υλικό αυτό.

Στη γενική περίπτωση τα έγγραφα έχουν κάποιου είδους χρωματική πληροφορία. Περνώντας από τη σάρωση σε 256 επίπεδα γκρίζου σε 2 χρώματα κερδίζουμε σε μέγεθος όπως ήδη αναφέραμε. Είναι όμως αυτονόητο πως ταυτόχρονα έχουμε πτώση της ποιότητας των εικόνων, που σε μερικές περιπτώσεις μπορεί να είναι αρκετά σημαντική, σε σχέση με αυτή των αρχικών εγγράφων. Και αυτό γιατί δεν είναι πάντα εύκολο να αποδώσει κανείς με δύο χρώματα όλη την πληροφορία που περιέχει μια εικόνα με 256 χρώματα (αποχρώσεις γκρίζου).

Στη δική μας περίπτωση η ποιότητα των εγγράφων ήταν αρκετά κακή (βλέπε Σχ. 3.2). Τα έγγραφα είναι αρκετά παλιά, μερικά χρονολογούνται στις αρχές του $17^{ου}$ αιώνα, και αυτό έχει σαν αποτέλεσμα το χαρτί να έχει υποστεί αλλοιώσεις και να έχει αλλάξει το χρώμα του. Πολλά από αυτά έχουν υποστεί φθορές από το χρόνο και έχουν σαπίσει σε μερικά σημεία, σε κάποια άλλα έχουν δημιουργηθεί τρύπες, ενώ υπάρχουν και αρκετά που εξαιτίας των πολλών διπλωμάτων που έχουν υποστεί τόσα χρόνια έχει μείνει έντονη η τσάκιση στο χαρτί. Άλλα έχουν ποτίσει από νερό ή άλλα υγρά που έχουν πέσει πάνω τους με αποτέλεσμα το κείμενο να έχει γίνει πιο αχνό στη περιοχή εκείνη. Σε μερικές περιπτώσεις στο ίδιο έγγραφο έχουν χρησιμοποιηθεί πολλά μελάνια. Τα μελάνια αυτά μπορεί να είναι διαφορετικών χρωμάτων ή μερικά να είναι πιο αχνά (ή πιο έντονα) από κάποια άλλα. Σε αρκετά από τα έγγραφα υπάρχουν με τη μορφή σημειώσεων κάποιες επιπλέον πληροφορίες που σε μερικές περιπτώσεις είναι και πολύ μεταγενέστερες. Οι πληροφορίες αυτές έχουν συνήθως γραφεί με διαφορετικό μελάνι, ακόμη και με μολύβι. Επιπρόσθετες δυσκολίες προέρχονται από την ύπαρξη χρωμάτων, όπως κατά τη χρήση βουλοκεριού, που συνήθως είναι κόκκινου χρώματος, από τη χρήση σφραγίδων (μπλε ή μωβ χρώματος). Ακόμη και ο τρόπος σφραγίσματος του εγγράφου δημιουργεί δυσκολίες, σε ορισμένες περιπτώσεις, αφού προκαλεί ανώμαλες επιφάνειες λόγω του υλικού που έχει χρησιμοποιηθεί.

Για όλους αυτούς τους λόγους το να δοκιμάζαμε να σαρώναμε τις εικόνες σε 2 χρώματα απευθείας ήταν κάτι σχεδόν απαγορευτικό. Σε κάποιες πρώτες, πειραματικές δοκιμές, τα αποτελέσματα ήταν πολύ φτωχά, γι'αυτό και μια τέτοια προσέγγιση δεν θα έπρεπε να γίνει αποδεκτή χωρίς πρώτα να υπάρξει διερεύνηση για τυχόν εναλλακτικές λύσεις.

Σαν μια τέτοια εναλλακτική λύση θεωρήθηκε η ιδέα της σάρωσης των εγγράφων σε επίπεδα γκρίζου και κατόπιν η μείωση του αριθμού των χρωμάτων σε δύο. Ακολουθήσαμε μια τέτοια προσέγγιση προσπαθώντας να εκμεταλλευτούμε τα πλεονεκτήματα και των δύο λύσεων: με την σάρωση σε 256 επίπεδα γκρίζου δεν έχουμε καθόλου χάσιμο πληροφορίας, και με το πέρασμα σε δύο χρώματα κατόπιν, έχουμε μείωση του όγκου στα αρχεία εικόνας. Το πέρασμα από τα 256 στα 2 χρώματα γίνεται με αλγοριθμικό



| | |
|---|---|
| α | β |
| γ | δ |
| ε | στ |

Σχήμα 3.2: Παραδείγματα εγγράφων κακής ποιότητας από α) πότισμα μελανιού, β) χρήση βουλοκεριού, γ) τρύπες στο χαρτί, δ) τσάκισμα του χαρτιού, ε) σφραγίδα και τσαλακωμένο χαρτί και στ) μουντζούρες στο χαρτί

τρόπο.

Στο σημείο αυτό πρέπει κανείς να πάρει μια σπουδαία απόφαση: να αποφασίσει τι θεωρεί πληροφορία και τι όχι. Η απόφαση αυτή είναι πολύ σημαντική γιατί θα έχει μεγάλες επιπτώσεις στις εικόνες που θα δημιουργηθούν σαν αποτέλεσμα της μείωσης των χρωμάτων. Σε κάθε έγγραφο υπάρχουν εκτός από το περιεχόμενο του κειμένου αυτό καθ'εαυτό, και άλλου είδους πληροφορίες που μπορεί κανείς έμμεσα να εξάγει. Τέτοιου είδους πληροφορίες είναι το είδος του χαρτιού που έχει χρησιμοποιηθεί, ο τρόπος με τον οποίον το κείμενο έχει γραφεί πάνω στο χαρτί, ο τρόπος με τον οποίο έχει σφραγιστεί το κείμενο, το αν είναι φθαρμένο ή όχι το χαρτί σαν αποτέλεσμα της συνεχής χρήσης του, το αν έχει χυθεί μελάνι σε κάποια σημεία του, κα. Όλα αυτά δίνουν πληροφορίες για τις συνθήκες που έχει η συγγραφή, για το ύφος και την ιδιοσυγκρασία του συγγραφέα και διάφορες άλλες πληροφορίες για την κοινωνική τάξη στην οποία βρισκόταν ο συγγραφές ή το πρόσωπο που έχει ζητήσει τη συγγραφή. Αυτές οι πληροφορίες είναι αρκετά σημαντικές σε μερικές περιπτώσεις αλλά δεν είναι αυτές που οι περισσότεροι μελετητές ψάχνουν. Έτσι πήραμε την απόφαση ότι τέτοιου είδους πληροφορίες δεν μας ενδιαφέρουν και πως δεν θα προσπαθήσουμε να τις διατηρήσουμε. Στις σπάνιες περιπτώσεις που κάποιος αναζητά τέτοιου είδους πληροφορίες θα πρέπει να χρησιμοποιήσει το πρωτότυπο αρχείο. Για τη δική μας εργασία θεωρήσαμε ότι θα έπρεπε να διατηρήσουμε όλη την πληροφορία που περιέχεται στα γράμματα του κειμένου, στις σφραγίδες που έχει το έγγραφο, στα τυχόν σχέδια που μπορεί να έχει, και σε υποσημειώσεις ή σε νεότερες προσθήκες που έχουν γίνει (και που πολλές φορές είναι με διαφορετικό μελάνι).

Μετά από κάποιες πρώτες δοκιμές που μας έδειξαν πως η αλγοριθμική μας προσέγγιση ήταν εφικτή και πως θα είχαμε καλύτερα αποτελέσματα από την κλασική μέθοδο, αποφασίσαμε να υιοθετήσουμε την ιδέα αυτή: σάρωση με ένα σαρωτή που θα δίνει σαν έξοδο εικόνες επιπέδων γκρίζου και κατόπιν πέρασμα σε ασπρόμαυρες εικόνες.

Επιλέχθηκε λοιπόν ένας σαρωτής υψηλής ανάλυσης ο οποίος θα μπορούσε να δώσει εικόνες υψηλής ευκρίνειας και ταυτόχρονα θα παρείχε την δυνατότητα να σαρώνει εικόνες σε αποχρώσεις γκρίζου. Επειδή τα έγγραφα προς σάρωση είχαν διάφορα μεγέθη, που σε μερικές περιπτώσεις φτάναν και το μέγεθος A3, ο σαρωτής θα έπρεπε να σαρώνει εικόνες ακόμη και τέτοιου μεγέθους. Η τελική επιλογή ήταν το μοντέλο 500 FB της εταιρίας Bell & Howell [14]. Το μοντέλο αυτό μπορεί να χειριστεί έγγραφα με μέγεθος έως και A3, δίνει σαν έξοδο της σάρωσης εικόνες με 256 επίπεδα γκρίζου, και μπορεί να σαρώσει έγγραφα με λεπτομέρεια ανάλυσης από 60 ως 800 κουκίδες για κάθε ίντσα (dpi - dots per inch).

### 3.1.2 Μετατροπή της εικόνας σε ασπρόμαυρη.

Το να περάσει κανείς από 256 χρώματα (επίπεδα γκρίζου) σε 2 (ασπρόμαυρη εικόνα) [34], [16], [31], είναι αρκετά εύκολο όταν έχουμε καλοδιατηρημένα έγγραφα. Όταν όμως η ποιότητα των εγγράφων δεν είναι αρκετά ικανοποιητική, όπως στη δική μας

περίπτωση, τότε η προσπάθεια αυτή δεν είναι και τόσο εύκολη. Στη συνέχεια θα αναπτυχθούν διάφορες μέθοδοι για την επίτευξη αυτού του σκοπού και θα συγκριθούν χρησιμοποιώντας την εικόνα 3.3 σαν εικόνα εισόδου. Ένα επιπλέον τμήμα εικόνας σε μεγάλη μεγέθυνση θα παρουσιαστεί για να διακριθούν καλύτερα οι λεπτομέρειες (Σχ. 3.8).

#### 3.1.2.1 Ενιαίο κατώφλι

Η απλούστερη μέθοδος μετατροπής της εικόνας σε ασπρόμαυρη είναι η μέθοδος ενιαίου κατωφλίου (global thresholding) [40], [3] η οποία βασίζεται στην εύρεση μιας τιμής για το κατώφλι, τιμή η οποία προκύπτει σαν η μέση τιμή των φωτεινοτήτων όλης της εικόνας. Έτσι αν η εικόνα μας έχει πλάτος, π, και ύψος, υ, και η φωτεινότητα σε κάθε της θέση (x, y) είναι I(x, y) τότε η μέθοδος του ενιαίου κατωφλίου δίνεται από τον ακόλουθο ψευδοκώδικα:

$$\text{Κατώφλι} = \frac{\sum_{i}^{\pi} \sum_{j}^{\upsilon} I(i,j)}{\pi \cdot \upsilon}$$

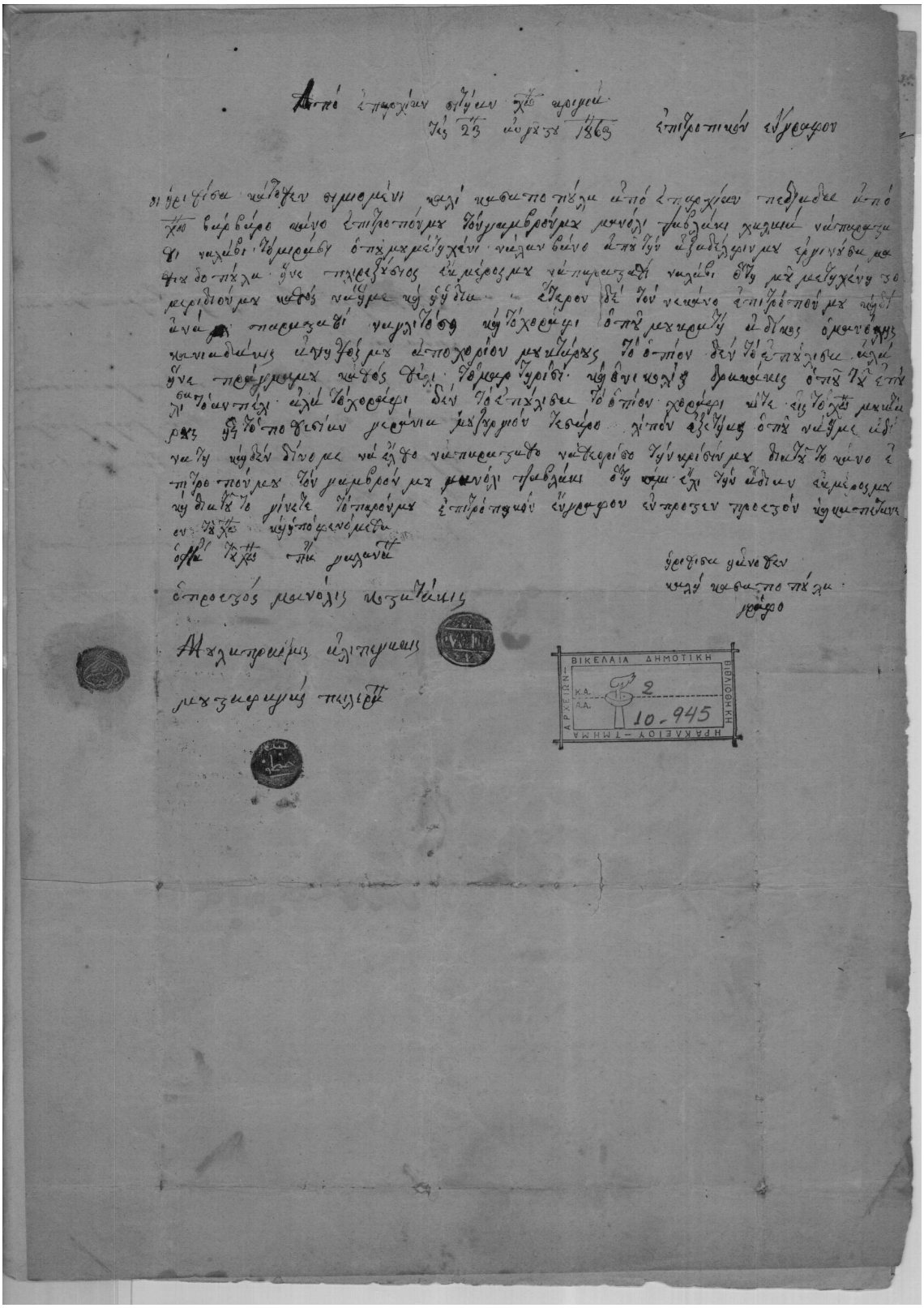
$$\mathrm{I}'(x,y) = \begin{cases} 0 & , \quad I(x,y) < \text{Κατώφλι} \\ 1 & , \quad \mathrm{I}(x,y) \geq \text{Κατώφλι} \end{cases} \qquad \text{για κάθε θέση } (x,y)$$

Πρακτικά, αυτή η μέθοδος προσομοιώνει ένα σαρωτή που δίνει σαν αποτέλεσμα της σάρωσης μια εικόνα με 2 χρώματα, έχοντας το πλεονέκτημα της επιλογής του κατωφλιού σε σχέση με τη συγκεκριμένη εικόνα που επεξεργαζόμαστε.

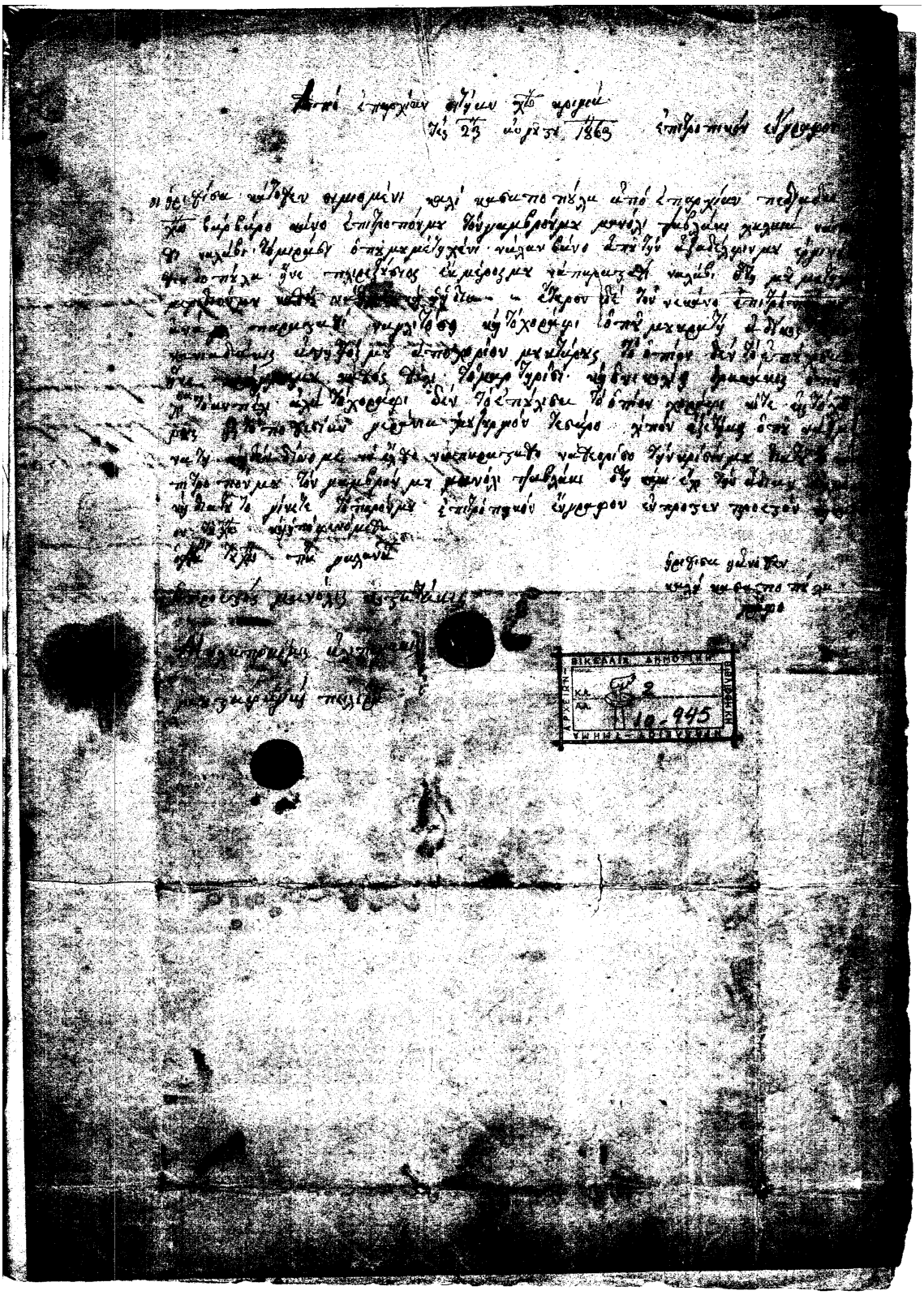
Τα αποτελέσματα, από την εφαρμογή της μεθόδου, όπως φαίνονται και στην εικόνα 3.4 δεν είναι ιδιαίτερα ικανοποιητικά. Το γεγονός αυτό οφείλεται στο ότι η μέθοδος αυτή δεν λαμβάνει υπ'όψιν της τις τοπικές διαφορές που υπάρχουν στην φωτεινότητα από σημείο σε σημείο της ίδιας εικόνας. Χρησιμοποιείται λοιπόν μία τιμή για κατώφλι που δεν είναι ικανή να διαχωρίσει το κείμενο από το χαρτί με καλή απόδοση. Ιδιαίτερα μεγάλο πρόβλημα έχει η μέθοδος στις πολύ σκοτεινές ή πολύ φωτεινές περιοχές. Εκεί οι τιμές των φωτεινοτήτων διαφέρουν αρκετά από την μέση τιμή με αποτέλεσμα να περνούν σχεδόν όλες στην ίδια πλευρά του κατωφλίου. Οπτικά αυτό ανιχνεύεται από την ύπαρξη μεγάλων κομματιών εικονοστοιχείων με μαύρο χρώμα. Πρόβλημα επίσης υπάρχει στη γειτονιά κάθε γράμματος. Εκεί υπάρχουν αρκετά εικονοστοιχεία με φωτεινότητα κοντά σε αυτή του γράμματος παρ'όλο που μπορεί εμείς να αναγνωρίζουμε πως πρόκειται για κομμάτι του χαρτιού. Και στην περίπτωση αυτή η μέθοδος δεν καταφέρνει να ξεχωρίσει καλά τα γράμματα από το χαρτί με αποτέλεσμα γύρω από κάθε γράμμα να υπάρχει αρκετός θόρυβος.

