

# Μάθημα 1

## Ψηφιακά έγγραφα

### 1.1 Ορισμός ψηφιακών εγγράφων

Ο ορισμός των ψηφιακών εγγράφων μπορεί να δοθεί ως:

- Αναπαράσταση δεδομένων κατανοητή από τα υπολογιστικά συστήματα π.χ. μία εικόνα της πραγματικότητας, μία ψηφιακή εγγραφή ήχου ή οποιαδήποτε πληροφορία σε μορφή κατανοητή από τον υπολογιστή (ένα πρόγραμμα υπολογιστή, ένα αρχείο δεδομένων κ.λ.π.). Στο σχήμα 1.1 δίδονται διάφορα παραδείγματα ψηφιακών εγγράφων.
- Ένας τρόπος δόμησης της πληροφορίας σε ψηφιακή μορφή
- Ένα σύνολο από bits (που έχουν συγκεκριμένη δομή, περιεχόμενο και πληροφορίες)
- Το βασικό δομικό στοιχείο κάθε ψηφιακής βιβλιοθήκης



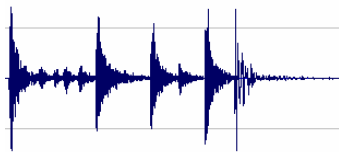
(α)



(β)



(γ)



(δ)



(ε)

**Σχήμα 1.1.** Παραδείγματα ψηφιακών εγγράφων: (α) κείμενο, (β) εικόνα, (γ) γραφικό, (δ) ήχος, (ε) frames ενός animation.

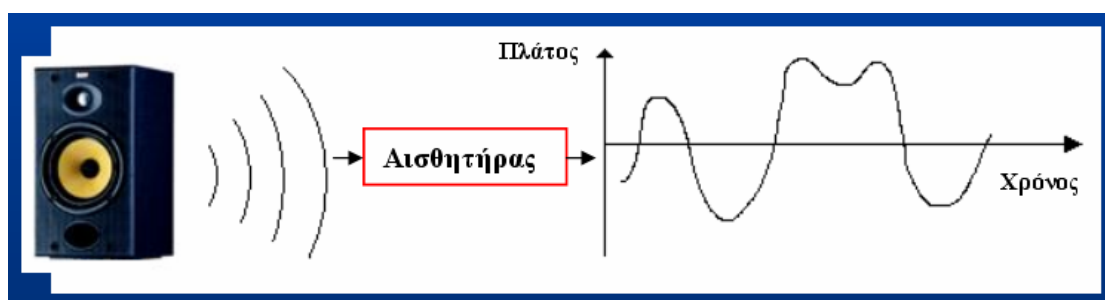
### 1.2 Πληροφορία ως σήμα

Η πληροφορία που αντιλαμβανόμαστε μέσω των αισθήσεων μας και επεξεργάζεται ο εγκέφαλος μας, μπορεί να περιγραφεί ως μια ή περισσότερες φυσικές μεταβλητές η τιμή των οποίων είναι μια συνάρτηση του χρόνου και / ή του χώρου. Να σημειωθεί ότι ως πληροφορία εννοούμε την μορφή της διέγερσης που λαμβάνουμε και όχι το σημασιολογικό περιεχόμενο που αυτή μεταφέρει. Για παράδειγμα, όταν αναφερόμαστε σε ηχητική πληροφορία, η φυσική μεταβλητή περιγράφει την πίεση του αέρα στη θέση ενός παρατηρητή ως συνάρτηση του

χρόνου. Αυτή η ηχητική πληροφορία έχει συνήθως και κάποια ερμηνεία, σημασιολογικό περιεχόμενο. Αν ακούμε μια ομιλία, οι λέξεις και οι ιδέες είναι το σημασιολογικό περιεχόμενο του ήχου.

Αυτή η φυσική μεταβλητή που περιγράφει ένα φαινόμενο, μπορεί να μετρηθεί με κάποιο ειδικά κατασκευασμένο όργανο που ονομάζεται αισθητήρας. Ένας αισθητήρας μετατρέπει αυτή την φυσική ποσότητα, στην περίπτωση του ήχου την πίεση του αέρα, σε μια άλλη ποσότητα, όπως μια ηλεκτρική τιμή, που ονομάζεται σήμα (βλ. Σχήμα 1.2). Αυτό το σήμα είναι τέτοιο ώστε να παριστά το φυσικό μέγεθος με πιστότητα και μπορεί εύκολα να μετρηθεί. Τα σήματα διακρίνονται σε δύο βασικές κατηγορίες:

- Αναλογικό ονομάζεται ένα σήμα το οποίο είναι συνεχής συνάρτηση του χρόνου και / ή του χώρου. Τότε λέμε επίσης ότι το σήμα είναι ανάλογο της φυσική μεταβλητής που περιγράφει.
- Ψηφιακό ονομάζεται ένα σήμα το οποίο αποτελείται από μια ακολουθία διακριτών τιμών που είναι κωδικοποιημένες στο δυαδικό σύστημα και εξαρτώνται από το χρόνο ή το χώρο.