#### 3.1.2.2 Προσαρμοζόμενο κατώφλι

Με βάση τα παραπάνω είναι φανερό ότι μια μέθοδος ενιαίου κατωφλίου είναι καταδικασμένη να αποτύχει. Έτσι, κρίθηκε αναγκαίο να εισάγουμε μεθόδους που βασίζονται




Σχήμα 3.3: Αρχική εικόνα



Σχήμα 3.4: Αποτέλεσμα μεθόδου ενιαίου κατωφλίου

σε προσαρμοζόμενο κατώφλι (adaptive thresholding) [40], [11]. Εδώ το κατώφλι μεταβάλλεται από περιοχή σε περιοχή πάνω στην εικόνα με αποτέλεσμα οι διακυμάνσεις της φωτεινότητας να μην επηρεάζουν την απόδοση. Στην πιο απλή μορφή της μεθόδου το κατώφλι εξακολουθεί να υπολογίζεται με βάση την μέση τιμή αλλά αυτή τη φορά η μέση τιμή προέρχεται από ένα παράθυρο πάνω στην εικόνα. Για ένα παράθυρο διαστάσεων $w \cdot w$, που ξεκινά στη θέση (i, j), η μέθοδος προσαρμοζόμενου κατωφλίου σε μορφή ψευδοκώδικα είναι:

$$\text{Κατώφλι}_{ij} = \frac{\sum_{k=i}^{i+w} \sum_{l=j}^{j+w} I(k,l)}{w \cdot w}$$

$$\text{I}'(x,y) = \begin{cases} 0 \quad , & I(x,y) < \text{Κατώφλι}_{ij} \quad \text{για κάθε θέση } (x,y) \text{ μέσα} \\ 1 \quad , & \text{I}(x,y) \geq \text{Κατώφλι}_{ij} \quad \text{στο τρέχον παράθυρο} \end{cases}$$

και επαναλαμβάνεται ο υπολογισμός του κατωφλίου για κάθε παράθυρο.

Τα αποτελέσματα (Σχ. 3.5) είναι καλύτερα από την απλή μέθοδο του ενιαίου κατωφλίου. Στις γειτονιές γύρω από κάθε γράμμα δεν παρατηρείται πια τόσο μεγάλη ποσότητα θορύβου, και αυτό έρχεται σαν αποτέλεσμα του καλύτερου διαχωρισμού. Εκεί που είναι εμφανές πως η μέθοδος πάσχει είναι στις ομοιόμορφες περιοχές. Στις περιοχές αυτές, και αν υποθέσουμε πως ο θόρυβος στο χαρτί είναι τυχαίος, περιμένουμε πως τα μισά εικονοστοιχεία θα περάσουν "πάνω" από το κατώφλι και τα άλλα μισά "κάτω". Το αποτέλεσμα μοιάζει σαν να έχει επιλεχθεί στην τύχη το αν ένα εικονοστοιχείο θα γίνει άσπρο ή μαύρο.

#### 3.1.2.3 Μέθοδος ελαχίστου-μεγίστου

Ένας τρόπος να αποφύγουμε τα προβλήματα των προαναφερθέντων μεθόδων είναι με χρήση της μεθόδου του ελαχίστου-μεγίστου (min-max). Εδώ το κατώφλι υπολογίζεται σαν η μέση τιμή δύο άλλων τιμών: της ελάχιστης και της μέγιστης σε ένα παράθυρο πάνω στην εικόνα. Αν ε είναι η ελάχιστη, και μ η μέγιστη τιμή σε ένα παράθυρο διαστάσεων $w \cdot w$ τότε η μέθοδος, σε μορφή ψευδοκώδικα, είναι η ακόλουθη:

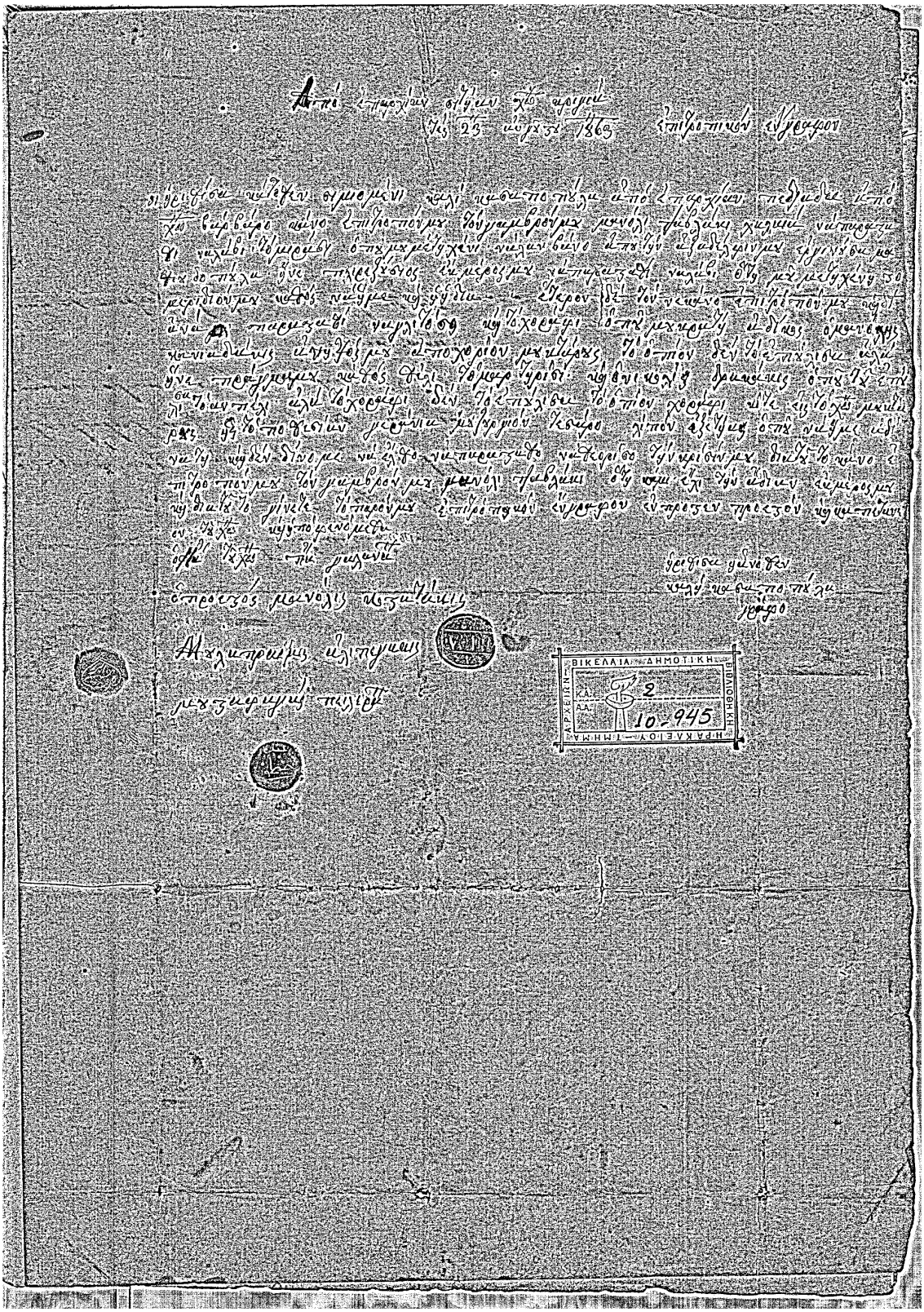
$$\varepsilon_{ij} = \min_{\substack{i \leq x \leq i+w \\ j \leq y \leq j+w}} \{I(x,y)\}, \qquad \mu_{ij} = \max_{\substack{i \leq x \leq i+w \\ j \leq y \leq j+w}} \{I(x,y)\}$$

$$\text{Κατώφλι}_{ij} = \frac{\varepsilon_{ij} + \mu_{ij}}{2}$$

$$\text{I}'(x,y) = \begin{cases} 0 \quad , & I(x,y) < \text{Κατώφλι}_{ij} \quad \text{για κάθε θέση } (x,y) \text{ μέσα} \\ 1 \quad , & \text{I}(x,y) \geq \text{Κατώφλι}_{ij} \quad \text{στο τρέχον παράθυρο} \end{cases}$$

και επαναλαμβάνεται ο υπολογισμός του κατωφλίου για κάθε παράθυρο.

Με τον τρόπο αυτό προσπαθούμε να αποφύγουμε την περίπτωση ενός παραθύρου που περιέχει μία μικρή πλειοψηφία εικονοστοιχείων με μικρές ή μεγάλες τιμές. Στη περίπτωση που υπολογίζαμε την μέση τιμή σε αυτό το παράθυρο θα παίρναμε ένα




Σχήμα 3.5: Αποτέλεσμα μεθόδου προσαρμοζόμενου κατωφλίου

κατώφλι που θα αγνοούσε ουσιαστικά τα εικονοστοιχεία αυτά. Με την μέθοδο ελαχίστου-μεγίστου αυτά τα εικονοστοιχεία δεν αγνοούνται. Έτσι παίρνουμε καλά αποτελέσματα στις περιοχές αυτές. Η δύναμη όμως της μεθόδου αυτής είναι ταυτόχρονα και το αδύνατο σημείο της αφού πάντα μετατοπίζει το κατώφλι προς την μεριά του τοπικού ελαχίστου ή μεγίστου. Και όσο μικρότερη η τιμή του ελαχίστου (αντίστοιχα: μεγαλύτερη η τιμή του μεγίστου) τόσο μεγαλύτερη και η μετατόπιση. Σαν αποτέλεσμα αυτής της συμπεριφοράς παρατηρούμε απώλειες γραμμάτων και εισαγωγή αρκετού θορύβου σε ομοιόμορφες περιοχές (Σχ. 3.6).

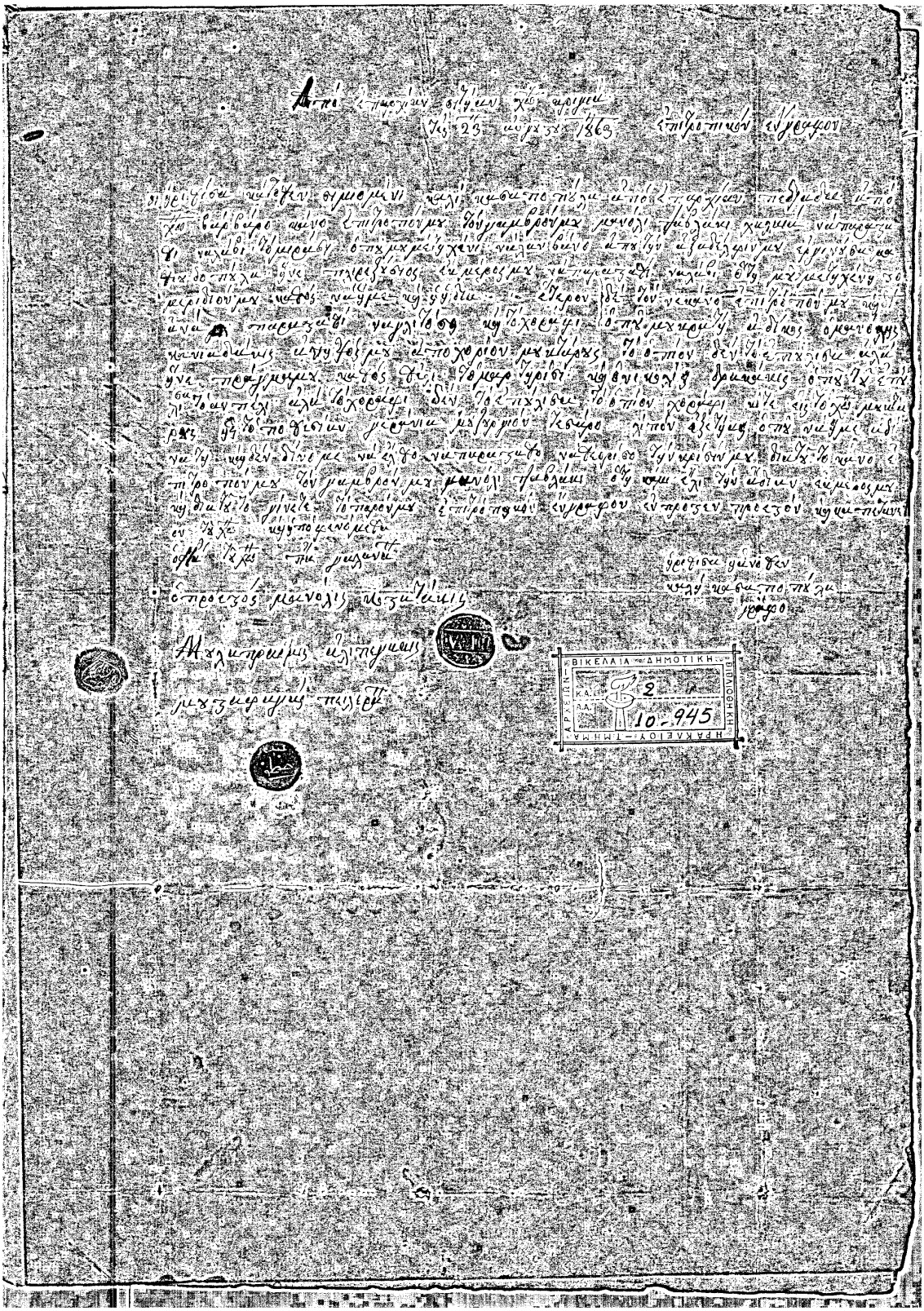
#### 3.1.2.4 Μέθοδος υποβιβασμού κατωφλίου

Παρατηρώντας τα θετικά και τα αρνητικά σημεία των παραπάνω μεθόδων μπορεί κανείς να βγάλει το εύλογο συμπέρασμα πως η μέθοδος που τα καταφέρνει καλύτερα είναι η μέθοδος του προσαρμοζόμενου κατωφλίου, και πως αν καταφέρναμε με κάποιο τρόπο να την μετατρέψουμε ώστε να δουλεύει και στις ομοιόμορφες περιοχές δίνοντας και εκεί καλά αποτελέσματα θα είχαμε πετύχει το σκοπό μας. Επικεντρώσαμε τις προσπάθειές μας στο σημείο αυτό και πετύχαμε κάποια αρκετά ικανοποιητικά αποτελέσματα.

Όπως αναφέρθηκε και προηγουμένως η μέθοδος του προσαρμοζόμενου κατωφλίου δίνει καλά αποτελέσματα στις περιοχές γύρω από τα γράμματα ενώ στις ομοιόμορφες περιοχές (όπως αυτές του χαρτιού) δεν δουλεύει ικανοποιητικά. Είναι προφανές πως όλες αυτές οι ομοιόμορφες περιοχές ανήκουν στο χαρτί και επομένως θα έπρεπε να γίνουν λευκές. Ας υποθέσουμε τώρα ότι έχουμε μια ομοιόμορφη περιοχή. Στην περιοχή αυτή η μέση τιμή των φωτεινοτήτων θα είναι περίπου σταθερή. Επειδή όμως οι φωτεινότητες έχουν κοντινές τιμές κάποιες από αυτές θα είναι μεγαλύτερες και κάποιες μικρότερες από την μέση τιμή, που είναι και η τιμή του κατωφλίου μας. Αν όμως αντί να κρατούσαμε αυτή την τιμή κρατούσαμε μια τιμή ελαφρώς μικρότερη από τις τιμές των γειτονικών εικονοστοιχείων τότε τα περισσότερα από αυτά θα οδηγούνταν στο λευκό. Η τιμή, βέβαια, που θα έπρεπε να κρατήσουμε δεν θα έπρεπε να απέχει και πολύ από την μέση τιμή γιατί τότε πιθανόν να χάναμε και πολύτιμη πληροφορία (εξαφανίζοντας π.χ. κάποιο κομμάτι από ένα γράμμα γραμμένο με αχνή μελάνη). Επίσης το πόσο θα απέχουμε από την μέση τιμή δεν θα πρέπει να είναι σταθερό γιατί σε αυτήν την περίπτωση θα έχουμε προβλήματα λόγω των διαφορετικών φωτεινοτήτων που υπάρχουν στην εικόνα σε διάφορα σημεία τους.

Ο τρόπος με τον οποίο επιτύχαμε όλους τους πιο πάνω στόχους είναι με το να χρησιμοποιούμε λιγότερες τιμές από το παράθυρο στο οποίο δουλεύουμε κάθε φορά, δηλαδή δεν χρησιμοποιούμε μερικές τιμές από τα άκρα του. Έτσι αντί να υπολογίζουμε το κατώφλι σε ένα παράθυρο μεγέθους $w \cdot w$ με τον τρόπο που το κάνουμε στη μέθοδο του προσαρμοζόμενου κατωφλίου, χρησιμοποιούμε τη σχέση

$$\text{Κατώφλι}_{ij} = \frac{\sum_{k=i+a}^{i+w-b} \sum_{l=j+c}^{j+w-d} I(k,l)}{w \cdot w}$$




Σχήμα 3.6: Αποτέλεσμα μεθόδου ελαχίστου-μεγίστου

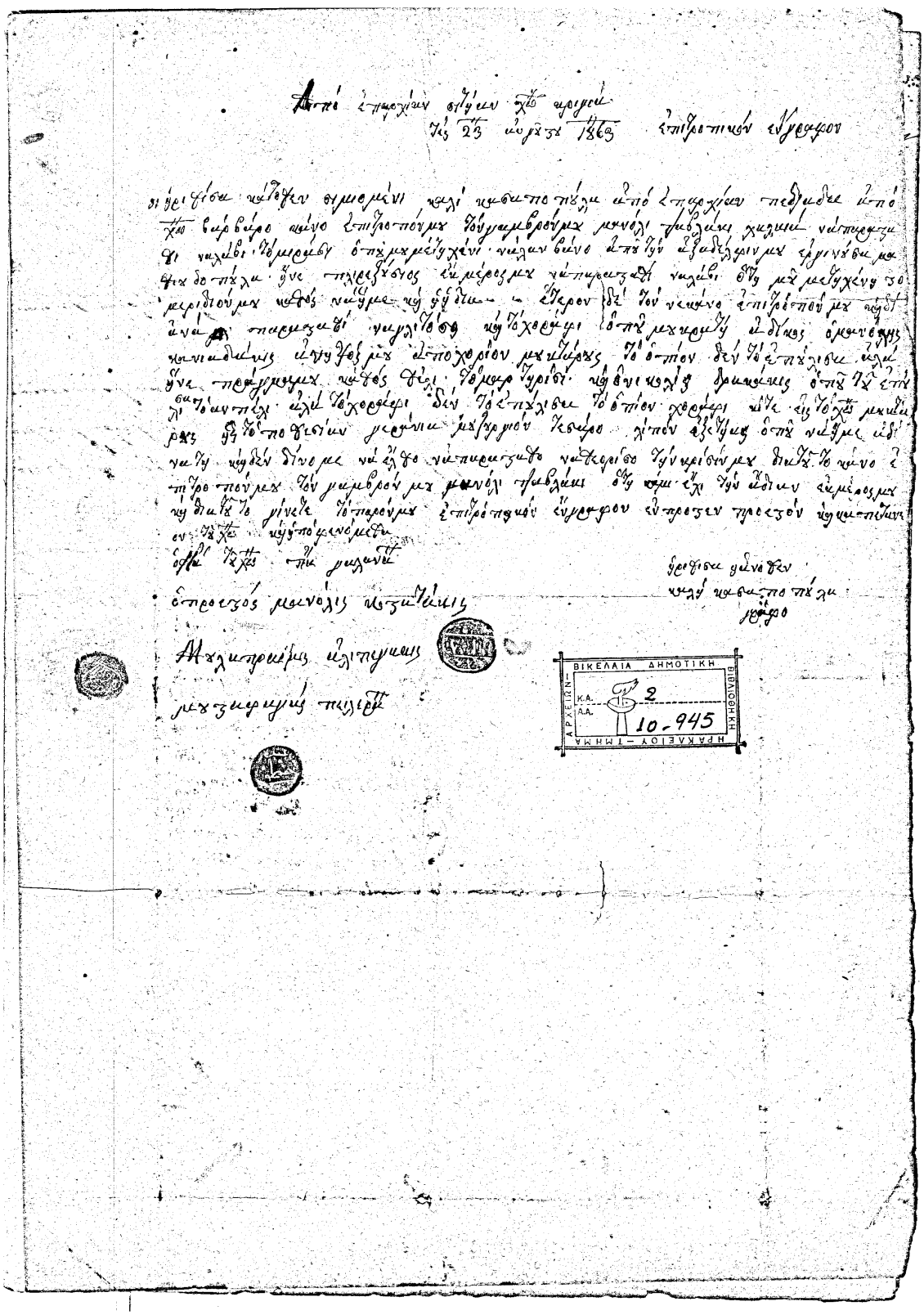
με $a > 1$, $b < w$, $c > 1$ και $d < w$, ενώ επίσης πρέπει να ισχύουν οι περιορισμοί $a < w - b$ και $c < w - d$ για να μην έχουμε προβλήματα στα όρια άθροισης. Ο έλεγχος για την τιμή της νέας φωτεινότητας για κάθε θέση του παραθύρου, και για κάθε παράθυρο, γίνεται ακριβώς όπως και στη μέθοδο προσαρμοζόμενου κατωφλίου.