Σχήμα 1.2. Ο αισθητήρας μετατρέπει τον ήχο σε σήμα.

### 1.3 Αναλογικό και ψηφιακό σήμα

Με την ψηφιοποίηση μετατρέπουμε το αναλογικό σήμα σε ψηφιακό. Τα βασικά στάδια της ψηφιοποίησης είναι τα ακόλουθα: α) Δειγματοληψία. Από το άπειρο πλήθος τιμών του συνεχούς σήματος κρατάμε ένα σύνολο διακριτών τιμών (διαφέρουν κατά σταθερό χρονικό διάστημα) β) Κωδικοποίηση. Επιλέγουμε τις στάθμες του αναλογικού σήματος που θα μεταφερθούν στο ψηφιακό σήμα. γ) Κβαντοποίηση. Βρίσκουμε την πλησιέστερη στάθμη κάθε τιμής που προέκυψε από την δειγματοληψία. Στο σχήμα 1.3 δίδεται ένα παράδειγμα ψηφιοποίησης.



Σχήμα 1.3. Διαδικασία ψηφιοποίησης.

Το μεγαλύτερο πλεονέκτημα της ψηφιακής αναπαράστασης είναι η ομοιομορφία. Όπως έχουμε αναφέρει και παραπάνω, όλα τα είδη πληροφορίας μπορούν να έρθουν σε ψηφιακή μορφή και να αντιμετωπισθούν με τον ίδιο τρόπο και από το ίδιο υλικό (ίδια μέσα αποθήκευσης, ίδια δίκτυα...). Αυτό έχει ως συνέπεια τη δυνατότητα χρησιμοποίησης των ίδιων μέσων αποθήκευσης και μετάδοσης δηλαδή την επίτευξη μεγαλύτερου βαθμού ολοκλήρωσης. Να υπενθυμίσουμε σε αυτό το σημείο ότι στην πράξη οι διαφορετικές απαιτήσεις μεγέθους αποθήκευσης και ταχύτητας μετάδοσης των διαφόρων μέσων διαταράσσουν αυτή την ομοιομορφία. Η μετάδοση ψηφιακών σημάτων αντί για αναλογικά έχει πολλά ακόμα πλεονεκτήματα πέραν της ολοκλήρωσης. Είναι λιγότερο ευαίσθητη στον

θόρυβο, η διαδικασία αναγέννησης του μεταδιδόμενου σήματος είναι πιο εύκολη, μπορεί να υλοποιηθεί διαδικασία ανίχνευσης και διόρθωσης λαθών και, τέλος, η κρυπτογράφηση της πληροφορίας είναι επίσης πιο εύκολη.

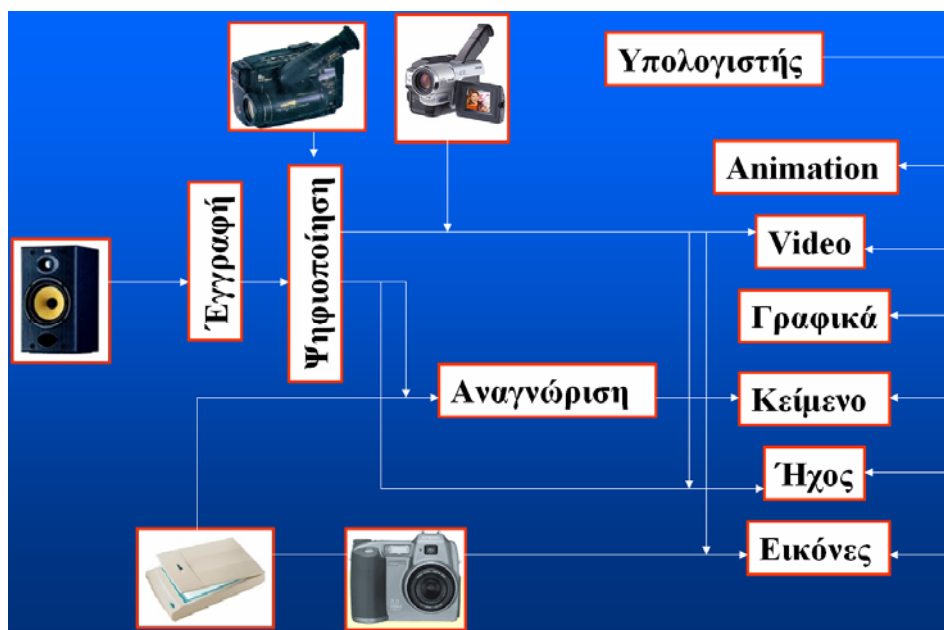
Η πληροφορία που βρίσκεται αποθηκευμένη στον υπολογιστή μπορεί να χρησιμοποιηθεί με διάφορους τρόπους:

- να υποστεί επεξεργασία με στόχο την ανάλυση της σημασιολογίας της ή την βελτίωση της ποιότητας της
- να δημιουργηθούν δομές δεδομένων που επιταχύνουν και διευκολύνουν την αναζήτηση
- να χρησιμοποιηθεί εύκολα για την δημιουργία νέων πολυμεσικών εγγράφων.