Με τη χρήση των παραμέτρων $a, b, c,$ και $d$ πετυχαίνουμε μείωση της μέσης τιμής (υποβιβασμό τιμής κατωφλίου), αφού ουσιαστικά απορρίπτουμε κάποια σημεία μη λαμβάνοντάς τα υπ'όψιν. Με τη μείωση της μέσης τιμής οδηγούμε τα περισσότερα από τα εικονοστοιχεία της γύρω περιοχής στο λευκό και δεν απομακρυνόμαστε αρκετά από την μέση τιμή. Το πόσο θα απομακρυνθούμε εξαρτάται από την επιλογή των $a, b, c$ και $d$. Μια αρκετά καλή επιλογή, είναι τέτοιες τιμές των παραμέτρων ώστε να κρατούμε τουλάχιστον τα $\frac{2}{3}$ του πλήθους των τιμών του παραθύρου.

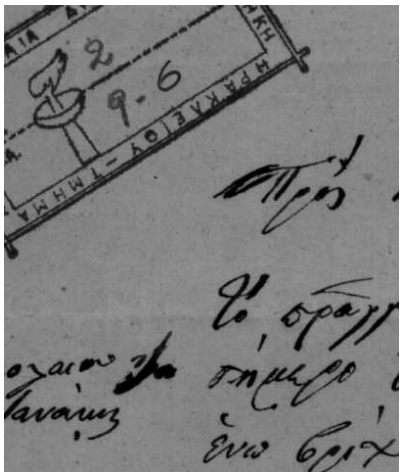
Αν και η μέθοδος μοιάζει αρκετά με τη μέθοδο του προσαρμοζόμενου κατωφλίου έχει εντούτοις μια ουσιώδη διαφορά: το κατώφλι που έμμεσα περιέχει δεν είναι μοναδικά ορισμένο για όλη την εικόνα αλλά προσαρμόζεται σε κάθε παράθυρο ακολουθώντας καλύτερα τις διακυμάνσεις της φωτεινότητας.

Τα αποτελέσματα της μεθόδου είναι αρκετά ικανοποιητικά όπως μπορούμε να δούμε και από το Σχ. 3.7. Ο θόρυβος έχει φύγει από τις ομοιόμορφες περιοχές ενώ τα γράμματα και η σφραγίδα έχουν διατηρηθεί απείραχτα. Στο σημείο αυτό θα θέλαμε να επισημάνουμε για μια ακόμη φορά ότι με την εφαρμογή της παραπάνω μεθόδου προσαρμοζόμενου κατωφλίου πετύχαμε να μετατρέψουμε τις εικόνες σε ασπρόμαυρες διατηρώντας τόσο την ποιότητα των εικόνων, όσο και (κυρίως) τη χρήσιμη πληροφορία των εγγράφων, όπως αυτή έχει οριστεί παραπάνω. Είναι προφανές ότι δεν διατηρείται το πιστό πρωτότυπο αλλά αυτός είναι ένας απαραίτητος συμβιβασμός προκειμένου, όπως έχουμε εξηγήσει, να έχουμε ένα διαχειρίσιμο όγκο αρχείων στη βάση μας.

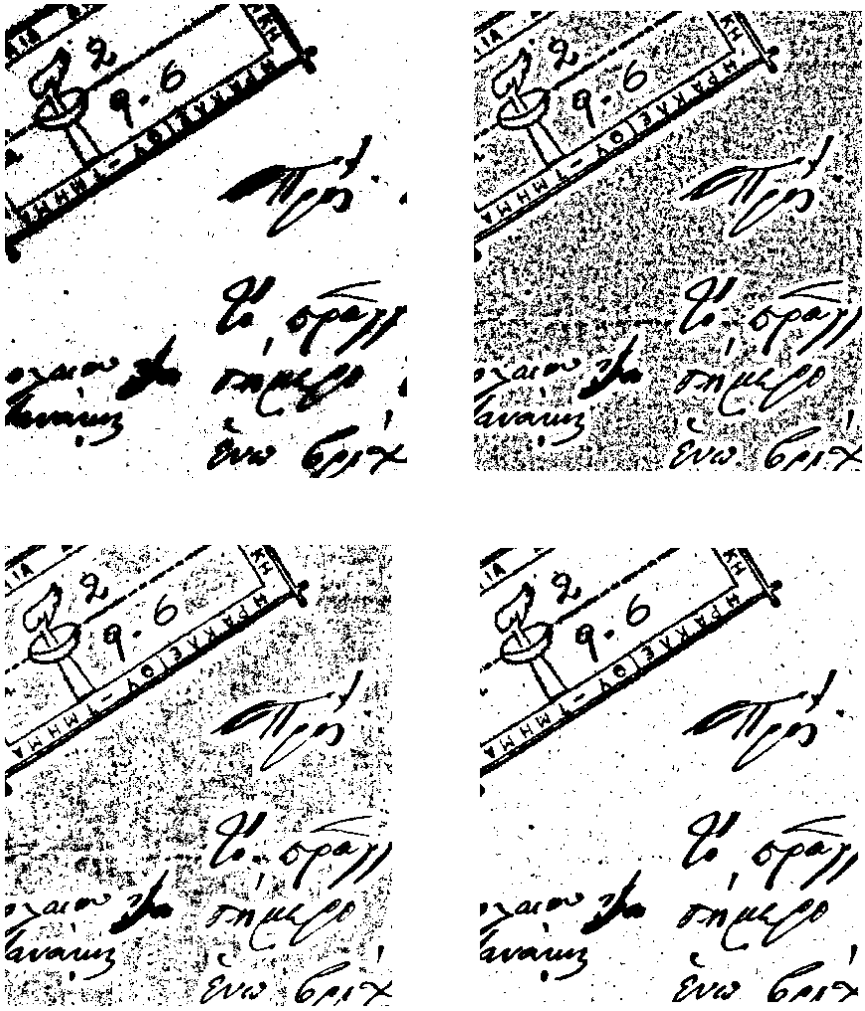
Στα σχήματα 3.8 και 3.9 έχουμε τα αποτελέσματα όλων των παραπάνω μεθόδων σε ένα τμήμα μιας εικόνας για να δειχθούν καλύτερα οι λεπτομέρειες που δεν φαίνονται στις προηγούμενες εικόνες λόγω του μεγάλου μεγέθους τους. Παρατηρούμε ότι με τη μέθοδο του ενιαίου κατωφλίου η εικόνα είναι σχετικά καλής ποιότητας αλλά αυτό συμβαίνει στη συγκεκριμένη περίπτωση μόνο μιας και η αρχική έχει αρκετά ομοιόμορφες αποχρώσεις γκρίζου. Παρ'όλα αυτά τα γράμματα και περιοχές στη σφραγίδα έχουν ενωθεί σαν αποτέλεσμα μη ικανοποιητικού διαχωρισμού. Με την μέθοδο του προσαρμοζόμενου κατωφλίου τα γράμματα και η σφραγίδα δίνονται αρκετά ικανοποιητικά αλλά προστίθεται αρκετός θόρυβος εκεί όπου θα έπρεπε να έχουμε λευκό (τμήμα που ανήκει στο χαρτί). Η μέθοδος ελαχίστου-μεγίστου δίνει ελαφρώς λιγότερο θόρυβο αλλά αρχίζει να εξαφανίζει γράμματα (τα διαβρώνει) αφού στις οριακές περιοχές μετάβασης από χαρτί σε γράμμα η συμπεριφορά της μεθόδου εξαρτάται πολύ από τις τιμές που έχουν οι περιοχές αυτές. Με τη δική μας μέθοδο το αποτέλεσμα είναι αρκετά καλύτερο αφού έχουμε την πληροφορία που θέλουμε με αρκετά λιγότερο θόρυβο.




Σχήμα 3.7: Αποτέλεσμα μεθόδου υποβιβασμού κατωφλίου



Σχήμα 3.8: Ένα τμήμα εικόνας σε επίπεδα γκρίζου



α β
γ δ

Σχήμα 3.9: Αποτελέσματα στο τμήμα εικόνας των μεθόδων α) ενιαίου κατωφλίου, β) προσαρμοζόμενου κατωφλίου, γ) ελαχίστου-μεγίστου και δ) υποβιβασμού κατωφλίου

| Πλήθος εικόνων | 306 |
|---|---|
| Μέση τιμή | 6,307189 |
| Τυπική απόκλιση | 7,371364 |

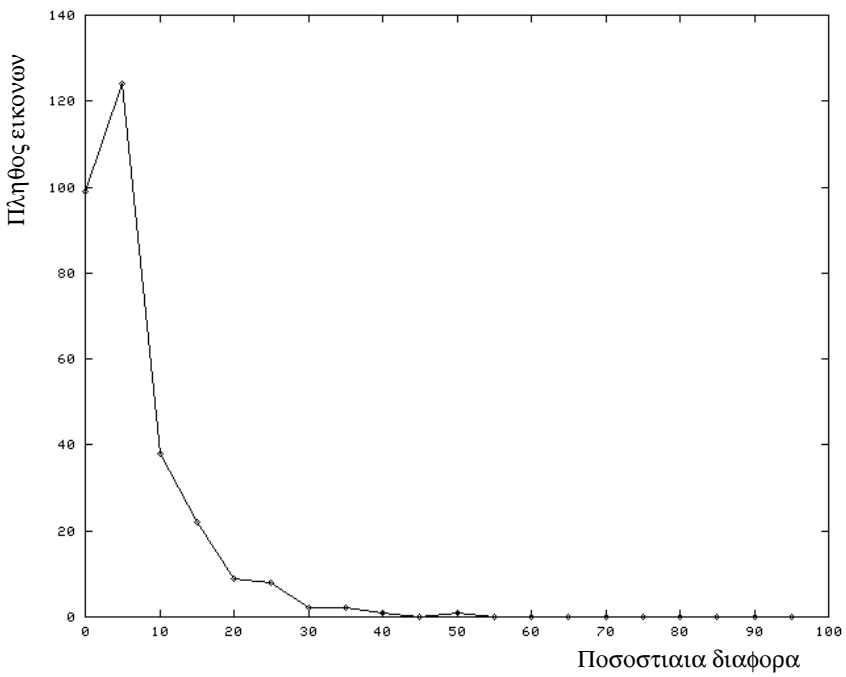
Πίνακας 3.2: Αριθμητικά μεγέθη της σύγκρισης

### 3.1.3 Πειραματική αποτίμηση μεθόδου

Για να γίνει μια ποσοτική εκτίμηση της αποδοτικότητας της προτεινόμενης μεθόδου έγινε το εξής πείραμα: Σε ένα πλήθος εικόνων χρησιμοποιήσαμε τη δική μας μέθοδο για μετατροπή από τα 256 χρώματα (επίπεδα γκρίζου) σε ασπρόμαυρες εικόνες 2 χρωμάτων. Την κάθε ασπρόμαυρη εικόνα που προέκυψε σαν αποτέλεσμα εξόδου της μεθόδου μας, εικόνα Α, τη χρησιμοποιήσαμε σαν εικόνα αναφοράς για την μέθοδο ενιαίου κατωφλίου. Ξεκινώντας από μια χαμηλή τιμή για το κατώφλι της μεθόδου του ενιαίου κατωφλίου, αυξάναμε την τιμή συνεχώς έως ότου να κρατήσουμε στη νέα εικόνα, εικόνα που θα προκύψει με τη μέθοδο του ενιαίου κατωφλίου, εικόνα Β, το 95% της πληροφορίας που περιέχεται στην εικόνα Α. Σαν πληροφορία θεωρούμε τα σημεία εκείνα της εικόνας Α που έχουν μείνει μαύρα με τη μέθοδό μας και που ως επί το πλείστον περιέχουν γράμματα και σφραγίδες. Όταν φτάσουμε να διατηρούμε το ποσό της πληροφορίας που θέλουμε, υπολογίζουμε το πλήθος των σημείων στα οποία οι δύο εικόνες διαφέρουν (έχει η μια τιμή χρώματος άσπρο και η άλλη μαύρο στο ίδιο σημείο).

Το πείραμα έγινε σε ένα πλήθος 306 εικόνων αντιπροσωπευτικών των διαφόρων ειδών εγγράφων που υπάρχουν στη Βικελαία. Το πλήθος αυτό είναι ικανό να δώσει μια καλή εικόνα της συμπεριφοράς της μεθόδου, χωρίς να υπάρχει το ενδεχόμενο της συμπτωματικής επιλογής καλών ή κακών εγγράφων. Η απαίτηση για διατήρηση του 95% της πληροφορίας εισήχθηκε για να εξασφαλιστεί ότι και οι δύο εικόνες, Α και Β, θα έχουν περίπου την ίδια ποσότητα χρήσιμης πληροφορίας.

Τα αποτελέσματα του πειράματος φαίνονται στο Σχ. 3.10 όπου έχει γίνει γραφική παράσταση του πλήθους των εικόνων, στον κατακόρυφο άξονα y, που έχουν μια ορισμένη τιμή ποσοστιαίας διαφοράς, στον οριζόντιο άξονα x. Η τιμή της ποσοστιαίας διαφοράς αναφέρεται στη διαφορά μεταξύ των εικόνων Α και Β σαν ποσοστό των εικονοστοιχείων των δύο εικόνων που διαφέρουν στο χρώμα, σε σχέση με το συνολικό πλήθος εικονοστοιχείων των δύο εικόνων. Οι τιμές ποσοστιαίας διαφοράς δίνονται σε βήματα της τάξεως του 5% η μία από την άλλη αφού δεν μας ενδιαφέρει η ακριβής τιμή της διαφοράς των δύο εικόνων αλλά η συμπεριφορά της μεθόδου σε γενικές γραμμές. Βλέπουμε λοιπόν, ότι η μέθοδός μας παρουσιάζει αρκετά καλύτερες επιδόσεις, κάτι το οποίο φαινόταν και οπτικά. Η μέση τιμή της διαφοράς των δύο μεθόδων είναι της τάξης του 6,3% (βλέπε και πίνακα 3.2), το οποίο μπορεί να μην ακούγεται μεγάλο σαν νούμερο, αλλά η διαφορά στην εμφάνιση δύο εικόνων με τέτοια απόσταση είναι πολύ σημαντική.



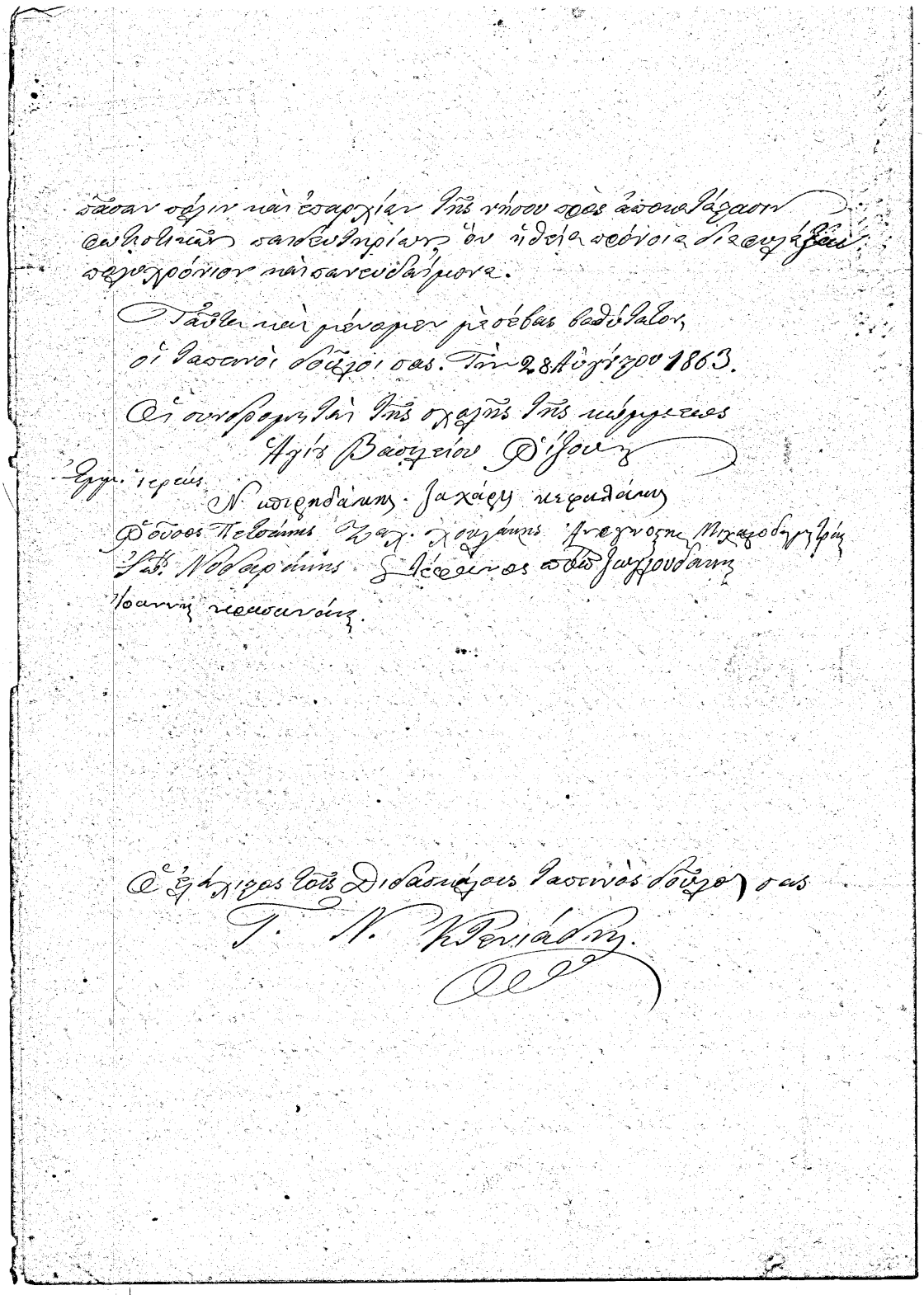

Σχήμα 3.10: Γραφική παράσταση της διαφοράς των δύο μεθόδων

Για την οπτική επαλήθευση του παραπάνω ισχυρισμού στο Σχ. 3.11 απεικονίζεται μια εικόνα που έχει προκύψει από τη δική μας μέθοδο ελλάτωσης χρωμάτων, ενώ στο Σχ. 3.12 φαίνεται η αντίστοιχη εικόνα που έχει προκύψει από τη μέθοδο του ενιαίου κατωφλίου και που διαφέρει από τη δική μας κατά 6%.

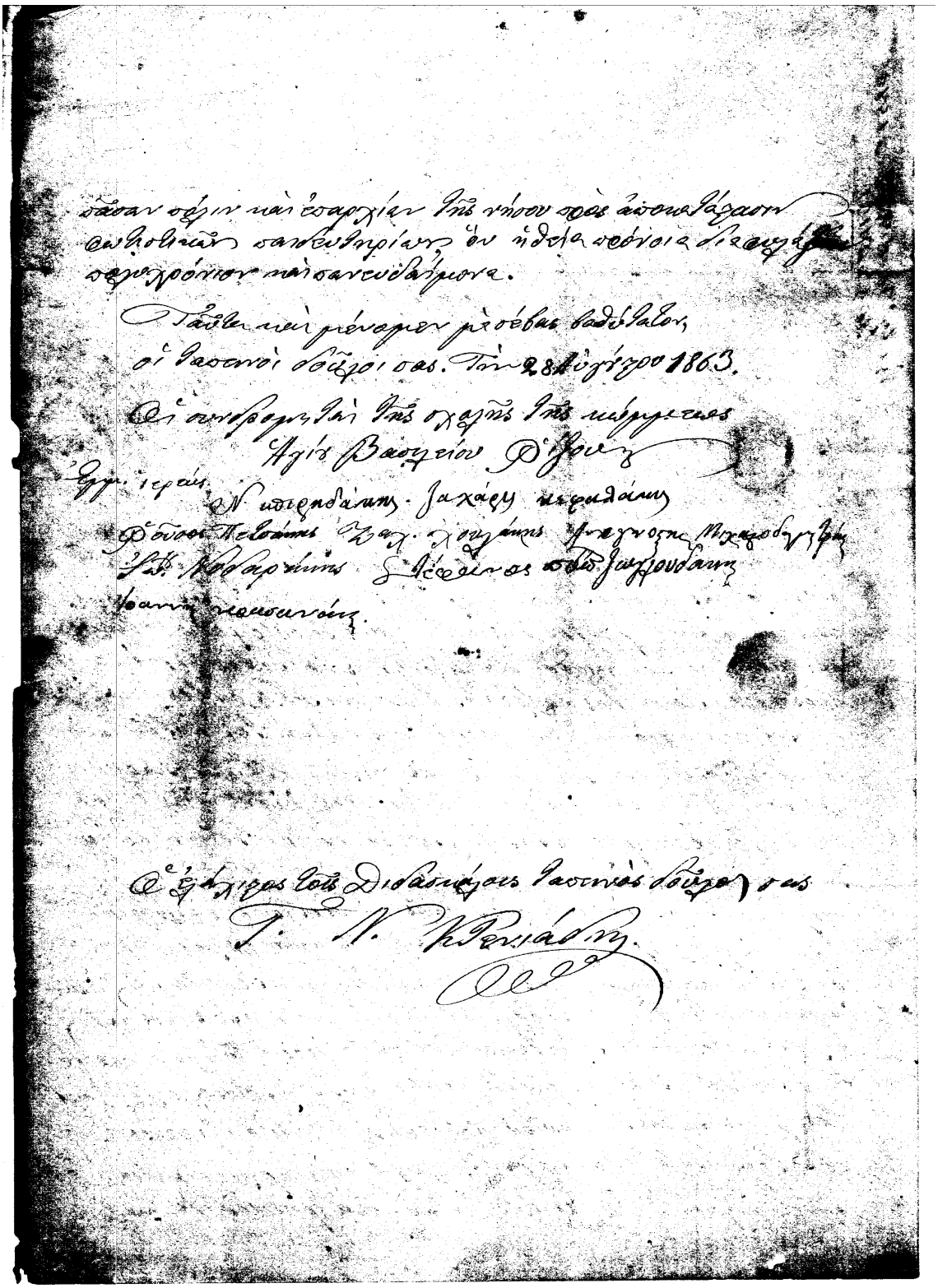
### 3.1.4 Καθαρισμός εικόνας

Μετά το στάδιο της μετατροπής της εικόνας σε δυαδική ακολουθεί ένα στάδιο καθαρισμού της εικόνας. Ο καθαρισμός αυτός είναι απαραίτητος προκειμένου να βελτιωθεί περισσότερο η ποιότητα της εικόνας που έχει προκύψει. Στο στάδιο αυτό αφαιρείται ο θόρυβος που έχει παραμείνει και που προέρχεται από παράγοντες όπως ανωμαλίες στην επιφάνεια του χαρτιού που στο στάδιο της σάρωσης μας δώσαν λανθασμένη τιμή στη φωτεινότητα, μελάνια που έχουν ποτίσει το χαρτί και αλλάζουν την φωτεινότητα των γραμμάτων και διάφορους άλλους λόγους.

Ο μεγάλος κίνδυνος με κάθε διαδικασία τέτοιου είδους είναι το να χαθεί πολύτιμη πληροφορία. Υπάρχει δηλαδή περίπτωση, αν δεν προσέξει κανείς, να αφαιρέσει τμήματα από γράμματα, κάποια κομμάτια από μια γραμμή (π.χ. υπογραφής) ή σημεία στίξης όπως τόνους, κόμματα και τελείες. Αφαίρεση κάποιων τέτοιων τμημάτων από την εικόνα θα μπορούσε να προκαλέσει αλλοίωση του νοήματος που σε μερικές περιπτώσεις θα μπορούσε να είναι αρκετά σημαντική. Έτσι ακολουθήθηκε μια ασφαλής προσέγγιση στη διαδικασία καθαρισμού, προσπαθώντας να αφαιρέσουμε μόνο περιοχές που είμαστε σίγουροι ότι περιέχουν τμήματα που προέρχονται από θόρυβο, και αφήνοντας κάποιες άλλες για τις οποίες δεν μπορούμε να είμαστε σίγουροι



Σχήμα 3.11: Εικόνα που προέρχεται από τη μέθοδο υποβιβασμού κατωφλίου



Σχήμα 3.12: Εικόνα που προέρχεται από τη μέθοδο ενιαίου κατωφλίου με διαφορά 6%, από αυτήν της μεθόδου υποβιβασμού κατωφλίου

για το περιεχόμενό τους. Προτιμήσαμε δηλαδή, να αφήσουμε μεγαλύτερη ποσότητα θορύβου στην εικόνα από το να αφαιρέσουμε πολύτιμη πληροφορία. Η προσέγγιση αυτή έγινε και για έναν ακόμη λόγο: ότι πάντα υπάρχει, μελλοντικά, η δυνατότητα ανάπτυξης μιας πιο έξυπνης διαδικασίας καθαρισμού που θα έχει, πιθανόν, ακόμη καλύτερα αποτελέσματα. Μια τέτοια μέθοδος δεν θα ήταν εύκολο να αναπτυχθεί αν η δική μας είχε καταστρέψει όλη τη χρήσιμη πληροφορία. Η επεξεργασία των εικόνων δεν επηρεάζει σημαντικά τη λειτουργία του συστήματος, μιας και μπορεί να γίνεται εκτός της κανονικής λειτουργίας του. Η προσέγγιση αυτή επομένως, συνδυάζει την ασφάλεια της πληροφορίας και ταυτόχρονα την ύπαρξη δυνατότητας διόρθωσης του καθαρισμού.

Η διαδικασία που αναπτύχθηκε για τον καθαρισμό των εικόνων βασίζεται σε δύο απλούς κανόνες:

- Αν μια μαύρη περιοχή είναι μεγαλύτερη από κάποιο προεπιλεγμένο μέγεθος μην την αφαιρείς. Με αυτό τον κανόνα δεν αφαιρούμε τις περιοχές που περιέχουν την ουσιαστική πληροφορία. Το προεπιλεγμένο μέγεθος είναι τέτοιο που να εξασφαλίζει πως δομικά στοιχεία του κειμένου, όπως σημεία στίξης, δεν θα αφαιρεθούν.

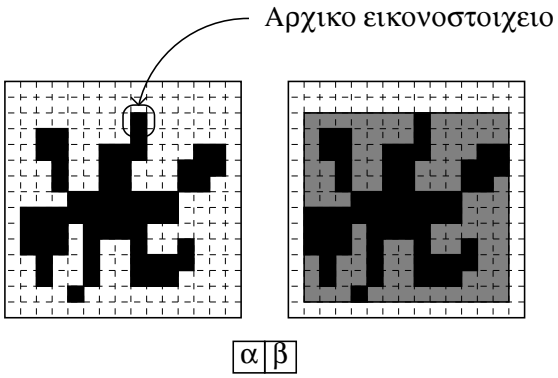
- Αν μια μαύρη περιοχή με μέγεθος μικρότερο από το προεπιλεγμένο κρίσιμο μέγεθος περιβάλλεται από μια ζώνη λευκών εικονοστοιχείων, ενός επίσης προαποφασισμένου μεγέθους, τότε και μόνο τότε αφαίρεσε την περιοχή. Η απαίτηση της ύπαρξης μιας τέτοιας λευκής ζώνης γύρω από την μαύρη περιοχή πηγάζει από την επιθυμία μας να μην διώξουμε περιοχές οι οποίες βρίσκονται κοντά η μία στην άλλη και πιθανόν να αποτελούν τμήματα του ίδιου σχηματισμού.

Με βάση αυτούς τους δύο κανόνες υλοποιήθηκε μια μέθοδος που ο αλγόριθμός της σε ψευδοκώδικα είναι ο ακόλουθος:

- Για κάθε μαύρο εικονοστοιχείο πάνω στην εικόνα, το οποίο συναντούμε κατά τη διαδοχική σάρωση της εικόνας από πάνω προς τα κάτω και από αριστερά προς τα δεξιά

    1. Επέκτεινε την περιοχή αυτή μεγαλώνοντας τα όριά της προς δεξιά, αριστερά, πάνω και κάτω
    2. Αν το μέγεθός της ξεπεράσει ένα προαποφασισμένο όριο για τη μαύρη περιοχή προχώρα στην επόμενη περιοχή
    3. Αλλιώς βρες αν περιβάλλεται από μια λευκή περιοχή ενός προκαθορισμένου μεγέθους
        - Αν όχι τότε προχώρα με την επόμενη μαύρη περιοχή
        - Αν ναι τότε αφαίρεσε την περιοχή αυτή

- Επανέλαβε την παραπάνω διαδικασία μέχρι να μην έχεις καμία αλλαγή

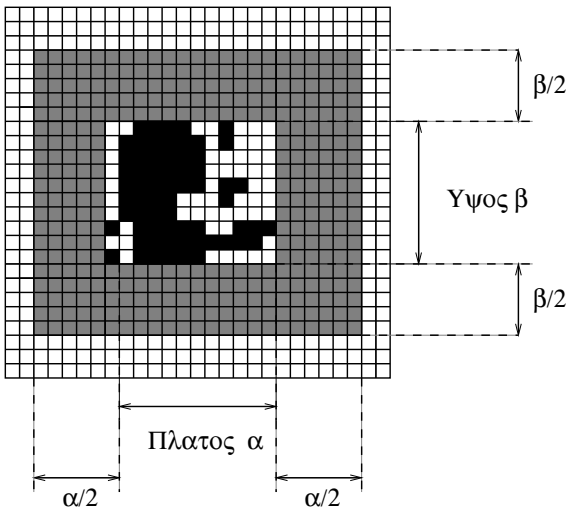



Σχήμα 3.13: Επέκταση περιοχής

Η επέκταση της περιοχής γίνεται έτσι ώστε τελικά να βρεθεί κλεισμένη μέσα σε κάποιο ορθογώνιο παραλληλόγραμμο (ουσιαστικά βρίσκουμε το ελάχιστο ορθογώνιο παραλληλόγραμμο που περιέχει ολόκληρη την μαύρη περιοχή). Έτσι αν π.χ. ξεκινούσαμε από την κατάσταση που φαίνεται στο Σχ. 3.13α με αρχικό εικονοστοιχείο αυτό που περικλείεται στον κύκλο, θα καταλήγαμε σε αυτή του Σχ. 3.13β, όπου όλη η μαύρη περιοχή περικλείεται από ένα ορθογώνιο παραλληλόγραμμο. Με γκρίζο φαίνεται το εσωτερικό του ορθογωνίου που περικλείει την περιοχή μαύρων εικονοστοιχείων. Για να πετύχει κάτι τέτοιο η μέθοδος μας επεκτείνει συνεχώς την κάθε πλευρά του αρχικού ορθογωνίου (στην αρχή είναι ένα τετράγωνο 1x1 που περικλείει μόνο το αρχικό εικονοστοιχείο) έως ότου να καταφέρει να καλύψει όλη τη μαύρη περιοχή.

Αν τώρα η περιοχή αυτή έχει ένα αποδεκτό μέγεθος για περιοχή μαύρων εικονοστοιχείων και περιβάλλεται από μια λευκή περιοχή κάποιου μεγέθους τότε μπορεί να αφαιρεθεί από την αρχική εικόνα. Στην υλοποίηση που έγινε στα πλαίσια του προγράμματος ΑΡΧΟΝ βρήκαμε πειραματικά ότι δεν πρέπει να αφαιρούνται περιοχές που έχουν πλάτος ή ύψος λευκής περιβάλλουσας ζώνης μικρότερο από το μισό της μαύρης περιοχής που έχουμε εντοπίσει (στο Σχ. 3.14 η ζώνη με λευκά εικονοστοιχεία που απαιτείται φαίνεται με το γκρίζο χρώμα). Έτσι αν η περιοχή με μαύρα εικονοστοιχεία έχει διαστάσεις 10x12 θα πρέπει να περιβάλλεται από μια ζώνη με λευκά εικονοστοιχεία που θα έχει δεξιά και αριστερά τουλάχιστον $\frac{10}{2}$=5 σειρές λευκών εικονοστοιχείων και από πάνω και από κάτω τουλάχιστον $\frac{12}{2}$=6 σειρές. Η απόφαση αυτή πάρθηκε έχοντας υπ'όψιν ότι με μικρότερο μέγεθος περιβάλλουσας ζώνης είναι πολύ πιθανόν ότι ο σχηματισμός αυτός της μαύρης περιοχής, θα είναι τμήμα ενός άλλου σχηματισμού και δεν θα έπρεπε να αφαιρεθεί.

Ένα παράδειγμα εφαρμογής της διαδικασίας καθαρισμού φαίνεται στο Σχ. 3.15 όπου αντιπαραβάλλεται η εικόνα πριν και μετά την διαδικασία. Στο Σχ. 3.16 φαίνονται τα αποτελέσματα της διαδικασίας για το μεγενθυμένο τμήμα μιας εικόνας όπου φαίνεται, σε αντιπαραβολή με την εικόνα του Σχ. 3.9δ, η απομάκρυνση όλων εκείνων των ανεπιθύμητων κουκίδων που προέρχονται από θόρυβο, με ταυτόχρονη διατήρηση όλων των υπόλοιπων λεπτομερειών σημαντικής πληροφορίας.



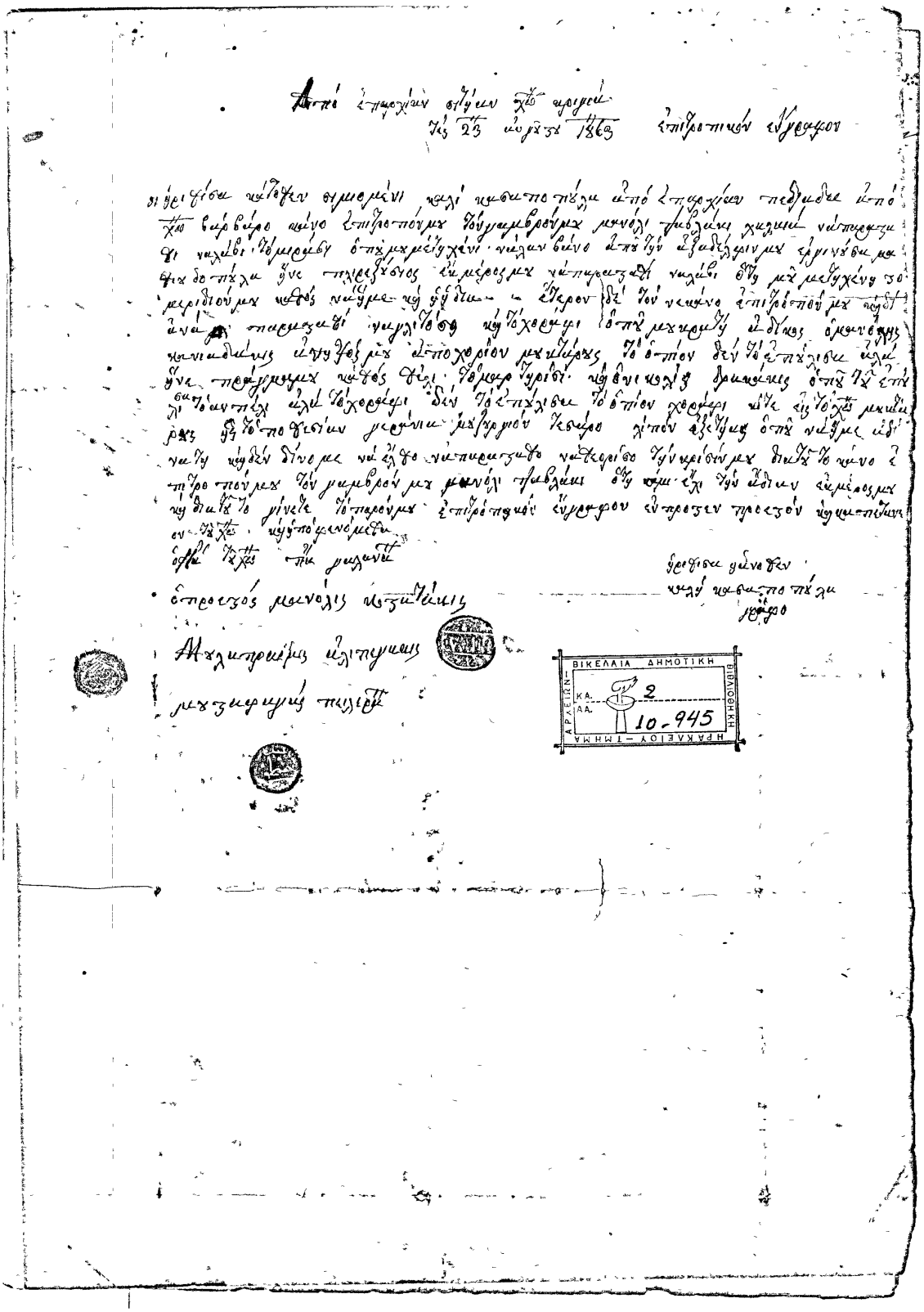

Σχήμα 3.14: Διαστάσεις λευκής ζώνης για την αφαίρεση μιας περιοχής

Όπως φαίνεται και από το ψευδοκώδικα η μέθοδος είναι επαναληπτική και εφαρμόζεται συνεχώς μέχρι να μην υπάρξει καμία αλλαγή. Αυτή η επαναληπτικότητα κρίθηκε αναγκαία επειδή δεν είναι δεδομένο ότι με ένα μόνο πέρασμα θα καθαρίσουν όλες οι περιοχές. Σαν παράδειγμα μπορούμε να δούμε την περίπτωση του Σχ. 3.17 όπου οι περιοχές α και γ δεν μπορούν να φύγουν πριν φύγει η περιοχή β. Αυτό συμβαίνει διότι η περιοχές α και γ απαιτούν για την απομάκρυνσή τους περιβάλλουσα λευκή ζώνη μεγαλύτερων διαστάσεων από αυτή που έχουν όταν είναι παρούσα η περιοχή β. Μετά την απομάκρυνση της β, που μπορεί να φύγει αφού έχει ζώνη λευκών εικονοστοιχείων με τις απαραίτητες διαστάσεις, στο επόμενο πέρασμα οι α και γ έχουν με τη σειρά τους την απαραίτητη λευκή περιβάλλουσα ζώνη, οπότε και μπορούν να απομακρυνθούν.

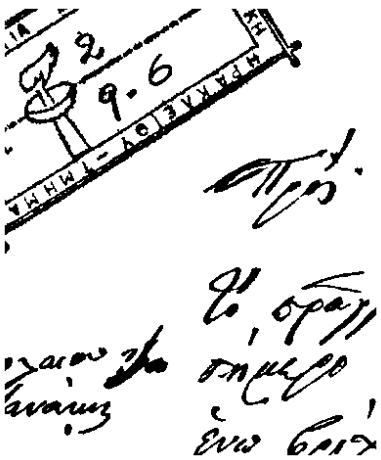
Η διαδικασία καθαρισμού εκτός του ότι βελτιώνει την εμφάνιση των εικόνων είναι αρκετά χρήσιμη και για έναν άλλο λόγο που θα γίνει πιο προφανής στη συνέχεια: βοηθά την διαδικασία της συμπίεσης να πετύχει ακόμη καλύτερες επιδόσεις και επομένως να ελλατώσει ακόμη περισσότερο το μέγεθος των τελικών αρχείων στο σύστημα.

### 3.1.5 Διαδικασία κλεισίματος

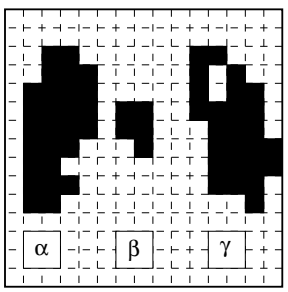
Στη συνέχεια δοκιμάστηκε μια μέθοδος που σαν σκοπό της είχε την περεταίρω βελτίωση των εικόνων. Αν κοιτάξει κανείς προσεκτικά τα γράμματα σε αρκετή μεγέθυνση θα δει πως αυτά μετά από τις παραπάνω διαδικασίες δείχνουν να είναι ελαφρώς "φαγωμένα". Μοιάζουν δηλαδή, σαν να τους έχουν αφαιρεθεί τμήματα που κάποιος θα περίμενε πως θα έπρεπε να υπάρχουν. Για παράδειγμα στην εικόνα 3.18β τόσο το γράμμα δ όσο και το γράμμα γ της φράσης "δεν γνοριζο", φαίνονται σαν να είναι "σπασμένα", χωρίς να έχουν συνοχή και σαν να τους λείπουν ορισμένα κομμάτια των γραμμών που τα αποτελούν. Το πρόβλημα προέκυψε διότι στην αντίστοιχη εικόνα επιπέδων γκρίζου (εικόνα 3.18α) οι φωτεινότητες των περιοχών αυτών αποτυπώθηκαν κατά τη διαδικασία της σάρωσης πολύ λανθασμένα, με αποτέλεσμα κατά τη μετατροπή σε



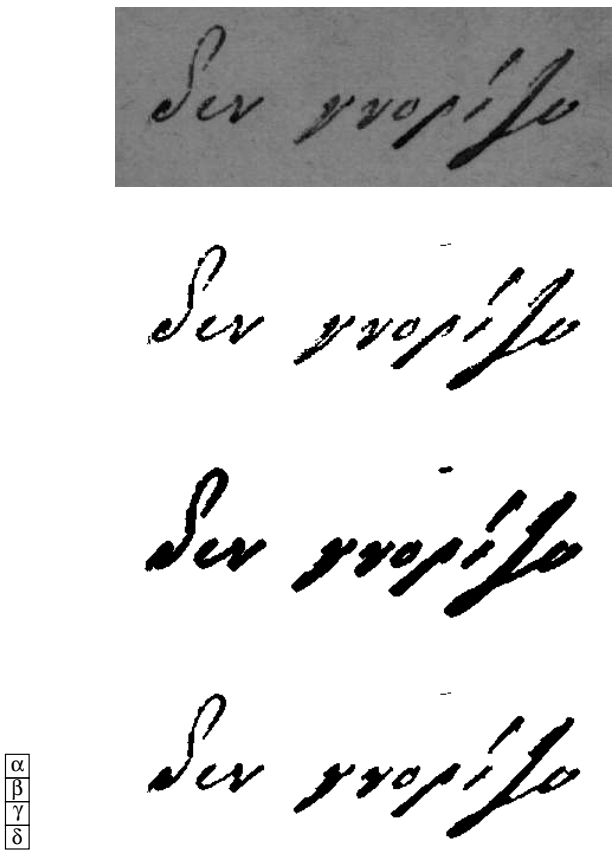

Σχήμα 3.15: Αποτέλεσμα της διαδικασίας καθαρισμού για την εικόνα του Σχ. 3.7



Σχήμα 3.16: Αποτέλεσμα της διαδικασίας καθαρισμού για το τμήμα εικόνας του Σχ. 3.9δ




Σχήμα 3.17: Ανάγκη επαναληπτικότητας της διαδικασίας καθαρισμού




Σχήμα 3.18: Αποτέλεσμα της διαδικασίας κλεισίματος

εικόνα 2 χρωμάτων (ασπρόμαυρη) να χαθούν τα κομμάτια αυτά.