Το κύριο μειονέκτημα της ψηφιακής αναπαράστασης συνεχών μέσων είναι η παραμόρφωση που εισάγει η διαδικασία δειγματοληψίας και κβαντοποίησης. Αφενός, αγνοώντας κάποιες τιμές του αναλογικού σήματος χάνουμε πληροφορία και αφετέρου, η προσέγγιση της πραγματικής τιμής του σήματος με μια από τις διαθέσιμες στάθμες περιέχει πάντοτε κάποιο ποσοστό λάθους. Αυτή η παραμόρφωση ελαττώνεται όσο αυξάνεται η συχνότητα δειγματοληψίας και το μήκος της λέξης. Τότε όμως αυξάνεται και ο όγκος που καταλαμβάνει η πληροφορία και κατά συνέπεια απαιτούνται μεγαλύτερα αποθηκευτικά μέσα, πιο γρήγορα μέσα μετάδοσης και ταχύτερες μονάδες επεξεργασίας. Η σημερινή τεχνολογία και οι προβλέψεις για το μέλλον δείχνουν ότι αυτό το μειονέκτημα θα ξεπεραστεί ακόμα και για τους πιο απαιτητικούς τύπους πληροφορίας.

## 1.4 Δημιουργία ψηφιακών εγγράφων

Στο σχήμα 1.4 φαίνονται τα διάφορα ήδη ψηφιακών εγγράφων (animation, video, γραφικά, κείμενο, ήχος, εικόνες) και ο τρόπος δημιουργίας τους. Από μία ψηφιακή φωτογραφική μηχανή μπορούμε να δημιουργήσουμε εικόνες. Με τον scanner ψηφιοποιούμε φωτογραφίες και δημιουργούμε εικόνες ή κείμενο αν χρησιμοποιούμε την διαδικασία αναγνώρισης του κειμένου (OCR). Από μία πηγή ήχου, μέσω του αισθητήρα καταγράφουμε το αναλογικό σήμα, στη συνέχεια το ψηφιοποιούμε και παράγουμε ψηφιακό ήχο ή κείμενο με την βοήθεια της αναγνώρισης (σύστημα αναγνώρισης φωνής). Με την χρήση της βιντεοκάμερας, ψηφιοποιούμε το σήμα αν πρόκειται για αναλογική κάμερα και στη συνέχεια παράγουμε video ή εικόνες και ήχο. Επίσης, με την βοήθεια του υπολογιστή μπορούμε να δημιουργήσουμε κυρίως animation και γραφικά καθώς και να επεξεργαστούμε όλα τα υπόλοιπα ήδη ψηφιακών εγγράφων.



Σχήμα 1.4. Τα διάφορα ήδη ψηφιακών εγγράφων και ο τρόπος δημιουργίας τους.

## 1.5 Εικόνα

Οι εικόνες χωρίζονται σε τρεις βασικές κατηγορίες:

- Στις ασπρόμαυρες ή διτονικές εικόνες όπου έχουμε ύπαρξη μόνο δύο χρωμάτων, συνήθως άσπρου και μαύρου (σχήμα 1.5α). Οι ασπρόμαυρες εικόνες βρίσκουν εφαρμογή σε συστήματα αρχειοθέτησης εγγράφων, σε συστήματα αναγνώρισης κειμένου (OCR), στην διαχείριση τεχνικών σχεδίων, διαγραμμάτων, χαρτών κ.λ.π.
- Στις εικόνες κλίμακας του γκριζου ή gray scale (σχήμα 1.5β). Οι εικόνες αυτές αποδίδονται με διαβαθμίσεις του γκριζου. Μια βασική εφαρμογή των gray scale εικόνων είναι οι ιατρικές εικόνες και τα υπερηχογραφήματα.
- Στις έγχρωμες εικόνες όπου έχουμε απόδοση της εικόνας με χρώματα (σχήμα 1.5γ). Ο αριθμός των χρωμάτων μπορεί να ποικίλει από 256 μέχρι εκατομμύρια χρώματα.

περιοχών κειμένου. Εξ' αιτίας της μεγάλης που προβλημάτων επεξεργασίας εικόνας και τον εσφαλμένε κειμένου οι περισσότερες εργασίες για την ΑΛΔ εγχει κωδία όπως η αναγνώριση ταχυδρομικών φακέλων ή Γενικά μια μέθοδος για την ΑΔΣ αποτελείται από δύο βήτ τμηματοποίηση του εγγράφου σε οπτικά διακριτά παξινόμηση κάθε μπλοκ.