Το πρόβλημα που τίθεται λοιπόν, και τίθεται περισσότερο για λόγους οπτικής καλαισθησίας και όχι για πρακτικούς λόγους, είναι αν μπορούμε με κάποια μέθοδο να διορθώσουμε τέτοιες ατέλειες. Η μέθοδος κλεισίματος (closing operation) είναι μια διαδικασία με την οποία γίνεται μια τέτοια προσπάθεια. Ο τρόπος με τον οποίο δουλεύει είναι σε δύο στάδια:

1. Στο πρώτο στάδιο περνάμε διαδοχικά *πάνω* από όλη την εικόνα μια μάσκα 3x3 σε σχήμα σταυρού, $\begin{array}{|c|c|c|}\hline & \bullet & \\ \hline \bullet & \bullet & \bullet \\ \hline & \bullet & \\ \hline \end{array}$. Όταν το κέντρο της μάσκας βρεθεί πάνω από ένα σημείο της εικόνας που είναι μαύρο, τότε γίνονται όλα τα σημεία της εικόνας που αντιστοιχούν σε σημεία της μάσκας με τιμή •, μαύρα. Το κομμάτι αυτό της διαδικασίας λέγεται διαστολή (dilation) γιατί προσθέτει πληροφορία στην εικόνα (εικόνα 3.18γ).

2. Στο δεύτερο στάδιο περνά από την εικόνα η ίδια 3x3 μάσκα, αλλά αυτή τη φορά με διαφορετικό τρόπο λειτουργίας. Όταν το κεντρικό σημείο της μάσκας βρεθεί πάνω από ένα σημείο της εικόνας που είναι άσπρο τότε όλα τα σημεία της εικόνας που αντιστοιχούν σε σημεία της μάσκας που έχουν τιμή • γίνονται λευκά. Το

κομμάτι αυτό της διαδικασίας λέγεται διάβρωση (erosion) γιατί αφαιρεί τμήματα από την πληροφορίας της εικόνας.

Το τελικό αποτέλεσμα είναι αυτό της εικόνας 3.18δ, όπου βλέπουμε ότι η εμφάνιση των γραμμάτων έχει βελτιωθεί αισθητά. Παρ'όλα αυτά, ακόμη και από αυτή την εικόνα, αρχίζουν να διαφαίνονται τα προβλήματα της μεθόδου. Λόγω του βήματος της διαστολής τα γράμματα που έχουν οπές, όπως το ο, τείνουν να γίνουν συμπαγή, να κλείσουν δηλαδή οι οπές αυτές. Έτσι τέτοιες λεπτομέρειες, όπως και λεπτομέρειες σε σφραγίδες χάνονται. Το επόμενο βήμα, της διάβρωσης δεν είναι σε θέση να επαναφέρει τέτοιου είδους πληροφορία με αποτέλεσμα η εικόνα να υφίσταται, ορισμένες φορές, σοβαρή αλλοίωση.

Ο φόβος μιας τέτοιας αλλοίωσης δεν μας επέτρεψε να χρησιμοποιήσουμε και να ενσωματώσουμε τη μέθοδο αυτή στο τελικό σύστημα. Είναι φανερό, πως για να είναι αποτελεσματική μια τέτοια μέθοδος θα έπρεπε να είναι σε θέση να διακρίνει περιοχές, στις οποίες αν εφαρμοζόταν θα δημιουργούσε πρόβλημα, και να τις αποφεύγει. Εκτός από τα κακά αποτελέσματα, τίθεται και το θέμα, του κατά πόσο θα έπρεπε κανείς να επέμβει και να διορθώσει τέτοιες ατέλειες. Αρκετοί θα ισχυρίζονταν πως κάτι τέτοιο δεν είναι αποδεκτό, μια και αλλοιώνει ακόμη περισσότερο την αυθεντική μορφή του κειμένου. Μια που τα αποτελέσματα της μεθόδου, δεν δικαιολογούσαν κάτι τέτοιο, παραιτηθήκαμε εντελώς από τις προσπάθειες για εύρεση καλύτερων αποτελεσμάτων.

### 3.1.6 Συμπίεση Εικόνων

Καθώς οι χρήστες και οι νέες εφαρμογές απαιτούν όλο και καλύτερη ποιότητα στις εικόνες των εφαρμογών τους, η ποσότητα των απαραίτητων δεδομένων αυξάνει πολύ γρήγορα. Αυτή η αυξανόμενη ανάγκη δημιούργησε τις διάφορες τεχνικές συμπίεσης δεδομένων, ή αλλιώς, τρόπους για να αποθηκεύσουμε την ίδια ποσότητα πληροφορίας χρησιμοποιώντας λιγότερα δυαδικά ψηφία. Καλύτερες και πιο αποδοτικές μέθοδοι συμπίεσης σημαίνουν μεγαλύτερες εικόνες, υψηλότερες αναλύσεις, αύξηση των χρωμάτων ή περισσότερες εικόνες.

Η συμπίεση (compression) είναι απλά ένας τρόπος να αναπαριστούμε την πληροφορία πιο αποδοτικά. Κάτι σαν να αφαιρούμε όλα τα "κενά" από τα δεδομένα, θα μπορούσαμε να πούμε. Συνήθως εκμεταλλεύονται τρεις συχνά απαντώμενες ιδιότητες των γραφικών δεδομένων:

- περιέχουν συχνά πλεονάζουσα πληροφορία
- είναι προβλέψιμα
- είναι πολλές φορές μη-χρήσιμα

Ένα σχήμα σαν την κωδικοποίηση Run-length encoding, RLE, που χρησιμοποιεί την ύπαρξη πλεονάζουσας πληροφορίας, λέει "Εδώ υπάρχουν τρία κόκκινα εικονοστοιχεία", αντί να πει, "Εδώ υπάρχει ένα κόκκινο εικονοστοιχείο, τώρα ένα κόκκινο, τώρα

ένα κόκκινο". Η κωδικοποίηση Huffman βασίζεται σε ένα στατιστικό μοντέλο για να μελετήσει την προβλεψιμότητα, χρησιμοποιώντας μικρότερους κώδικες για συχνά επαναλαμβανόμενες τιμές εικονοστοιχείων. Η ύπαρξη μη-χρήσιμων δεδομένων προτρέπει στη χρήση ενός σχήματος συμπίεσης με απώλειες. Για παράδειγμα, το ανθρώπινο σύστημα όρασης δεν δίνει τόση μεγάλη σημασία στη χρωματική πληροφορία (στο πλήθος των χρωμάτων δηλαδή) όση στη πληροφορία φωτεινότητας (στις διαβαθμίσεις της). Θα μπορούσαμε επομένως να εκμεταλλευτούμε το γεγονός αυτό, για να κωδικοποιήσουμε τη πληροφορία με λιγότερα δυαδικά ψηφία.

Στη δική μας εργασία μελετήθηκαν ως προς την απόδοση αρκετοί από τους πιο διαδεδομένους αλγορίθμους. Παρακάτω θα παρουσιάσουμε σε γενικές γραμμές τον τρόπο λειτουργίας του καθενός.

#### 3.1.6.1 Συμπίεση με τον αλγόριθμο PackBits

Ο αλγόριθμος PackBits είναι μια από τις δημοφιλέστερες υλοποιήσεις της Run-length συμπίεσης. Η συμπίεση RLE αντικαθιστά μια σειρά από επαναλαμβανόμενες τιμές (φωτεινότητας εικονοστοιχείων) με μια τιμή και ένα πλήθος επαναλήψεων. Για παράδειγμα, χρησιμοποιώντας γράμματα για να αναπαραστήσουμε τιμές, μια σειρά τιμών όπως η αβββββββγγδδδδεεδδδ θα μπορούσε να αναπαρασταθεί σαν 1α6β2γ4δ2ε3δ. Η προσέγγιση αυτή είναι αρκετά απλή να υπολοιηθεί και δουλεύει αρκετά ικανοποιητικά με μεγάλες ακολουθίες επαναλαμβανόμενων τιμών. Εικόνες με μεγάλες περιοχές σταθερής φωτεινότητας είναι καλές υποψήφιες για τέτοιου είδους συμπίεση.

Ο PackBits υλοποιεί την μέθοδο RLE θεωρώντας δεδομένα 8 δυαδικών ψηφίων. Οι επαναλαμβανόμενες τιμές κωδικοποιούνται σαν δύο δυαδικές λέξεις (bytes). Η πρώτη περιέχει το πλήθος επαναλήψεων και η δεύτερη την τιμή που θα επαναληφθεί. Στη περίπτωση μη επαναλαμβανόμενων τιμών, όπως για τη σειρά αβγδε, χρησιμοποιεί ένα κωδικό για να δηλώσει την ύπαρξη μιας τέτοιας σειράς χαρακτήρων. Ακολουθεί το μήκος της σειράς και έπειτα οι διαδοχικοί κωδικοί.

Ο PackBits είναι υλοποιημένος σαν διαδικασία συμπίεσης για το είδος (format) αρχείων εικόνας TIFF. Αυτή την υλοποίηση χρησιμοποιήσαμε στις συγκριτικές δοκιμές μας.

#### 3.1.6.2 Συμπίεση LZW (Lempel, Ziv και Welch)

Ο αλγόριθμος LZW ανήκει στο είδος αλγορίθμων που μπορούμε να τους χαρακτηρίσουμε σαν προσαρμοζόμενους. Ξεκινά με ένα απλό πίνακα που περιέχει ένα κωδικό για κάθε πιθανή τιμή δεδομένων και δυναμικά, κατά την διαδικασία εκτέλεσης, προσθέτει καινούργιους κωδικούς στον πίνακα για κάθε μοναδική ακολουθία τιμών που δεν αναγνωρίζει. Είναι απαραίτητο να ορίσει κανείς το μέγιστο μήκος του πίνακα ώστε να οριστεί και ένα επιτρεπτό μήκος κωδικών (αν και μερικές υλοποιήσεις του επιτρέπουν και μεταβλητό μήκος).

Ας θεωρήσουμε για παράδειγμα την ακολουθία αβαβαααγαααααδ, όπου κάθε γράμμα αναπαριστά, για ευκολία, δεδομένα 2 δυαδικών ψηφίων. Ο αρχικός πίνακας θα ήταν:

α:00 β:01 γ:10 δ:11

Ο αλγόριθμος LZW ψάχνει για τη μεγαλύτερη ακολουθία που μπορεί να αναγνωρίσει, τη μεγαλύτερη που υπάρχει στον πίνακα. Βρίσκει λοιπόν τη πρώτη τιμή, α, την αναγνωρίζει, και προσπαθεί να αναγνωρίσει την αβ, αλλά αποτυγχάνει. Δημιουργεί τότε μια νέα θέση στο πίνακα για την ακολουθία αυτή και ο πίνακας γίνεται:

α:000 β:001 γ:010 δ:011 αβ:100

Κατόπιν μεταδίδει την τιμή για την ακολουθία που έχει αναγνωρίσει (000). Παίρνει έπειτα την τελευταία τιμή, β, και δοκιμάζει να αναγνωρίσει μια ακολουθία μαζί με την επόμενη τιμή. Έχει την βα, που δεν την αναγνωρίζει, δημιουργεί μια νέα θέση για την ακολουθία αυτή με κωδικό βα:101, και στέλνει τον κωδικό για το β (001).

Τώρα έρχεται το όφελος. Η επόμενη ακολουθία, αβ, είναι αναγνωρίσιμη. Έτσι ένας κωδικός 3 ψηφίων μπορεί να αντικαταστήσει δύο 2-ψήφιους κωδικούς. Το κέρδος μικρό στην περίπτωση αυτή αλλά σε μια ρεαλιστική κατάσταση αρκετά μεγαλύτερο.

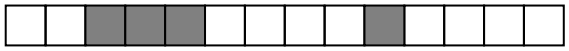
Ο LZW βασίζει την απόδοσή του στην εύρεση ακολουθιών τιμών. Η ύπαρξη θορύβου καταστρέφει τις ακολουθίες αυτές, με αποτέλεσμα εικόνες με υψηλό θόρυβο να μην συμπιέζονται σε μεγάλο βαθμό με τον LZW.

Ο LZW χρησιμοποιείται στα αρχεία εικόνων TIFF και GIF. Η μόνη διαφορά στην υλοποίησή του ανάμεσα στα δύο διαφορετικά είδη αρχείων, είναι ότι στην υλοποίηση για TIFF οι κώδικες έχουν σταθερό μήκος πχ. 12 ψηφίων και όταν υπερβούμε το πλήθος των κωδικών που μπορούν να σχηματιστούν με 12 ψηφία τότε επαναφέρουμε τον πίνακα κωδικών στην αρχική του κατάσταση και ξεκινάμε σχηματίζοντας κωδικούς από την αρχή. Στην υλοποίηση για GIF ο πίνακας δεν έχει σταθερό μέγεθος κωδικών αλλά μεταβάλλεται καθώς ο αλγόριθμος συμπιέζει τα δεδομένα. Και όταν φτάσουμε ένα μέγιστο αριθμό κωδικών μηδενίζουμε τον πίνακα τους και ξεκινάμε από την αρχή.

#### 3.1.6.3 Συμπίεση με τους αλγορίθμους CCITT Group3 και Group4

Πρόκειται για δύο αλγόριθμους συμπίεσης που χρησιμοποιούνται και για την συμπίεση δεδομένων που αποστέλλονται με χρήση τηλεομοιότυπων (fax) [15]. Και οι δύο συμπιέζουν της εικόνες, κωδικοποιώντας τις οριζόντιες γραμμές, μία κάθε φορά. Αυτός ο τρόπος κωδικοποίησης αντιστοιχεί στο τρόπο με τον οποίο σαρώνονται και εκτυπώνονται τα έγγραφα σε μια μηχανή τηλεομοιοτυπίας. Η διαφορά τους έγκειται στον τρόπο με τον οποίο οι δύο μέθοδοι χειρίζονται τις διαδοχικές γραμμές δεδομένων στην εικόνα: στη μέθοδο Group3, κάθε γραμμή κωδικοποιείται ξεχωριστά από την άλλη, ενώ στη μέθοδο Group4, οι γραμμές κωδικοποιούνται με αναφορά στη προηγούμενη, έχοντας σαν αποτέλεσμα καλύτερα ποσοστά συμπίεσης.

Στη μέθοδο Group3, κάθε γραμμή κωδικοποιείται σαν ένα σύνολο από επαναλαμβανόμενες τιμές, που η κάθε μία αντιπροσωπεύει ένα πλήθος από λευκά ή μαύρα εικονοστοιχεία. Η μέθοδος μοιάζει δηλαδή με τον αλγόριθμο PackBits όταν αυτός



| 2A | 3M | 4A | 1M | 4A |
|---|---|---|---|---|
| 0111 | 10 | 1011 | 010 | 1011 |

Σχήμα 3.19: Παράδειγμα κωδικοποίησης για τον αλγόριθμο CCITT Group3

περιοριστεί σε δυο τιμές χρώματος (λευκό, μαύρο). Κάθε ακολουθία κωδικοποιείται με ένα μεταβλητό αριθμό δυαδικών ψηφίων, και μπορεί μοναδικά να αναγνωριστεί κατά τη διαδικασία της αποσυμπίεσης. Με το τρόπο αυτό συχνά επαναλαμβανόμενες τιμές μπορούν να κωδικοποιηθούν με πολύ αποδοτικό τρόπο, εις βάρος των σπανίως εμφανιζόμενων τιμών. Για παράδειγμα, ένα πλήθος επαναλήψεων μαύρων εικονοστοιχείων μπορεί να κωδικοποιηθεί χρησιμοποιώντας μόνο 2 δυαδικά ψηφία, ενώ ένα πλήθος 1000 επαναλήψεων χρειάζεται 25 δυαδικά ψηφία.

Έτσι αν έχουμε, για παράδειγμα (Σχ. 3.19), την ακολουθία εικονοστοιχείων αα μμμ αααα μ ααα, με α να αντιστοιχεί σε άσπρο και μ σε μαύρο εικονοστοιχείο, και γνωρίζοντας, από πίνακες ότι η μέθοδος Group3 κωδικοποιεί 2 άσπρα εικονοστοιχεία με τα δυαδικά ψηφία 0111, 3 μαύρα με τα 10, 4 άσπρα με τα 1011 και 1 μαύρο με τα 010, μπορούμε πολύ εύκολα να δούμε ότι η ακολουθία θα κωδικοποιηθεί σαν 0111 10 1011 010 1011.

Φυσικά η απόδοση της μεθόδου εξαρτάται πάρα πολύ από το είδος της εικόνας προς συμπίεση. Κάτω από κάποιες ιδιαίτερες συνθήκες η μέθοδος θα μπορούσε να δημιουργήσει ένα αρχείο σαν έξοδο που να είναι ακόμη και μεγαλύτερο από το αρχικό (στη περίπτωση πχ. που εναλλάσσεται διαρκώς άσπρο με μαύρο εικονοστοιχείο) αλλά γενικά έχουμε συμπίεση περίπου 10:1 σε κανονικές εικόνες.

Ο αλγόριθμος για τη συμπίεση με τη μέθοδο Group4 είναι αρκετά πιο πολύπλοκος. Όπως αναφέραμε και προηγουμένως κωδικοποιεί τις γραμμές με βάση την προηγούμενή τους, και με αυτό το τρόπο κατακόρυφοι σχηματισμοί στην εικόνα κωδικοποιούνται πολύ αποτελεσματικά. Δεν θα προχωρήσουμε σε περισσότερες λεπτομέρειες αλλά θα παραπέμψουμε τον αναγνώστη που τον ενδιαφέρει το θέμα σε κάποιο βιβλίο που να αναλύει τις μεθόδους συμπίεσης [2], [13], αφού πούμε ότι η μέθοδος αυτή είναι αρκετές φορές καλύτερη από τη μέθοδο Group3 και δίνει μεγέθη συμπίεσης μέχρι και 25:1.

#### 3.1.6.4 Συμπίεση με τον αλγόριθμο JBIG

Ο αλγόριθμος JBIG (Joint Bi-level Image experts Group) είναι ένας άλλος αλγόριθμος κωδικοποίησης ασπρόμαυρων εικόνων χωρίς απώλειες (όπως και οι Group3 και Group4 που αναφέρθηκαν παραπάνω) [27].

Βασίζεται σε μία εντελώς διαφορετική τεχνική που ονομάζεται Αριθμητική Κωδικοποίηση (Arithmetic Coding). Για κάθε εικονοστοιχείο στην εικόνα ένας τύπος ”συμφραζόμενων” (context) εξάγεται για αυτό από τα γειτονικά του εικονοστοιχεία. Για κάθε δεδομένο τύπο συμφραζόμενων υπάρχει μια στατιστική πιθανότητα το εικο-

νοστοιχείο για το οποίο έχει βγει το συμφραζόμενο αυτό να είναι μαύρο ή να είναι άσπρο. Αν η τιμή του εικονοστοιχείου συμφωνεί με την πρόβλεψη, μπορεί να αναπαρασταθεί με λιγότερη πληροφορία από ότι αν δεν συμφωνεί. Καθώς κάθε εικονοστοιχείο κωδικοποιείται ο σχετικός τύπος συμφραζόμενου ανανεώνεται. Τα δυαδικά ψηφία που προκύπτουν με αυτή τη τεχνική κωδικοποιούνται με μια μέθοδο γνωστή σαν "υποδιαίρεση διαστήματος" (interval subdivision) η οποία παρέχει μια σημαντική βελτίωση στη συμπίεση σε σχέση με όλες τις μεθόδους που βασίζονται σε κωδικοποίηση Huffman.

Η μέθοδος είναι αρκετά πολύπλοκη αλλά αποφέρει συμπίεση από 1,1 έως και 1,5 φορές μεγαλύτερη από αυτή του Group4 σε σελίδες με χαρακτήρες που προέρχονται από σάρωση.