Οι προτεινόμενες τεχνικές ΑΔΣ μπορούν να τε τρόπο τμηματοποίησης που χρησιμοποιούν, ως "από πά ή *pattern driven*" [PAZH92], [WONWS2], [WAWCS "από κάτω προς τα πάνω", (*bottom-up* ή *data dri* ICGOR93] και *εθελούσι-* [FAI W94], [FARR90] [13]



(α)



(β)



(γ)

**Σχήμα 1.5.** Παραδείγματα εικόνων: (α) ασπρόμαυρη, (β) gray scale, (γ) έγχρωμη.

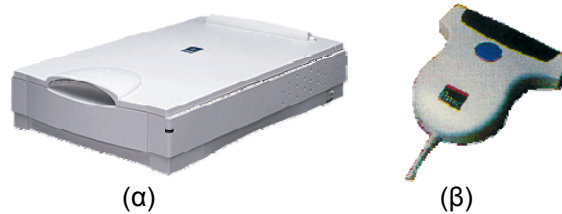
Βασικά χαρακτηριστικά των εικόνων είναι οι διαστάσεις (οριζόντιος και κάθετος αριθμός εικονοστοιχείων) και η ανάλυση ψηφιοποίησης η οποία μετρείται σε dpi (dots per inch). Η ανάλυση της εικόνας δίνει τον αριθμό των εικονοστοιχείων στον οποίο αντιστοιχίζεται μία τετραγωνική ίντσα της εικόνας που ψηφιοποιείται. Η ρύθμιση της ανάλυσης εξαρτάται από την χρήση της ψηφιοποιημένης εικόνας. Μερικά παραδείγματα ανάλυσης είναι τα 400-600dpi για αρχειοθέτηση εικόνων με μεγάλη πληροφορία, τα 300-400dpi για αναγνώριση των εγγράφων με χρήση της οπτικής αναγνώρισης χαρακτήρων (OCR), τα 150-300dpi για εικόνες που θα μεταδοθούν στο διαδίκτυο, 75 dpi για εικόνες που θα εμφανιστούν στην οθόνη του υπολογιστή. Άλλα χαρακτηριστικά των εικόνων είναι το πρότυπο αποθήκευσης (format) και ο τρόπος συμπίεσης.

Υπάρχει μεγάλη ποικιλία στα πρότυπα αποθήκευσης των εικόνων. Τα κυριότερα πρότυπα είναι τα ακόλουθα:

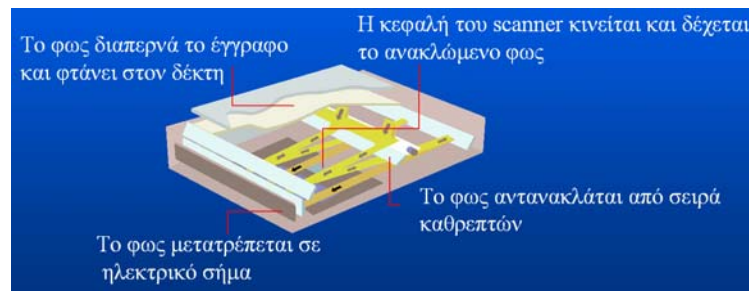
- **TIF πρότυπο.** Είναι το πιο διαδεδομένο πρότυπο. Είναι το πιο δημοφιλές format για ψηφιακή αποθήκευση. Βγάζει σχετικά μεγάλο αρχείο για τις έγχρωμες εικόνες, χρησιμοποιεί LZW συμπίεση χωρίς απώλειες και έχει πολλές παραλλαγές οι οποίες δεν διαβάζονται από όλα τα προγράμματα. Το TIF πρότυπο χρησιμοποιείται συνήθως για ασπρόμαυρες εικόνες αφού πετυχαίνει μεγάλα ποσοστά συμπίεσης (μέθοδος συμπίεσης, CCITT Group 4). Ένα άλλο χαρακτηριστικό των TIF εικόνων είναι ότι αποθηκεύουν και αρκετές πληροφορίες σχετικά με τις εικόνες (δημιουργός, ημερομηνία τελευταίας αλλαγής κ.α.).
- **GIF πρότυπο.** Το GIF πρότυπο αποθήκευσης επιτρέπει το πολύ 256 χρώματα και εφαρμόζεται κυρίως για απλές εικόνες, για παράδειγμα ένα λογότυπο εταιρείας, έχει καλή συμπίεση, όμως η χρήση της GIF κωδικοποίησης έχει πατέντα και χρειάζεται άδεια για την χρήση του από τα διάφορα προγράμματα.
- **JPEG πρότυπο.** Το JPEG πρότυπο είναι το πιο συνηθισμένο για δημιουργία μικρών αρχείων ιδανικών για αποστολή με e-mail ή μεταφορά μέσω internet. Χρησιμοποιεί συμπίεση με απώλειες (ρύθμιση του παράγοντα JPG ποιότητας). Είναι σχεδιασμένο για εικόνες με συνεχείς τόνους ενώ δεν έχει καλά αποτελέσματα για εικόνες με ακμές και γωνίες, για κείμενα, λογότυπα κ.λ.π.