### 3.1.7 Αποφάσεις για τη δική μας εργασία

Από αυτά που αναφέρθηκαν στην παράγραφο 3.1.6 καταλαβαίνουμε πως ένα άλλο πολύ σημαντικό κομμάτι του όλου συστήματος είναι το υποσύστημα της συμπίεσης. Λόγω του εξαιρετικά μεγάλου όγκου εγγράφων είναι αυτονόητο πως χρειάζεται κάποιος τρόπος για να μειωθεί το μέγεθος των αρχείων. Για να κάνουμε ακόμη πιο σαφές στον αναγνώστη το πόσο σημαντική είναι η διαδικασία της συμπίεσης ας κάνουμε μερικούς υπολογισμούς. Ας υποθέσουμε ότι στη βιβλιοθήκη μας θα εισάγουμε εικόνες εγγράφων με διαστάσεις A4. Σύμφωνα με τον πίνακα 3.1 η σελίδα A4 απαιτεί για μια ασπρόμαυρη εικόνα περίπου 1 MB[4]. Επομένως για 1.500.000 έγγραφα, όσα περίπου διαθέτει η Βικελαία Δημοτική Βιβλιοθήκη, θα απαιτούνταν 1.500.000 MB, αστρονομικό ποσό για τα σημερινά δεδομένα, ποσό που στη πραγματικότητα θα οδηγούσε στη μη πραγμάτωση του προγράμματος ευθύς εξ'αρχής. Με ένα παράγοντα συμπίεσης 25:1, που πετυχαίνει ο αλγόριθμος συμπίεσης CCITT Group4, το απαιτούμενο μέγεθος πέφτει σε 60.000 MB, ενώ με χρήση του αλγορίθμου JBIG μπορούμε να ρίξουμε το μέγεθος σε 30.000 MB περίπου ή ισοδύναμα σε 30 GB[5]. Τέτοια ποσότητα πληροφορίας είναι αρκετά διαχειρίσιμη και μπορεί να αποθηκευτεί χωρίς εξαιρετικά μεγάλο κόστος. Η χρήση εικόνων επιπέδων γκρίζου όταν μιλάμε για τέτοιο πλήθος εγγράφων και τέτοια μεγέθη εικόνων ανήκει απλά στη σφαίρα της επιστημονικής φαντασίας.

Οι παραπάνω αλγόριθμοι συμπίεσης έχουν εξαιρετικά - σε ορισμένες περιπτώσεις - αποτελέσματα. Αν και θα μπορούσαμε να δοκιμάσουμε και κάποιους νέους τρόπους συμπίεσης, στην δική μας εργασία ακολουθήσαμε μια άλλη γραμμή: έχοντας ως γνώμονα τη συμβατότητα του συστήματος που αναπτύσσεται με άλλα συστήματα, προτιμήσαμε να χρησιμοποιήσουμε έναν από τους ήδη υπάρχοντες αλγορίθμους συμπίεσης ώστε να έχουμε αρχεία εικόνων σε κάποιον από τους ήδη γνωστούς τύπους αρχείων (format). Με αυτό τον τρόπο τα αρχεία μπορούν εύκολα και γρήγορα να χρησιμοποιηθούν και από οποιονδήποτε άλλο ενδιαφερόμενο.

Προκειμένου να γίνει επιλογή κάποιας μεθόδου έγινε συγκριτική μελέτη των υποψή-

[4] το 1 MB είναι ποσότητα πληροφορίας που αντιστοιχεί σε 1048576 δυαδικές λέξεις.

[5] ισχύει 1 GB = 1024 MB.

| Μέθοδος | Μέση τιμή | Τυπική απόκλιση |
|---|---|---|
| GIF | 87,301323 | 4,558516 |
| TIFF LZW | 88,178810 | 4,548613 |
| TIFF PACKBITS | 80,033112 | 7,568598 |
| TIFF CCITT G3 | 90,397354 | 3,381052 |
| TIFF CCITT G4 | 92,284767 | 2,947392 |
| JBIG | 93,890732 | 1,941124 |

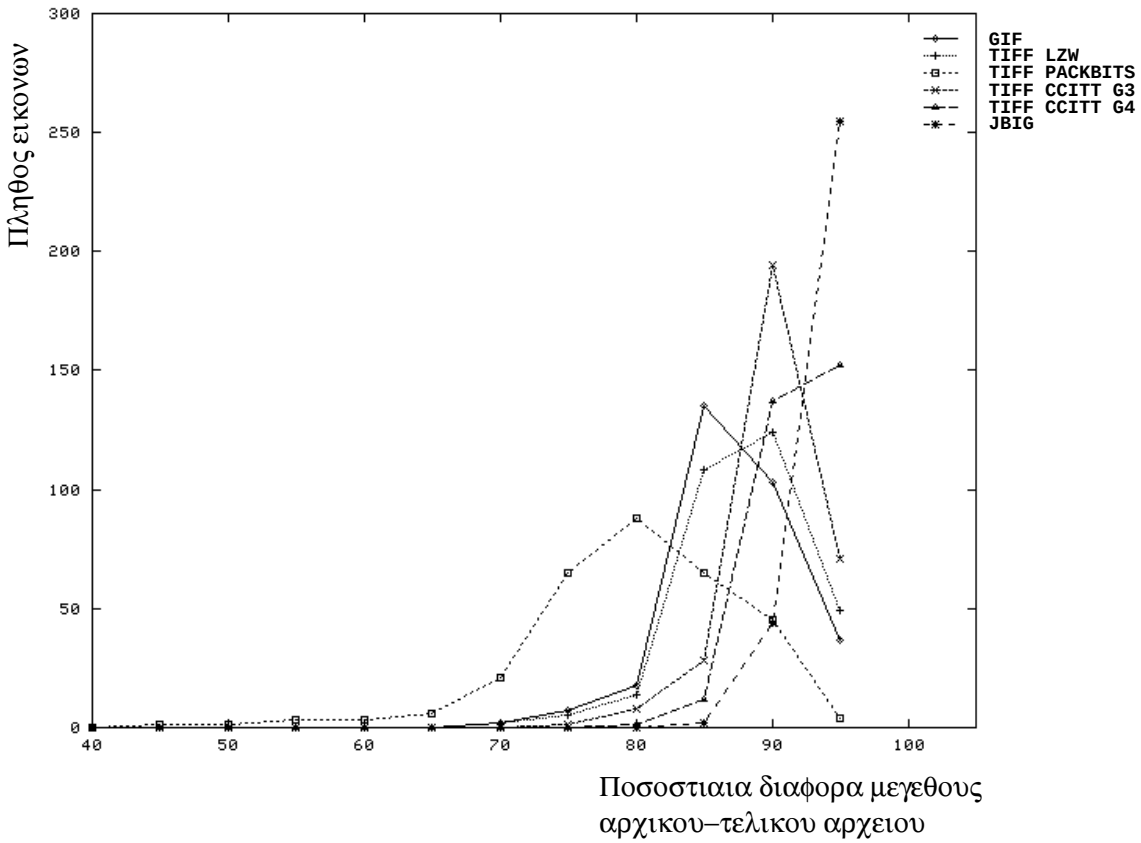
Πίνακας 3.3: Επιδόσεις των αλγορίθμων κωδικοποίησης

φιων μεθόδων, βάσει διαφόρων κριτηρίων απόδοσης. Το πιο σημαντικό ήταν αυτό της μεγαλύτερης συμπίεσης (σε μέγεθος τελικού αρχείου), αλλά ρόλο έπαιξαν και άλλοι παράγοντες όπως χρόνος συμπίεσης, χρόνος αποσυμπίεσης κ.α.

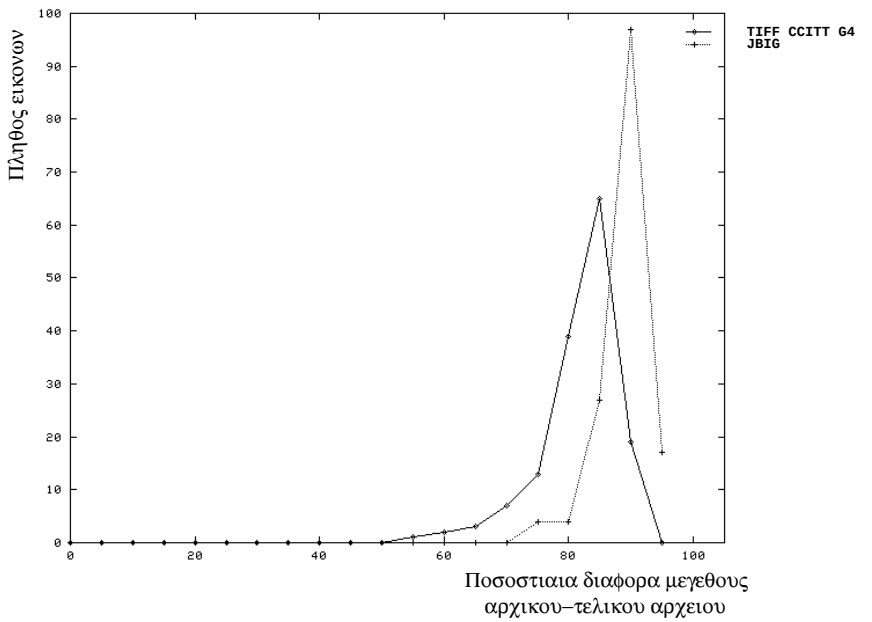
Στην μελέτη αυτή δεν έγιναν δεκτές μέθοδοι συμπίεσης με απώλειες (lossy compression methods). Ο λόγος είναι πως τα αρχεία στη μορφή που βρίσκονται πριν την συμπίεση, περιέχουν χρήσιμη πληροφορία που δεν θα ήταν αποδεκτό να χαθεί κατά την διαδικασία συμπίεσης. Κάτι τέτοιο θα μπορούσε να οδηγήσει σε σημαντικές αλλοιώσεις στο κείμενο των εγγράφων, και να δυσκολέψει τη μετέπειτα μελέτη τους. Αυτό δεν είναι καθόλου αποδεκτό αν αναλογιστούμε την κακή αρχική ποιότητα των εγγράφων. Ακολουθήθηκε λοιπόν και σε αυτή τη περίπτωση μια ασφαλής προσέγγιση προκειμένου να μην χάσουμε ούτε το παραμικρό τμήμα πληροφορίας.

Τα αποτελέσματα των διαφόρων αλγορίθμων συμπίεσης φαίνονται στο Σχ. 3.20 και στον πίνακα 3.3. Από εκεί μπορούμε να δούμε πως η μέθοδος συμπίεσης JBIG έχει πολύ καλύτερα αποτελέσματα από όλες τις υπόλοιπες μεθόδους. Με μέση τιμή συμπίεσης 93,9% μπορεί να συμπιέσει ένα αρχείο μεγέθους 1.087.579 δυαδικών ψηφίων σε 65.254 δυαδικά ψηφία, προσφέροντας τεράστια εξοικονόμηση αποθηκευτικού χώρου, χρημάτων αλλά και χρόνου αποστολής στη περίπτωση που η εικόνα θα πρέπει να μεταφερθεί πάνω από το δίκτυο. Ο αμέσως επόμενος, στις επιδόσεις αλγόριθμος, ο αλγόριθμος TIFF CCITT Group4, διαφέρει από τον JBIG κατά $93,9 - 92,2 = 1,7\%$. Στο προαναφερθέν μέγεθος αρχείου αυτό μεταφράζεται σε 18.500 περίπου δυαδικά ψηφία. Μικρό νούμερο αλλά αν λάβει κανείς υπόψιν το μεγάλο μέγεθος του αρχείου της βιβλιοθήκης, θα εκτιμήσει τη διαφορά. Για να γίνει πιο προφανές αυτό, συνεχίζοντας τον προηγούμενο υπολογισμό βλέπουμε ότι στις 1.000 εικόνες έχουμε διαφορά 18,5 MB. Ένα όχι και τόσο αμελητέο νούμερο για ένα τόσο μικρό αριθμό εικόνων.

Στο Σχ. 3.21 και στον πίνακα 3.4 φαίνεται το πόσο σημαντική είναι η διαδικασία καθαρισμού στην απόδοση κάθε μεθόδου συμπίεσης. Συγκρίνοντας και με τον πίνακα 3.3 βλέπουμε ότι οι δύο καλύτεροι, σε επιδόσεις, αλγόριθμοι, πετυχαίνουν αρκετά χαμηλότερες επιδόσεις. Η διαφορά τους είναι της τάξεως του 10,5% για τον αλγόριθμο TIFF CCITT Group4, και 5% για τον αλγόριθμο JBIG. Συγκρίνοντας τις δύο τιμές βλέπουμε ότι ο JBIG είναι ελαφρώς πιο ανθεκτικός στην ύπαρξη θορύβου στην εικόνα




Σχήμα 3.20: Σύγκριση αλγορίθμων κωδικοποίησης




Σχήμα 3.21: Επιδόσεις στη συμπίεση εικόνων, στις οποίες δεν έχει προηγηθεί η διαδικασία καθαρισμού

| Μέθοδος | Μέση τιμή | Τυπική απόκλιση |
|---|---|---|
| TIFF CCITT G4 | 81,812080 | 6,430182 |
| JBIG | 88,993286 | 3,885089 |

Πίνακας 3.4: Αριθμητικά μεγέθη συμπίεσης χωρίς τη διαδικασία καθαρισμού

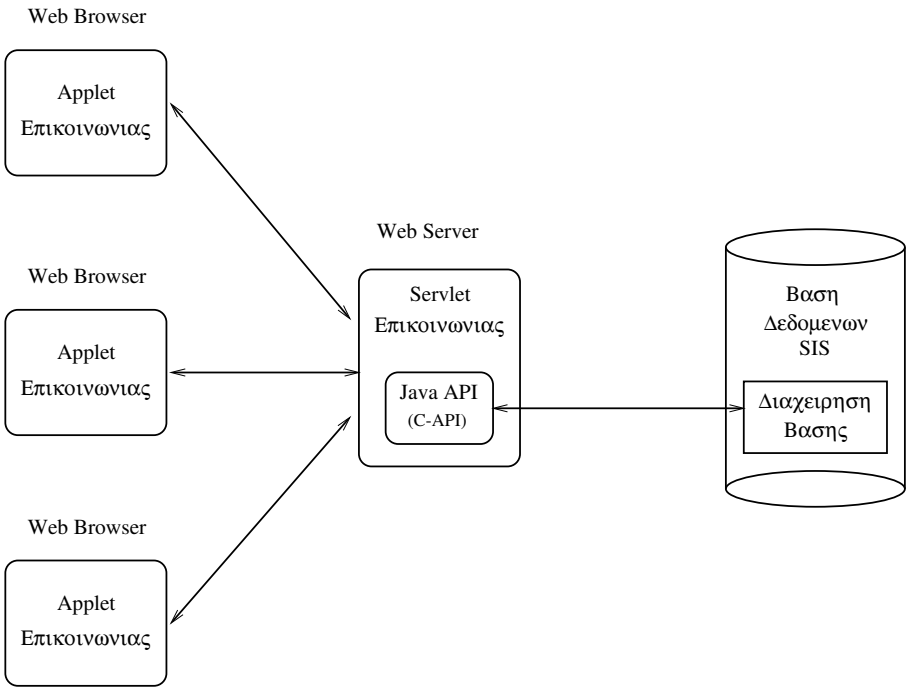
αφού δεν έχει τόσο μεγάλη πτώση στην απόδοσή του όσο ο TIFF CCITT Group4.

Τα μεγέθη συμπίεσης πάντως, πριν και μετά τη διαδικασία καθαρισμού, δικαιολογούν απολύτως την ύπαρξη της διαδικασίας αυτής. Πολύ δε περισσότερο, αφού αυτή επεμβαίνει αφαιρώντας μόνο θόρυβο από την εικόνα, χωρίς να καταστρέψει πολύτιμη πληροφορία.

## 3.2 Στάδιο ανάκτησης και εμφάνισης

Στο στάδιο αυτό οι εικόνες έχουν ήδη αποθηκευτεί και οργανωθεί από τη βάση δεδομένων του Συστήματος Σημασιολογικού Ευρετηριασμού (ΣΣΕ). Ο χρήστης είναι πια σε θέση να ζητήσει από το σύστημα κάποια έγγραφα.

Το σύστημα που αναπτύχθηκε για τη Βικελαία Βιβλιοθήκη βασίζεται στο μοντέλο επικοινωνίας εξυπηρέτη/εξυπηρετητή (βλέπε [17] και Σχ. 3.22). Το ρόλο του εξυπηρετητή έχει ένα μηχάνημα το οποίο περιέχει τόσο το σύστημα ΣΣΕ όσο και τα αρχεία των εικόνων και όλες τις άλλες πληροφορίες με τις οποίες είναι οργανωμένη η βάση μας. Η επικοινωνία μεταξύ των διαφόρων υποσυστημάτων γίνεται με τη βοήθεια




Σχήμα 3.22: Μοντέλο επικοινωνίας του συστήματος

του υποσυστήματος επικοινωνίας. Τα απαραίτητα προγράμματα είναι γραμμένα στη γλώσσα προγραμματισμού Java.

### 3.2.1 Γλώσσα προγραμματισμού Java

Η ανάπτυξη τόσο του υποσυστήματος εμφάνισης εικόνων και παροχής υπηρεσιών, που είναι τμήμα της παρούσας εργασίας, όσο και των προγραμμάτων επικοινωνίας, έγινε αρκετά πιο εύκολα και σε πιο σύντομο χρονικό διάστημα με την χρήση της γλώσσας προγραμματισμού Java [26].

Η Java είναι μια γλώσσα προγραμματισμού που εγγενώς υποστηρίζει τη λειτουργία των προγραμμάτων, που έχουν γραφτεί με αυτή, πάνω από περιβάλλον δικτύων [12], [33]. Μια και η λειτουργία πάνω από περιβάλλον δικτύων ήταν από τα πιο σημαντικά θέματα στη σχεδίαση του όλου συστήματος, υιοθετήθηκε η χρήση της γλώσσας αυτής στο πρόγραμμα ΑΡΧΟΝ. Με κομψό και συνάμα εύκολο και γρήγορο τρόπο δόθηκαν λύσεις στα προβλήματα που έχει κανείς να αντιμετωπίσει όταν προσπαθεί να υλοποιήσει ένα τέτοιο τρόπο λειτουργίας.

Η επιλογή της Java βασίστηκε επίσης στο γεγονός ότι η γλώσσα αυτή είναι αρκετά δυνατή σε υλοποίηση διεπιφανειών χρήσης καθώς παρέχει αρκετά από τα δομικά στοιχεία των διεπιφανειών αυτών ενσωματωμένα στο κορμό της. Τα προγράμματα που προκύπτουν είναι αρκετά μικρά σε μέγεθος και επομένως δεν επιβαρύνουν τη μνήμη των υπολογιστών. Η ταχύτητα εκτέλεσης, αν και δεν είναι από τα δυνατά σημεία της γλώσσας, είναι αρκετά ικανοποιητική. Με την χρήση νέων μεταφραστών καθώς και

μεταφραστών τεχνολογίας JIT[6] (Just In Time compilers), που υπάρχουν αποκλειστικά για τη γλώσσα Java, η ταχύτητά της μπορεί να γίνει ίση με την ταχύτητα εκτέλεσης της γλώσσας C, και να εξαλειφθεί και αυτό το μειονέκτημα.

Όλα τα παραπάνω καθώς και η δυνατότητά της για εκτέλεση του ίδιου προγράμματος χωρίς μετατροπές σε πολλά, διαφορετικά συστήματα υπολογιστών, την καθιστούν μια αρκετά σοβαρή υποψήφια στη περίπτωση που κάποιος έχει να αντιμετωπίσει παρόμοια προβλήματα.