## Σάρωση

Οι ψηφιακές εικόνες παράγονται κυρίως με την βοήθεια της σάρωσης. Δύο βασικές κατηγορίες σαρωτών είναι οι επιτραπέζιοι σαρωτές και οι σαρωτές χειρός (σχήμα 1.6). Στον επιτραπέζιο σαρωτή τοποθετούμε το έγγραφο στο εσωτερικό του σαρωτή ενώ τον σαρωτή χειρός τον σέρνουμε πάνω από την εικόνα που θέλουμε να ψηφιοποιήσουμε. Ένας σαρωτής αποτελείται από τρία βασικά μέρη (σχήμα 1.7). Την πηγή φωτός, τον χώρο τοποθέτησης του εγγράφου και τον ανιχνευτή του φωτός. Το φως ανακλάται από σειρά καθρεπτών, διαπερνά το έγγραφο και φτάνει στον δέκτη. Ο δέκτης βρίσκεται στην κεφαλή του scanner η οποία κινείται και δέχεται το ανακλώμενο φως. Το φως αυτό μετατρέπεται σε ηλεκτρικό σήμα το οποίο μετατρέπεται σε ψηφιακό σήμα εικόνας.



Σχήμα 1.6. Είδη σαρωτών: (α) επιτραπέζιος, (β) χειρός.



Σχήμα 1.7. Λειτουργία σαρωτή.

## Δημιουργία ψηφιακής βιβλιοθήκης

Οι ψηφιακές εικόνες σε μία ψηφιακή βιβλιοθήκη μπορούν να κρατούνται σε διάφορα αντίγραφα καθένα από τα οποία έχει διαφορετικά χαρακτηριστικά. Για παράδειγμα στην ψηφιακή βιβλιοθήκη της California (California Digital Library) έχουμε ένα βασικό αντίγραφο (Master file) το οποίο είναι υψηλής ανάλυσης (600dpi), gray scale ή ασπρόμαυρο και με κωδικοποίηση TIF, ένα αρχείο πρόσβασης (Access file) το οποίο είναι πιο χαμηλής ανάλυσης, με ενδιάμεση ποιότητα, gray scale ή έγχρωμο και με κωδικοποίηση JPEG, έχουμε τα thumbnails τα οποία είναι μικρά εικονίδια GIF gray scale 4 bit ή έγχρωμο 8 bit, και τέλος υπάρχει ένα αντίγραφο για εκτύπωση σε μορφή pdf (σχήμα 1.8). Για να αποθηκευτούν όλα τα αντίγραφα χρειάζονται βέβαια μεγάλη χωρητικότητα αποθηκευτικών μέσων. Εναλλακτικά τα διάφορα είδη των εικόνων μπορεί να προκύπτουν με αλγορίθμους μετασχηματισμού κατά την διάρκεια της πρόσβασης στην πληροφορία. Αυτό επιβαρύνει την ταχύτητα πρόσβασης στην ψηφιακή βιβλιοθήκη.

## 1.6 Κείμενο

Ο πιο διαδεδομένος τύπος κειμένου είναι ο ASCII. Αναπτύχθηκε από τον Αμερικανικό Οργανισμό πιστοποίησης και είναι απλή κωδικοποίηση χαρακτήρων με χρήση 7 bits. Περιλαμβάνει τους λατινικούς χαρακτήρες, τα νούμερα και ένα σύνολο ειδικών συμβόλων. Οι κωδικοσελίδες ISO καλύπτουν την αδυναμία του ASCII κειμένου να υποστηρίξει γλώσσες με μη λατινικό αλφάβητο. Έχουμε επίσης το markup κείμενο το οποίο είναι κείμενο με οδηγίες παρουσίασης και χρησιμοποιείται στις περιπτώσεις περίπλοκων κειμένων όπως τα κείμενα με πίνακες και εξισώσεις. Τέλος, στους βασικούς τύπου κειμένου συγκαταλέγεται και το Hypertext όπου δεν υπάρχει η γραμμικότητα του παραδοσιακού κειμένου. Το κείμενο

διαμερίζεται σε κόμβους (nodes) που συνδέονται μεταξύ τους δημιουργώντας την αναπαράσταση του γράφου.

Η συμπίεση του κειμένου βασίζεται κυρίως σε τεχνικές αντικατάστασης προτύπων και στην Huffman κωδικοποίηση. Τέλος, ιδιαίτερο ενδιαφέρον παρουσιάζει η κωδικοποίηση κειμένου όταν δεν θέλουμε την δημοσιοποίησή του. Χρησιμοποιείται το πρότυπο DES (Data Encryption Standard) το οποίο δημιουργήθηκε από την IBM και γίνεται χρήση δημοσίου και ιδιωτικού κλειδιού. Σύμφωνα με αυτή την τακτική, χρησιμοποιούμε ένα δημόσιο κλειδί (public key) για την κωδικοποίηση του κειμένου το οποίο δημοσιοποιείται. Το κείμενο πάει στον παραλήπτη και ο παραλήπτης χρησιμοποιεί ένα ιδιωτικό κλειδί (private key) που έχει στην κατοχή του για την αποκρυπτογράφηση του κειμένου.

<b>Master Files</b>	<ul style="list-style-type: none"> <li>• 600 pixels per inch (PPI) capture resolution; less for oversize originals (see Alternative Minimum below)</li> <li>• TIFF, lossless compression</li> <li>• 8-bit greyscale; 24-bit color; (bitonal for typeset pages with typeface 7 pt and above)</li> </ul>
<b>Alternative Minimum for Originals Larger than 8.5"x11"</b>	<ul style="list-style-type: none"> <li>• 300 PPI (otherwise, same as main guidelines above)</li> </ul>
<b>Access Files</b>	<ul style="list-style-type: none"> <li>• 800, 1500 or 3000 pixels across the long dimension</li> <li>• JPEG, medium quality compression</li> <li>• 8-bit greyscale; 24-bit color</li> </ul>
<b>Alternative Access File Format for Better Text Legibility</b>	<ul style="list-style-type: none"> <li>• 100 PPI (i.e., resample image master file to 100 pixels per inch of original document)</li> <li>• GIF; 4-bit greyscale; 8-bit color; or</li> <li>• JPEG, higher quality compression</li> </ul>
<b>Thumbnail Files</b>	<ul style="list-style-type: none"> <li>• 200-400 pixels across the long dimension</li> <li>• GIF</li> <li>• 4-bit grayscale, 8-bit color</li> </ul>
<b>Print</b>	<ul style="list-style-type: none"> <li>• 300-600 DPI</li> <li>• PDF or TIFF w/LZW compression for B&amp;W materials, otherwise, JFIF - medium to high quality</li> <li>• 8-bit greyscale; 24-bit color</li> </ul>

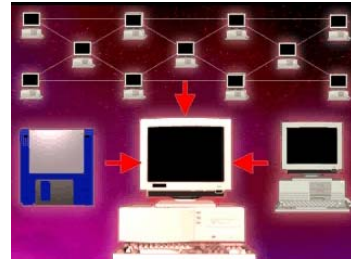
**Σχήμα 1.8.** Αποθήκευση εικόνων στην ψηφιακή βιβλιοθήκη.

## 1.7 Ήχος

Ο ήχος μπορεί να είναι ένα ψηφιοποιημένο ηχητικό σήμα ή ένα σήμα με σημασιολογική σημασία όπως το πρότυπο MIDI όπου υπάρχει κωδικοποίηση της παρτιτούρας και των οργάνων. Οι πιο διαδεδομένοι τύποι (formats) ήχου είναι το WAV, το οποίο είναι ψηφιοποιημένο ηχητικό σήμα, το MP3 και το MID που είναι αρχεία με σημασιολογική πληροφορία ήχου. Στην επεξεργασία των εγγράφων ήχου, ιδιαίτερο ενδιαφέρον παρουσιάζει η τεχνολογία αναγνώρισης ομιλίας με την οποία μετατρέπουμε τον ήχο σε αντικείμενο κειμένου το οποίο μπορούμε να επεξεργαστούμε, να αναζητήσουμε κ.λ.π. Ένα βασικός διαχωρισμός αυτών των συστημάτων είναι τα συστήματα τα οποία είναι εκπαιδευμένα μόνο σε συγκεκριμένο ομιλητή και σε εκείνα που είναι ανεξάρτητα ομιλητή.

## 1.8 Γραφικά

Μία μορφή εικόνων είναι και τα γραφικά. Η διαφορά τους είναι ότι περιέχουν πληροφορίες για την δομή των αντικειμένων, δεν είναι δηλαδή ένα σύνολο εικονοστοιχείων. Ένα χαρακτηριστικό των γραφικών είναι ότι χρησιμοποιούν διάφορα μοντέλα για την περιγραφή τους. Τα μοντέλα αυτά μπορεί να είναι γεωμετρικά, στερεά, φυσικά κ.λ.π. Στο σχήμα 1.9 φαίνονται δύο παραδείγματα γραφικών, εικόνων κατασκευασμένων με την βοήθεια του ηλεκτρονικού υπολογιστή.



Σχήμα 1.9. Παραδείγματα γραφικών.

Τα γραφικά μας δίνουν μεγάλη ευχέρεια παρεμβάσεων. Μπορούμε να έχουμε εισαγωγή, αφαίρεση, μετακίνηση, στροφή ή μεγέθυνση των αντικειμένων. Μπορούμε να έχουμε σκίαση των αντικειμένων ή αντιστοίχιση κάποιας εικόνας στην επιφάνεια ενός αντικειμένου (δημιουργία υφής), αλλαγή του φωτισμού κ.λ.π. Με την βοήθεια του rendering τα διάφορα μοντέλα γραφικών μετατρέπονται σε εικόνες με συγκεκριμένη ανάλυση, χρώμα, μέγεθος κ.λ.π.

## 1.9 Video

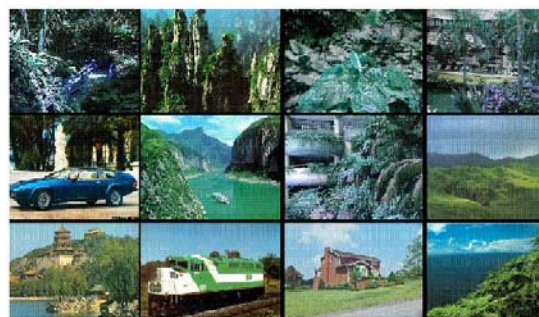
Τα video είναι από τα πιο διαδεδομένα ψηφιακά έγγραφα. Μπορεί να είναι ψηφιοποιημένα video ή animation. Τα χαρακτηριστικά ενός video είναι η συχνότητα δειγματοληψίας, η ταχύτητα εναλλαγής των πλαισίων, το μέγεθος και το χρωματικό βάθος της εικόνας. Τα κυριότερα video formats είναι τα: mpeg, avi, mov, flc, fli, wmv, qt.