### 3.2.2 Επικοινωνία χρήστη

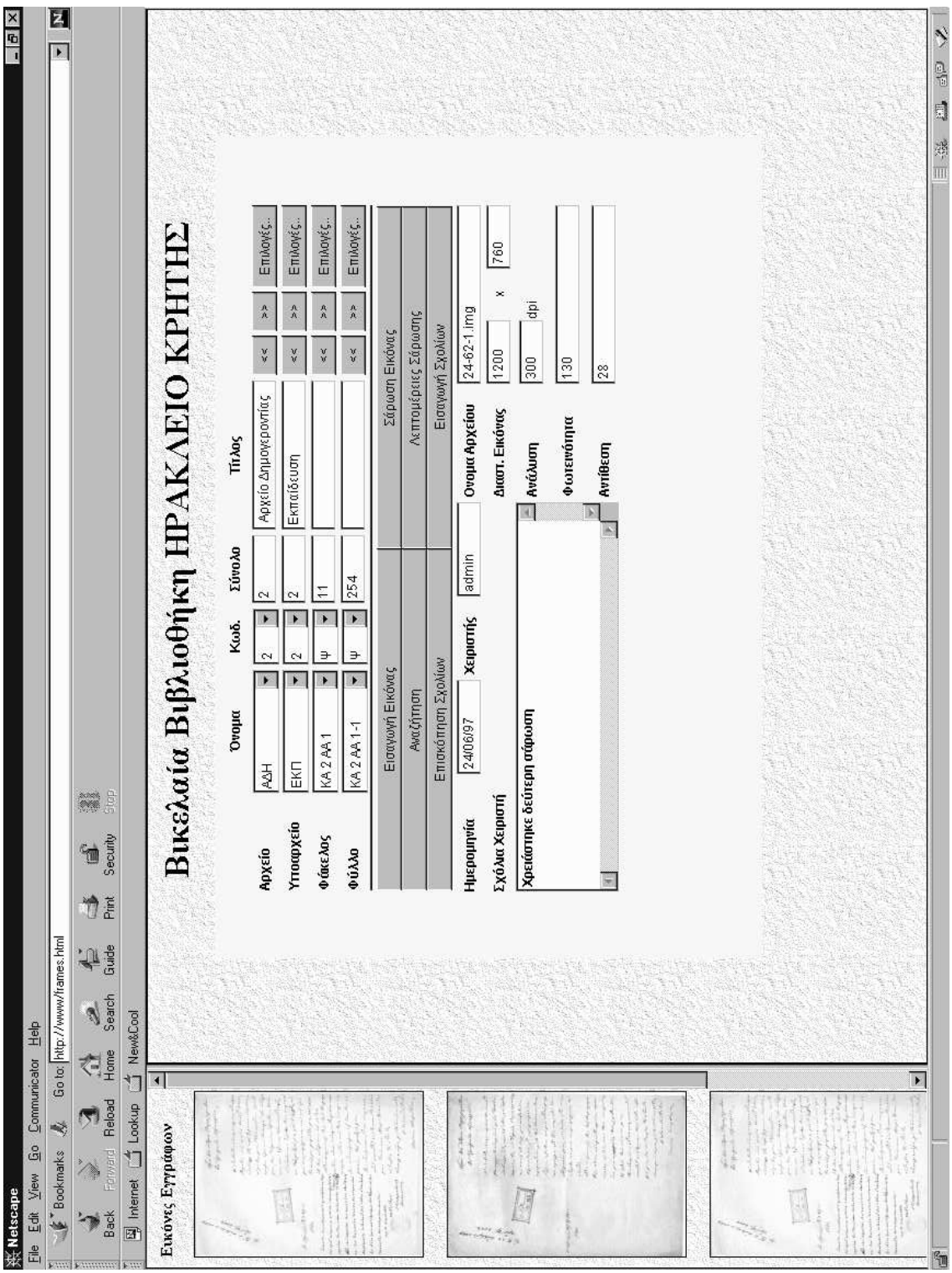
Την επικοινωνία μεταξύ του χρήστη και του συστήματος ΣΣΕ αναλαμβάνουν κάποια προγράμματα που υλοποιούν μια διεπιφάνεια χρήσης (interface). Τα προγράμματα αυτά εκτελούν τις λειτουργίες τους εργαζόμενα μέσα από το περιβάλλον ενός φυλλομετρητή (browser) αφού έχουν υλοποιηθεί με τη γλώσσα προγραμματισμού Java που υποστηρίζει τέτοια δυνατότητα. Με τον τρόπο αυτό πετυχαίνουμε λειτουργία πάνω από περιβάλλον δικτύων, επιτρέποντας την επικοινωνία υπολογιστών (που παίζουν το ρόλο του εξυπηρέτη) με τον εξυπηρετητή ακόμη και από μακρινές αποστάσεις χρησιμοποιώντας το διαδίκτυο.

Στη περίπτωση που ο χρήστης ζητήσει κάποια έγγραφα από το σύστημα οι ενέργειες που γίνονται από αυτό είναι οι ακόλουθες:

- Το πρόγραμμα επικοινωνίας που εκτελείται στο φυλλομετρητή αναλαμβάνει να στείλει την εντολή αναζήτησης στον εξυπηρετητή
- Ο εξυπηρετητής όταν λάβει μια εντολή αναζήτησης ενεργοποιεί το κομμάτι εκείνο του κώδικα που έχει και το οποίο είναι υπεύθυνο για την επικοινωνία με το σύστημα ΣΣΕ. Επειδή το ΣΣΕ είναι γραμμένο στη γλώσσα C++ χρειάστηκε να χρησιμοποιήσουμε τη δυνατότητα της Java για κλήσεις κώδικα σε άλλη γλώσσα προγραμματισμού (Java Native Interface).
- Χρησιμοποιώντας κλήσεις στον κώδικα του ΣΣΕ ο εξυπηρετητής μεταφέρει την επερώτηση του χρήστη στο σύστημα και λαμβάνει πίσω τις απαντήσεις.
- Οι απαντήσεις (που είναι με την μορφή ονομάτων αρχείων ή/και δεικτών σε αυτά) αποστέλλονται πίσω στο χρήστη προκειμένου αυτός να αποφασίσει για τις περαιτέρω ενέργειές του.

Αυτή τη στιγμή ο χρήστης έχει στην οθόνη του τα αποτελέσματα της επερώτησης που έκανε και βλέπει με την μορφή μικρών εικόνων (thumbnails) το περιεχόμενο των εγγράφων που πήρε σαν απάντηση (βλέπε Σχ. 3.23).

[6]Οι μεταφραστές αυτοί κάνουν τη μετάφραση μόνο τη στιγμή που έχει ζητηθεί η εκτέλεση ενός προγράμματος. Με τον τρόπο αυτό δεν υπάρχει η έννοια του εκτελέσιμου και αυξάνεται ακόμη πιο πολύ η λειτουργικότητα των προγραμμάτων πάνω από δίκτυα (μπορούν να γίνονται αλλαγές στον κώδικα χωρίς να χρειάζεται επαναμετάφραση και επανεγκατάσταση του εκτελέσιμου).




Σχήμα 3.23: Εμφάνιση αποτελεσμάτων επερώτησης στη βάση

### 3.2.3 Εμφάνιση εγγράφων

Ο χρήστης τώρα έχει τη δυνατότητα να δει κάποια εικόνα εγγράφου. Στη περίπτωση αυτή ενεργοποιείται το πρόγραμμα που εμφανίζει τις εικόνες στο χρήστη. Το πρόγραμμα αυτό υλοποιήθηκε στα πλαίσια της παρούσας εργασίας με σκοπό να προσφέρει ένα μικρό εργαλείο εμφάνισης και κάποιες ενσωματωμένες δυνατότητες επεξεργασίας των εικόνων.

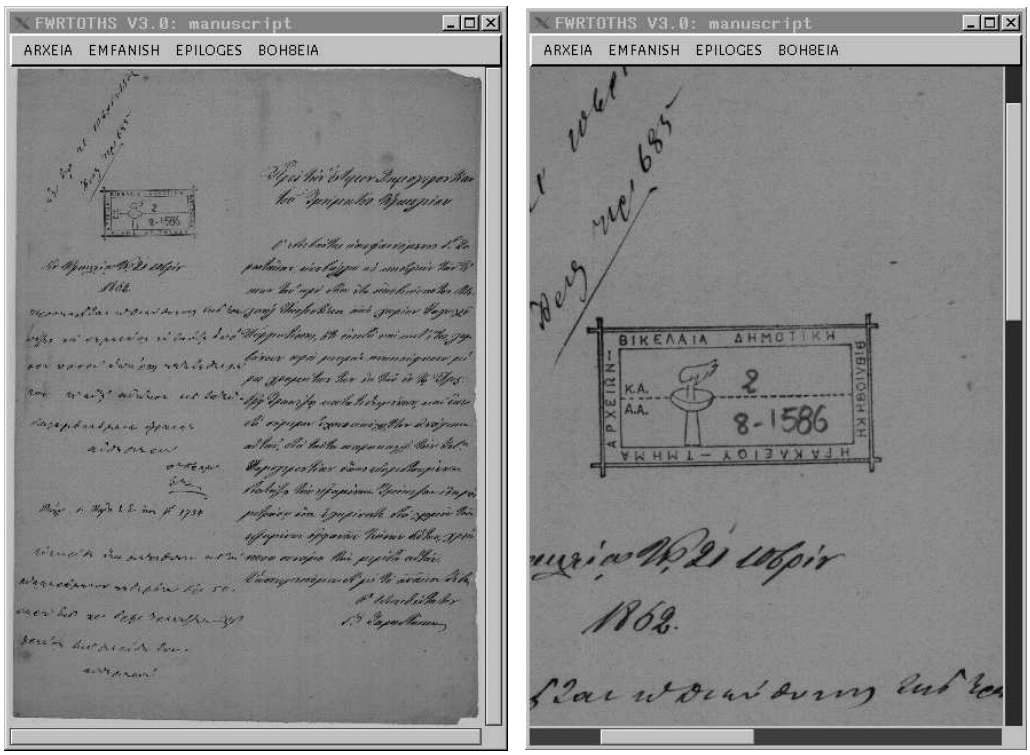
Θα μπορούσε κανείς να παρατηρήσει για ποιόν λόγο έγινε η ανάπτυξη ενός τέτοιου εργαλείου τη στιγμή που υπάρχουν πάρα πολλά ήδη διαθέσιμα, και αρκετά με πολλές και σημαντικές δυνατότητες. Δεν επιλέχθηκε η χρήση ενός έτοιμου εργαλείου γιατί τα περισσότερα δεν προσφέρουν τη δυνατότητα εκτέλεσης μέσα από το περιβάλλον ενός φυλομετρητή. Το χαρακτηριστικό αυτό είναι απολύτως αναγκαίο στη δική μας προσέγγιση της ψηφιακής βιβλιοθήκης, μιας και δίνει τη δυνατότητα της επικοινωνίας πάνω από ένα δίκτυο. Επίσης τα περισσότερα από τα έτοιμα προγράμματα είναι αρκετά απαιτητικά τόσο σε μνήμη όσο και σε ταχύτητα επεξεργαστή. Αυτή η απαίτηση ανεβάζει το κόστος ενός συστήματος που θα χειρίζεται τη βάση δεδομένων, και επιπλέον επιβαρύνει το σταθμό εργασίας με περισσότερη δουλειά. Στη περίπτωση που έχουμε ταυτόχρονη πρόσβαση περισσοτέρων του ενός χρήστη, ο χρόνος απόκρισης του συστήματος[7] θα αυξάνεται σημαντικά αφού θα πρέπει να εκτελεί κάθε φορά που κάποιος ζητά μια εικόνα ένα μεγάλο πρόγραμμα εμφάνισης. Ένας τελευταίος λόγος είναι πως θα θέλαμε να έχουμε πλήρη έλεγχο των δυνατοτήτων του προγράμματος, ώστε να μην δημιουργηθεί κάποιο πρόβλημα από λάθος χρήση του, όπως στη περίπτωση λανθασμένης διαγραφής. Τα έτοιμα προγράμματα δεν υποστηρίζουν κάτι τέτοιο, και γι'αυτό αποκλείστηκαν.

Με την επιλογή λοιπόν κάποιας εικόνας, αναλαμβάνει το πρόγραμμα εμφάνισης τις λειτουργίες εκείνες που πρέπει να επιτελεστούν ώστε να εμφανιστεί η εικόνα. Επικοινωνεί με τον εξυπηρετητή για να πάρει τα δεδομένα της εικόνας, δίνοντάς του το όνομα αρχείου της εικόνας, και κατόπιν αρχίζει να διαβάζει τα πραγματικά δεδομένα της εικόνας πάνω από το δίκτυο. Όταν διαβάσει όλα τα δεδομένα της εικόνας είναι σε θέση να την εμφανίσει στην οθόνη.

Οι εικόνες με τις οποίες ασχολούμαστε είναι αρκετά μεγάλες (οι διαστάσεις τους είναι της τάξης των $3500 \times 2500$ εικονοστοιχείων) και συνεπώς δεν είναι δυνατή η εμφάνισή τους στις φυσικές τους διαστάσεις. Το πρόγραμμα εμφάνισης πρέπει λοιπόν να υποστηρίζει δυνατότητες μεγέθυνσης και σμίκρυνσης των εικόνων και μάλιστα σε πολλαπλά επίπεδα, μια δυνατότητα που ενσωματώθηκε στη δική μας υλοποίηση (βλέπε Σχ. 3.24).

Αφού δεν μπορεί παρά να εμφανιστεί ένα κομμάτι της εικόνας κάθε φορά, είναι αυτονόητο ότι ο χρήστης θα πρέπει να έχει κάποια δυνατότητα κίνησης/πλοήγησης στο υπόλοιπο τμήμα της εικόνας. Επιλέχθηκε στη δική μας περίπτωση να υλοποιηθεί

[7] είναι ο χρόνος που χρειάζεται το σύστημα, από τη στιγμή που κάποιος χρήστης του δώσει μια εντολή, μέχρι να απαντήσει το σύστημα.




Σχήμα 3.24: Η αρχική εικόνα ενός εγγράφου και μια μεγέθυνση περιοχής

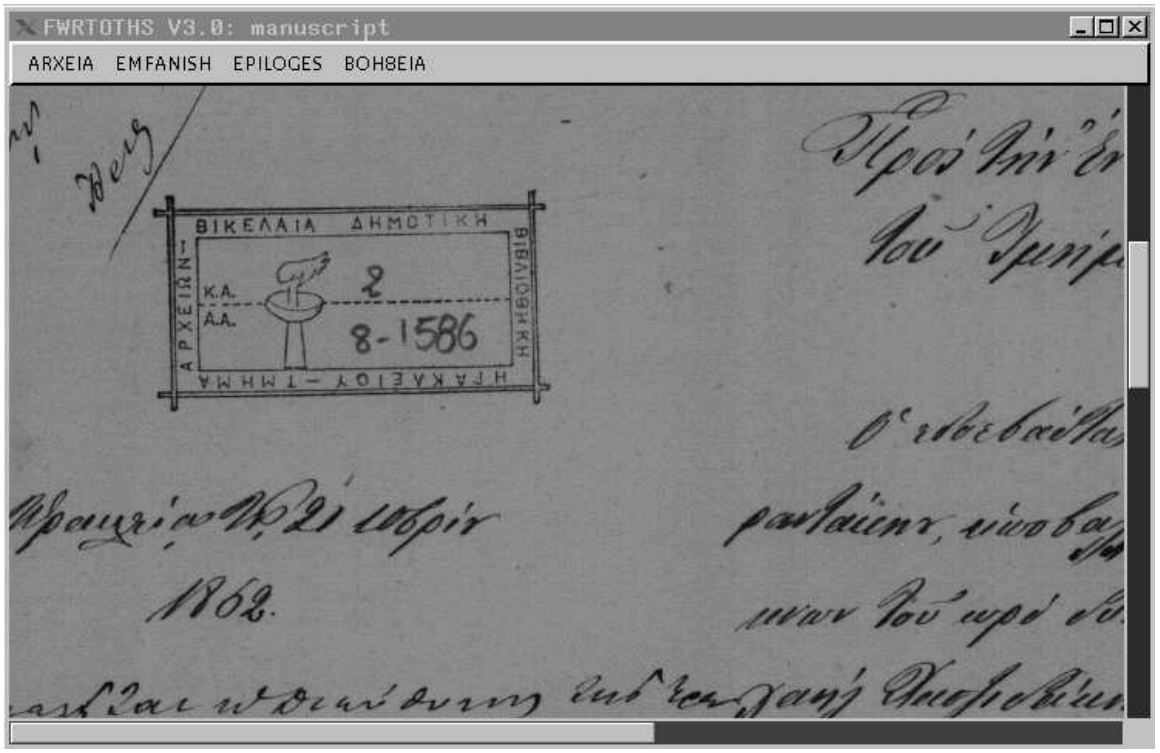
κάτι τέτοιο με τη χρήση μπαρών κύλισης (scrollbars) μια και είναι αρκετά εύκολο να το κατανοήσει και να το χρησιμοποιήσει ακόμη και κάποιος με ελάχιστες γνώσεις από υπολογιστές (έχουμε πάντα στο μυαλό μας αυτό που αναφέρθηκε στην παράγραφο 1.2 σχετικά με είδος των χρηστών στους οποίους απευθύνεται το παρόν σύστημα).

Είναι αρκετά πιθανό λόγω της μορφής ενός εγγράφου (λόγω των διαστάσεών του ή της κατανομής του κειμένου πάνω σε αυτό) να μην είναι εύχρηστο το μέγεθος του παραθύρου στο οποίο εμφανίζεται η εικόνα του εγγράφου. Στην περίπτωση αυτή είναι αρκετά χρήσιμο το πρόγραμμα εμφάνισης να υποστηρίζει την αλλαγή μεγέθους (resizing) του παραθύρου εμφάνισης (Σχ. 3.25). Και αυτή η δυνατότητα έχει υλοποιηθεί στην παρούσα εργασία.

Πολλές φορές μία εικόνα δεν έχει ικανοποιητική φωτεινότητα. Το πρόβλημα αυτό γίνεται πιο εμφανές όταν προκύπτουν εικόνες με αυτοματοποιημένες μεθόδους (όπως και στη δική μας περίπτωση). Οι μέθοδοι αυτές δεν λαμβάνουν ικανοποιητικά υπ'όψιν τους της μεταβολές της φωτεινότητας από έγγραφο σε έγγραφο με αποτέλεσμα άλλα έγγραφα να είναι πιο φωτεινά και άλλα πιο σκοτεινά.

Πολλές φορές επίσης, λόγω διαφορετικής κατασκευής των συσκευών, οι ίδιες εικόνες φαίνονται διαφορετικά σε διαφορετικές οθόνες. Το πρόβλημα αυτό οφείλεται στη μη γραμμική απόκριση της οθόνης στη τιμή της φωτεινότητας. Έτσι ένα εικονοστοιχείο με τιμή φωτεινότητας 200 δεν εκπέμπει πάντα διπλάσιο φως από ένα με τιμή φωτεινότητας 100.

Για να διευκολύνουμε το χρήστη όταν αντιμετωπίζει τέτοιου είδους προβλήματα,




Σχήμα 3.25: Αλλαγή στη γεωμετρία του παραθύρου εμφάνισης

υλοποιήθηκαν και ενσωματώθηκαν στο πρόγραμμα εμφάνισης μερικά απλά φίλτρα επεξεργασίας εικόνων.

Το πρώτο πρόβλημα αντιμετωπίζεται με την ύπαρξη μιας επιλογής στη μπάρα επιλογών (menu) με την οποία μπορούμε να διορθώσουμε το ιστόγραμμα της φωτεινότητας της εικόνας (brightness enhancement). Με την επιλογή αυτή εμφανίζεται ένα παράθυρο εισαγωγής στοιχείων, στο οποίο εισάγουμε το ποσό κατά το οποίο θέλουμε να μετακινηθούν οι τιμές της φωτεινότητας (Σχ. 3.26). Η αλλαγή στις τιμές της φωτεινότητας δίνεται από τη σχέση

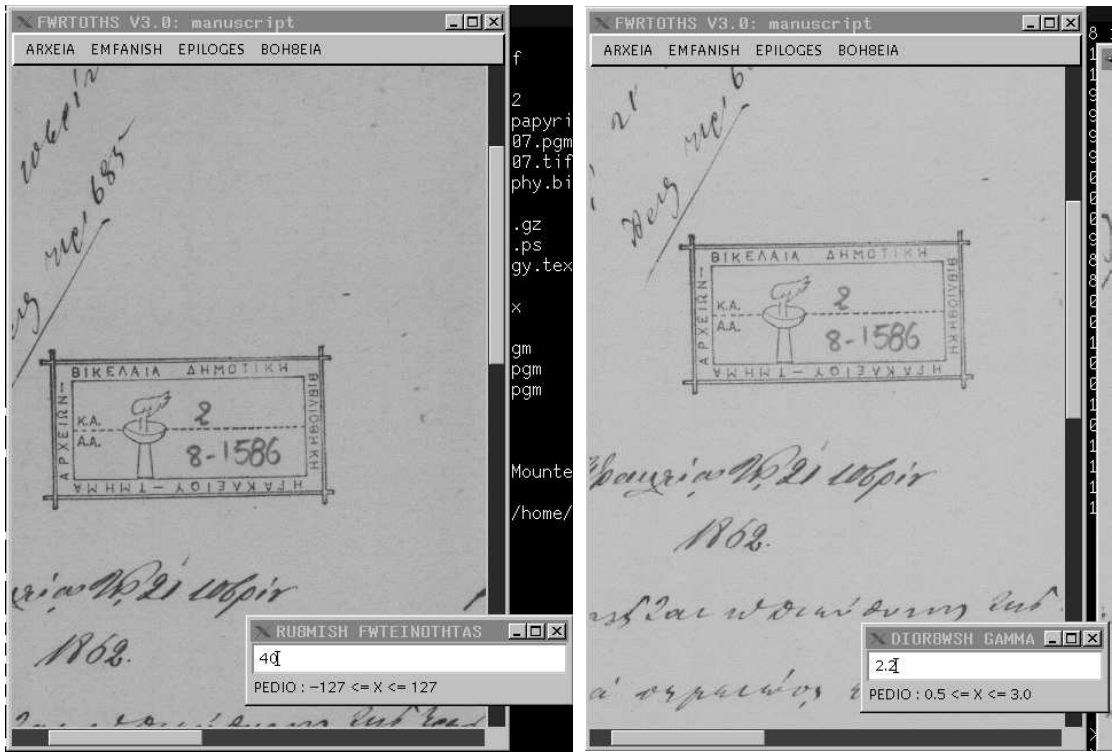
$$I(x,y) = I(x,y) + shift$$

όπου $I(x,y)$ η τιμή της φωτεινότητας στη θέση $(x,y)$ με τιμές στο διάστημα $[0, I_{max}]$ και $shift$ το πόσο θα μετατοπιστεί η φωτεινότητα.