Δύο βασικές μέθοδοι συμπίεσης των videos είναι η MPEG και η συμπίεση με Wavelets. Το πρότυπο MPEG αποτελείται από έναν αριθμό διαφορετικών προτύπων και εκδόσεων (MPEG 1, MPEG 2, MPEG 4) για κάλυψη πολλαπλών αναγκών. Η λειτουργία του MPEG ακολουθεί τις ίδιες αρχές με το JPEG, αλλά απλώς εφαρμόζεται σε κινούμενες εικόνες ή video. Η συμπίεση με wavelets προκύπτει από διακριτούς μετασχηματισμούς Fourier (DCT) με επιλογή συγκεκριμένων στοιχείων συχνότητας. Το κάθε στοιχείο συχνότητας μελετάται με ανάλυση (resolution) αντίστοιχη της κλίμακάς του. Εισάγονται μεν απώλειες κατά την συμπίεση, όμως επιτυγχάνονται μεγάλοι ρυθμοί συμπίεσης. Πιο συγκεκριμένα, τα wavelets μπορούν να επιτύχουν μέχρι και 10 φορές μεγαλύτερη συμπίεση από τα MPEG-1, MPEG-2 για την ίδια ποιότητα εικόνας βίντεο.

Η επεξεργασία του video περιλαμβάνει την εξαγωγή του κειμένου είτε αναγνωρίζοντας τις εικόνες (σχήμα 1.10) είτε αναγνωρίζοντας τον ήχο, την αυτόματη εξαγωγή περιγραφικών χαρακτηριστικών για την ομαδοποίηση εικόνων με παρόμοιο θέμα ή χρώμα (σχήμα 1.11) κ.λ.π.



Σχήμα 1.10. Εξαγωγή κειμένου από video



Σχήμα 1.11. Εικόνες παρόμοιου θέματος

## 1.10 Αποθηκευτικά μέσα

Για την επιλογή των μέσων που θα αποθηκεύσουμε τα ψηφιακά έγγραφα πρέπει να λάβουμε υπόψη μας τους παρακάτω παράγοντες:

- Την ποσότητα αποθήκευσης
- Τον απαιτούμενο χρόνο προσπέλασης
- Το είδος πληροφορίας
- Τον ρυθμό με τον οποίο λαμβάνεται και αλλάζει η πληροφορία
- Τον αριθμό αντιγράφων
- Την επιθυμητή διάρκεια ζωής
- Το κόστος προετοιμασίας για την αποθήκευση
- Το κόστος αποθηκευτικών μέσων
- Την αναγκαιότητα προσπέλασης από υπάρχοντα μηχανήματα

Τα αποθηκευτικά μέσα χωρίζονται σε δύο βασικές κατηγορίες: Στα μαγνητικά αποθηκευτικά μέσα τα οποία είναι κατάλληλα για δυναμικά δεδομένα που απαιτούν συχνές αλλαγές και προσπελάσεις. Χρησιμοποιούνται κατά την επεξεργασία των δεδομένων. Το σύστημα RAID (Random Arrays of Inexpensive Disks) επιτρέπει την κατανομή των δεδομένων σε πολλούς δίσκους. Ο τρόπος σύνδεσης των δίσκων εξαρτάται από τις επιδόσεις και το κόστος του συστήματος και την διαθεσιμότητα των δεδομένων. Στα οπτικά αποθηκευτικά μέσα τα οποία χρησιμοποιούνται για την αρχειοθέτηση των δεδομένων. Είναι τα CD-ROM (Compact Disk – Read Only Memory), CD-R, Photo CD (Kodak), DVD (Digital Video Disk). Στο δίσκο υπάρχουν κοιλάδες (pits) και νησίδες (lands) που χαράσσονται σε ένα πλαστικό υλικό πάνω από μία ανακλαστική επιφάνεια σχηματίζοντας σπείρα από το κέντρο του δίσκου προ τα έξω. Η μετάβαση από κοιλάδα σε νησίδα (και το αντίστροφο) αντιστοιχεί στο 1. Η ακτίνα laser ανακλάται (νησίδες) και ανιχνεύεται από έναν φακό δημιουργώντας ψηφιακό σήμα.

## 1.11 Συστήματα διαχείρισης εγγράφων (EDMS)

Τα συστήματα διαχείρισης εγγράφων (EDMS – Electronic Document Management Systems) προσφέρουν ψηφιοποίηση, κεντρική αποθήκευση, διαχείριση, δεικτοδότηση και αναζήτηση στα ψηφιακά έγγραφα. Τα EDMS οργανώνουν δυναμικά και ελέγχουν την διανομή εγγράφων και πληροφοριών, αλληλεπιδρούν και οργανώνουν την επαγγελματική δραστηριότητα στην οποία χρησιμοποιούνται τα έγγραφα, αποτελούν βασικό κορμό των υποδομών για κάθε ηλεκτρονική δραστηριότητα (e-business) που βασίζεται στα ψηφιακά έγγραφα. Μπορούν να διαχειριστούν πληροφορία σε μορφή: χαρτιού, εικόνες, εφαρμογών, κειμένων, ήχου, video, e-mails, ιστοσελίδων κ.λ.π.