Το πρόβλημα της διαφορετικής απεικόνισης λόγω οθόνης μπορεί να διορθωθεί (Σχ. 3.26) με το φίλτρο Γάμμα (Gamma filter) που είναι ένα μη γραμμικό φίλτρο που δίνεται από την σχέση

$$I(x,y) = I_{max} \cdot \left( \frac{I(x,y)}{I_{max}} \right)^{\frac{1}{\gamma}}$$

όπου $\gamma$ ο παράγοντας διόρθωσης. Οι τυπικές τιμές του παράγοντα διόρθωσης για μία οθόνη Silicon Graphics είναι 1,2, για μια οθόνη NTSC 2,2 ενώ για αυτή ενός συστήματος SUN το εύρος τιμών είναι από 2,0 ως 2,5. Οι περισσότεροι εκτυπωτές έχουν τιμή $\gamma$=1, 0. Στο παράθυρο εισαγωγής στοιχείων μπορούμε να εισάγουμε την επιθυμητή τιμή της παραμέτρου γάμμα.



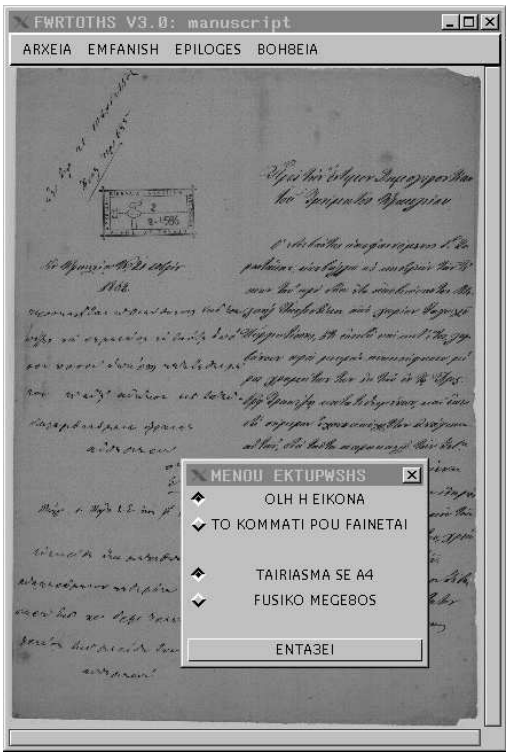

Σχήμα 3.26: Εφαρμογή των φίλτρων διόρθωσης φωτεινότητας και διόρθωσης γάμμα αντίστοιχα

Αν ο χρήστης δεν μείνει ικανοποιημένος με το αποτέλεσμα των αλλαγών που έχει κάνει μπορεί να επαναφέρει την αρχική εικόνα μέσω μιας επιλογής επαναφοράς της εικόνας στην αρχική της κατάσταση που υπάρχει στη μπάρα επιλογών.

Μια πολύ σημαντική δυνατότητα είναι αυτή της εκτύπωσης της εικόνας. Μέσω ενός παράθυρου επιλογών μπορούμε να επιλέξουμε αν θα εκτυπωθεί όλη η εικόνα ή μόνο το κομμάτι που φαίνεται στο παράθυρο εμφάνισης, καθώς και το αν θα γίνει εκτύπωση σε φυσικό μέγεθος ή θα γίνει ταίριασμα σε μέγεθος Α4 (Σχ. 3.27). Αφού γίνουν οι επιλογές, ο χρήστης έχει τη δυνατότητα να εκτυπώσει την εικόνα ή να την αποθηκεύσει σε ένα αρχείο σε μορφή postscript (μορφή που επιτρέπει την άμεση εκτύπωση του). Η δυνατότητα της εκτύπωσης είναι αρκετά σημαντική γιατί επιτρέπει στους χρήστες να αποκτήσουν υλικό και τους δίνει τη δυνατότητα να κάνουν τη μελέτη τους εκτός του χώρου της βιβλιοθήκης. Με τον τρόπο αυτό η βιβλιοθήκη αυξάνει την ευχρηστία της και ταυτόχρονα τον αριθμό των ερευνητών που μπορεί να υποστηρίξει (μιας και θα υπάρχουν περισσότεροι διαθέσιμοι σταθμοί εργασίας/αναζήτησης).

### 3.2.4 Σύνοψη λειτουργιών

Ανακεφαλαιώνοντας, το υποσύστημα εμφάνισης εικόνων και παροχής υπηρεσιών αναλαμβάνει τη προβολή των εικόνων από τη βάση δεδομένων στην οθόνη του χρήστη. Σαν σκοπό του έχει τη παροχή λειτουργιών με απλό και κατανοητό τρόπο, ακόμη και




Σχήμα 3.27: Εκτύπωση μέσα από το πρόγραμμα εμφάνισης

σε ανθρώπους με λίγες γνώσεις στο χώρο της τεχνολογίας των υπολογιστών.

Οι λειτουργίες που προσφέρει είναι όλες οι απαραίτητες για την σωστή εμφάνιση μιας εικόνας. Για το σκοπό αυτό έχουν υλοποιηθεί λειτουργίες όπως μεγέθυνση, σμίκρυνση, αλλαγή της γεωμετρίας ενός παραθύρου και κύλιση με μπάρες για εμφάνιση όλων των τμημάτων μιας εικόνας. Υπάρχουν λειτουργίες βελτίωσης της οπτικής εμφάνισης μιας εικόνας με τη χρήση απλών φίλτρων, όπως το φίλτρο διόρθωσης της φωτεινότητας και της παραμέτρου γάμμα. Τέλος, έχει προστεθεί η λειτουργία της εκτύπωσης τμήματος ή και ολόκληρης της εικόνας.

Με αυτές τις λειτουργίες δεν θέλουμε με κανένα τρόπο να ισχυριστούμε ότι προσφέρουμε ένα πλήρες πρόγραμμα εμφάνισης, εφάμιλου των έτοιμων προγραμμάτων της αγοράς, αλλά μια αρκετά λειτουργική και κομψή απάντηση για τις ανάγκες του συστήματος ΑΡΧΟΝ.

# Κεφάλαιο 4

# Μελλοντικές κατευθύνσεις

Οπωσδήποτε ένα τόσο σημαντικό θέμα που σχετίζεται με την ανάπτυξη και λειτουργία μιας ψηφιακής βιβλιοθήκης δεν τελειώνει εδώ, ούτε μπορεί εύκολα να καλυφθεί στα πλαίσια μιας μεταπτυχιακής εργασίας. Οι προοπτικές του είναι πολύ μεγάλες και οι πειραματισμοί και οι βελτιώσεις που θα μπορούσε να κάνει κανείς σχεδόν ατελείωτες.

Η παρούσα εργασία επικεντρώθηκε στα υποσυστήματα εκείνα της ψηφιακής βιβλιοθήκης που είναι απαραίτητα για την δημιουργία υλικού σε μορφή εικόνων και τη μετέπειτα χρήση τους. Προτάθηκαν, αναπτύχθηκαν και δοκιμάστηκαν πειραματικά, διάφορες ρουτίνες επεξεργασίας εικόνων με στόχο τη βελτίωση της ποιότητάς τους μετά τη διαδικασία της σάρωσης. Έγινε σύγκριση πολλών μεθόδων συμπίεσης και αποτίμησης των επιδόσεών τους. Και τέλος υλοποιήθηκε ένα πρόγραμμα εμφάνισης των εικόνων και παροχής υπηρεσιών.

Τα κέρδη που αποκομίσθησαν με την μορφή εμπειρίας και επιπλέον γνώσεων είναι πολλά. Μέσα από τη διαδικασία ανάπτυξης των συνιστωσών του προγράμματος ΑΡΧΟΝ στις οποίες επικεντρώθηκε η παρούσα εργασία, ήρθαμε αντιμέτωποι με ποικίλα και ενδιαφέροντα προβλήματα τόσο σε θεωρητικό, όσο και σε πρακτικό επίπεδο. Δόθηκαν αρκετές λύσεις σε μερικά προβλήματα αλλά σίγουρα υπάρχουν και πολλά ακόμη που θα μπορούσαν να γίνουν.

Ένα από τα πιο ενδιαφέροντα θέματα είναι η χρήση υδατογραφημάτων (watermarks) για την προστασία των εικόνων από κακόβουλη χρήση. Η υδατογράφηση των εικόνων είναι από τα θέματα που έχουν γνωρίσει αρκετή δημοσιότητα τον τελευταίο καιρό και έχει ήδη γίνει αρκετή δουλειά επάνω τους. Βασίζεται στην ενσωμάτωση μιας οποιαδήποτε εικόνας, που ονομάζεται υδατογράφημα και που συνήθως είναι κάποιας μορφή λογότυπο, στην αρχική εικόνα με τέτοιο τρόπο που το υδατογράφημα να είναι ευδιάκριτο πίσω από την εικόνα, ενώ ταυτόχρονα η εικόνα δεν υφίσταται μεγάλη αλλοίωση παραμένοντας εύχρηστη και λειτουργική. Η προσθήκη του υδατογραφήματος γίνεται με τέτοιο τρόπο ώστε να είναι δύσκολη η αφαίρεσή του από κάποιο χρήστη έτσι ώστε να μην μπορεί εύκολα να πάρει την εικόνα στην αρχική της μορφή. Με τη χρήση του υδατογραφήματος προστατεύονται τα πνευματικά δικαιώματα του ιδιοκτήτη και ταυτόχρονα γίνεται εφικτή η εμφάνιση των εικόνων στο διαδίκτυο. Για μια πρώτη

επαφή μπορεί κανείς να επισκεφθεί την [23] όπου υπάρχουν κάποια πρώτα αποτελέσματα πάνω στη χρήση του υδατογραφήματος καθώς και πιο αναλυτικές πληροφορίες για τη μέθοδο αυτή.

Η χρήση της τεχνολογίας VRML (Virtual Reality Modeling Language - Γλώσσα Μοντελοποίησης Εικονικής Πραγματικότητας) για την επαύξηση των δυνατοτήτων μιας ψηφιακής βιβλιοθήκης είναι μια άλλη πολύ ενδιαφέρουσα ιδέα. Η VRML μπορεί να περιγράψει τον τρισδιάστατο κόσμο, βασισμένη πάνω σε ένα μοντέλο που της παρέχει ένας σχεδιαστής, και κατόπιν είναι σε θέση να κάνει αναπαράσταση του χώρου αυτού. Μπορεί να αποδώσει οπτικά, όψεις του χώρου από οποιαδήποτε γωνία και από οποιοδήποτε σημείο. Η μοντελοποίηση δεν περιορίζεται σε μία περιοχή μόνο (πχ. σε εξωτερικούς χώρους), αλλά καλύπτει οτιδήποτε είναι δυνατό να παρασταθεί με χρήση εικόνων και σχεδίων. Θα ήταν αρκετά ενδιαφέρον να μπορούσε κάποιος να συνδέσει τις περιγραφές και απεικονίσεις της VRML με το περιεχόμενο ενός εγγράφου ή ενός κειμένου γενικότερα. Έτσι, αν ένα έγγραφο αναφερόταν σε κάποιο ιστορικό κάστρο, θα μπορούσε ένα σύστημα να του δώσει τη δυνατότητα να δει το κάστρο στην οθόνη του, να περιηγηθεί στους χώρους του, και γενικά να βιώσει και να αντιληφθεί με τα ίδια του τα μάτια το χώρο αυτό. Το ίδιο θα μπορούσε να γίνει στη περιγραφή ενός αντικειμένου (όπλου, κοσμήματος, αντικειμένου τέχνης), με την εμφάνιση του αντικειμένου στην οθόνη και τη δυνατότητα που προσφέρει η VRML για θέαση από οποιαδήποτε γωνία και σημείο. Κάτι τέτοιο θα πρόσφερε αυξημένες δυνατότητες κατανόησης, ιδίως εκείνων των πληροφοριών που αφορούν περιγραφές χώρων, γεωγραφικών περιοχών, αντικειμένων κα. Μια πρώτη προσπάθεια τέτοιας προσέγγισης, έχει γίνει στα πλαίσια μιας συνεργασίας στο εργαστήριο Μηχανικής Όρασης και Ρομποτικής του Ινστιτούτου Πληροφορικής στο ΙΤΕ, και κάποια αποτελέσματα όπως και άλλες πληροφορίες γενικότερης φύσης για την VRML υπάρχουν στο [30].

Μια άλλη ιδέα που εμφανίστηκε σχετικά πρόσφατα στο εργαστήριο Μηχανικής Όρασης και Ρομποτικής του Ινστιτούτου Πληροφορικής στο ΙΤΕ, είναι και αυτή του σημαδέματος (marking) περιοχών [7]. Πρόκειται για μια τεχνική που επιτρέπει το σημάδεμα διαφόρων οντοτήτων (όπως εικόνας, ήχου, βίντεο) για την γρήγορη εισαγωγή σημασιολογικής πληροφορίας. Προς το παρόν υπάρχει μόνο η υλοποίησή της για εικόνα (σύστημα ImageTagger), που επιτρέπει το σημάδεμα οποιωνδήποτε περιοχών μιας εικόνας [38]. Για παράδειγμα σε μια εικόνα εγγράφου που περιέχει μια σφραγίδα, θα μπορούσε κάποιος να σημαδέψει τη σφραγίδα με το ποντίκι και να εισάγει κατόπιν πληροφορίες που την αφορούν (όπως πχ. ημερομηνία). Ακόμη, στο σημείο που αναφέρεται το όνομα μιας ιστορικής προσωπικότητας, θα μπορούσε να υπάρχει κινηματογραφικό υλικό με μια σύντομη βιογραφία της προσωπικότητας αυτής. Με τον τρόπο αυτό γίνεται μεγάλη εξοικονόμηση χρόνου και ταυτόχρονα η πληροφορία είναι άμεσα διαθέσιμη στο σημείο που κανείς θα περίμενε να υπάρχει (κάποια άλλη χρονική στιγμή ο χρήστης βλέποντας τη περιοχή σημαδεμένη μπορεί να ζητήσει τις πληροφορίες που υπάρχουν, γνωρίζοντας από την αρχή πως θα αφορούν τη σφραγίδα). Με την εύκολη και γρήγορη εισαγωγή στοιχείων δημιουργούνται καλύτερες προϋποθέσεις αναζήτησης

κατόπιν. Αν κάποιος μπορεί να εισάγει εύκολα πληροφορίες, λογικό είναι να εισάγει περισσότερες από ότι αν η διαδικασία εισαγωγής ήταν χρονοβόρα ή δύσκολη. Με περισσότερες όμως πληροφορίες έχουμε ένα σύστημα αναζήτησης, που εντοπίζει με μεγαλύτερη ακρίβεια τα έγγραφα που του ζητά ο χρήστης. Άρα με αυτό τον τρόπο αυξάνουμε και τη λειτουργικότητα του όλου συστήματος. Θα ήταν λοιπόν αρκετά χρήσιμο να ενσωματωθεί μια τέτοια δυνατότητα εισαγωγής στοιχείων και στο σύστημα ΑΡΧΟΝ.

Τέλος ένα θέμα με μεγάλες δυνατότητες ανάπτυξης είναι και αυτό της κατασκευής ενός εργαλείου εμφάνισης εικόνων με ακόμη καλύτερες και περισσότερες λειτουργίες. Θα μπορούσαν να προστεθούν νέα και πιο πολύπλοκα φίλτρα για την περαιτέρω βελτίωση της οπτικής εμφάνισης των εικόνων ή η δυνατότητα υποστήριξης ξεχωριστών ρυθμίσεων για κάθε χρήστη. Με τον τρόπο αυτό δεν θα είναι αναγκασμένος κάποιος να κάνει συνεχώς τις ίδιες διορθώσεις αλλά αυτές θα υπάρχουν κρατημένες σε κάποιο αρχείο (user profile), και θα ενεργοποιούνται αυτόματα από το σύστημα με το που κάποιος χρήστης αναγνωριστεί.

# Βιβλιογραφία


[1] VENIVA Venetian Virtual Archive.
http://www.tin.it/veniva/home/project.html.

[2] Jerry D. Gibson, Toby Berger and David Lindbergh. *Digital Compression for Multimedia: Principles and Standards.* Morgan Kaufman Publishers, 1998.

[3] Kenneth R. Castleman. *Digital Image Processing.* Prentice-Hall, second edition, 1995.

[4] Chiou-Ting and Ja-Ling Wu. Hidden Signatures In Images. In *Proc. of ICIP'96*, volume III, pages 743--746, September 1996.

[5] Venetian Castles Of Crete.
http://www.munch.org.gr/scripts/veniva/forts.exe/.

[6] I. Dionysiadou and M. Doerr. Mapping of material culture to a semantic network. In *International Council of Museums Documentation Committee and Computer Network*. Wasington, USA, August 1994.

[7] K.V. Chandrinos, J. Immerkær, Martin Doerr and P.E. Trahanias. A Visual Tagging Technique for Annotating Large-Volume Multimedia Databases. $5^{th}$ *DELOS Workshop on Filtering and Collaborative Filtering*, 1998.

[8] Martin Doerr and Irini Fundulaki. SIS - TMS : A Thesaurus Management System for Distributed Digital Collections. In *Second European Conference, ECDL'98*, pages 215--234. Heraklion, Crete, Greece, September 1998.

[9] ECDL'98. *Research and Advanced Technology for Digital Libraries*, Heraklion, Crete, Greece, September 1998.

[10] Digital Formats for Content Reproductions.
http://memory.loc.gov/ammem/formats.html.

[11] Jaakko Sauvola, Tapio Seppänen, Sami Haapakoski and Matti Pietikäinen. Adaptive Document Binarization. In $4^{th}$ *International Conference Document Analysis and Recognition (ICDAR '97)*, 1997.

[12] Elliotte R. Harold. *Java Network Programming.* O'Reilly & Associates, 1st edition, February 1997.

[13] Gilbert Held and Thomas R. Marshall. *Data and Image Compression: Tools and Techniques.* John Wiley & Sons, 1996.

[14] Bell & Howell.
http://www.bellhowell.com/.

[15] R. Hunter and A. H. Robinson. International digital fascimile coding standards. In *Proc. of IEEE*, volume 68, pages 854--867, July 1980.

[16] Zahid Hussain. *Digital Image Processing - (Practical Applications of Parallel Processing Techniques.* Ellis Horwood Limited, first edition, 1991.

[17] K. V. Chandrinos, J. Immerkær and P. E. Trahanias. ARHON: A Multimedia Database Design for Image Documents. *EUSIPCO 98, Special Session on Multimedia Signal Processing*, 1998.

[18] Project Gutenberg Index.
http://www.promo.net/pg/.

[19] John Mylopoulos, Alex Borgida, Matthias Jarke and Manolis Koubarakis. Telos: Representing Knowledge About Information Systems. *ACM Transactions on Information Systems, Vol. 8, No. 4*, pages 325--362, October 1990.

[20] I.J. Cox, J. Kilian, T. Leighton and T. Shamoon. A Secure, Robust Watermark for Multimedia. In *Workshop on Information Hiding*. Newton Institute, Univ. of Cambridge, May 1996.

[21] IBM DB2 Digital Library : The Vatican Library.
http://www.software.ibm.com/is/dig-lib/vatican/vatican.html.

[22] Library of Congress Home Page.
http://lcweb.loc.gov/.

[23] Visible Watermarking of Historical Documents.
http://www.ics.forth.gr/~ghislain/icsrapp/icsrapp.html.

[24] Electronic Text Center University of Virginia.
http://etext.lib.virginia.edu/.

[25] Toward on-line, worldwide access to Vatican Library materials.
http://eagle.almaden.ibm.com/journal/rd/mintz/mintzer.html.

[26] Java Technology Home Page.
http://java.sun.com/.

[27] JBIG Home Page.
http://www.jpeg.org/public/jbighomepage.htm.

[28] Perseus Project Home Page.
http://www.perseus.tufts.edu/.

[29] Query By Image Content (QBIC) Home Page.
http://wwwqbic.almaden.ibm.com/.

[30] Heraklion Pre-rapport de stage a l'Institute of Computer Science.
http://www.ics.forth.gr/~pierre/rapportonline/rapportonline.html.

[31] Processing Recording and Archiving Carbonized Papyri.
http://www.cs.hut.fi/papyrus/.

[32] Manuscript Digitization Demonstration Project Summary Report.
http://memory.loc.gov/ammem/pictel/index.html.

[33] Prashant Sridharan, Bill Rieken and Laraine Peterson. *Advanced Java Networking.* Prentice Hall Professional Technical Reference, April 1997.

[34] Azriel Rosenfeld and Avinash C. Kak. *Digital Image Processing.* Academic Press, Inc., second edition, 1982.

[35] Digital Collections (DL SunSITE).
http://sunsite.berkeley.edu/collections/.

[36] LibraryWeb: Columbia University Libraries WWW Information System.
http://www.columbia.edu/cu/libraries/.

[37] Information Systems and Software Technology Division (SIS System).
http://www.ics.forth.gr/proj/isst/systems/sis/.

[38] K.V. Chandrinos, P.E. Trahanias. Beyond HTML: Web-based information systems. $6^{th}$ *DELOS Workshop on Digital Preservation*, 1998.

[39] G. Voyatzis and I. Pitas. Problems and Challenges in Multimedia Networking and Content Protection. In *Proc. of Workshop on Trends and Important Challenges in Signal Processing.* Kirkkonummi, Finland, June 1998.

[40] Robert Fisher, Simon Perkins, Ashley Walker and Erik Wolfart. *Hypermedia Image Processing Reference (HIPR).* John Wiley & Sons Ltd, May 1997.