Τα βασικά τμήματα ενός ολοκληρωμένου EDMS είναι τα ακόλουθα:

- **Hardware.** Απαραίτητη είναι η ύπαρξη σαρωτή και ενός server εικόνων ενώ προαιρετική είναι η ύπαρξη μέσων αποθήκευσης αρχειακού υλικού, οπτικά jukebox, υψηλής ανάλυσης monitors.
- **Software.** Σύστημα επεξεργασίας και κατανόησης εγγράφων, σύστημα διαχείρισης εγγράφων, σύστημα πρόσβασης και προβολής των εγγράφων. Προαιρετικό είναι το σύστημα παρακολούθησης της ροής των ενεργειών (workflow).

## 1.12 Εισαγωγή στην επεξεργασία και κατανόηση των εγγράφων

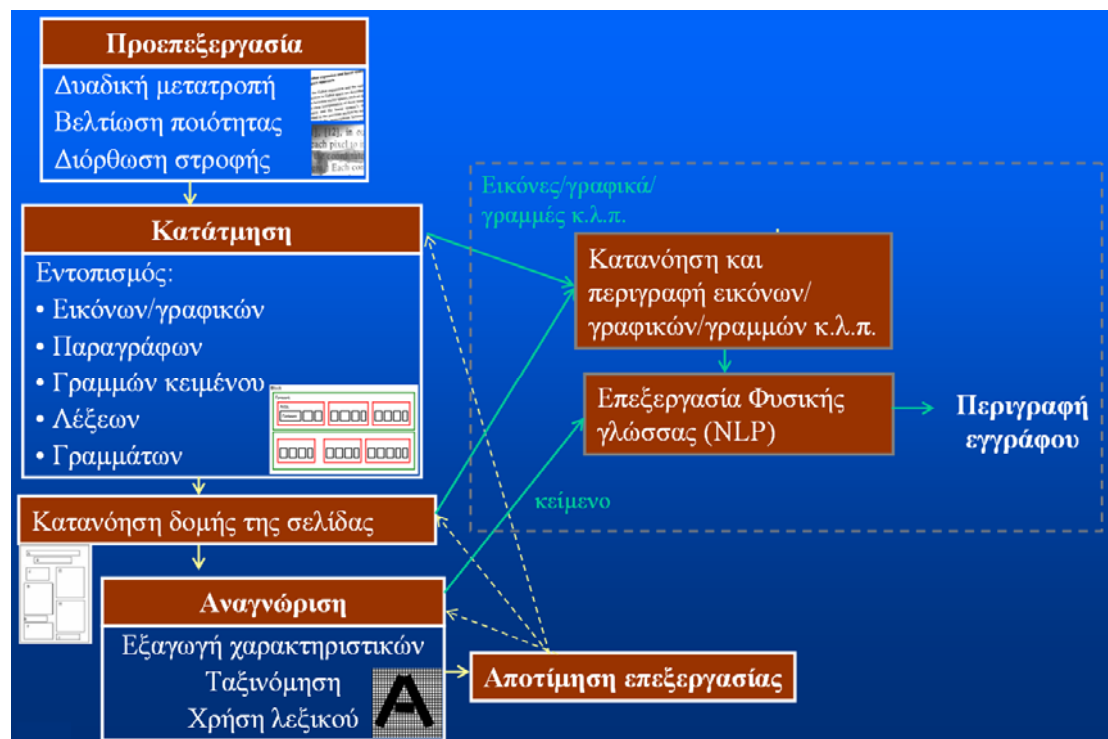
Τα έγγραφα περιέχουν γνώση. Πιο συγκεκριμένα, τα έγγραφα αποτελούν το μέσο μεταφοράς της γνώσης. Η εξαγωγή της γνώσης από ένα έγγραφο το οποίο μπορεί να είναι τεχνική



αναφορά, ανακοίνωση, εφημερίδες, βιβλία, περιοδικά, γράμματα, τραπεζικές επιταγές κ.λ.π. είναι το πιο κρίσιμο σημείο ενός πληροφοριακού συστήματος. Η εξαγωγή της γνώσης προϋποθέτει χρόνο και ανθρώπινη προσπάθεια, στοιχεία τα οποία περιορίζουν τις δυνατότητες του πληροφοριακού συστήματος. Έτσι, η αυτοματοποίηση της εξαγωγής της γνώσης από τα έγγραφα αποτελεί ένα σημαντικό ερευνητικό τομέα με πολλές εφαρμογές. Ήδη από την δεκαετία του 1960 η έρευνα σχετικά με την αυτόματη επεξεργασία των εγγράφων ξεκίνησε με την οπτική αναγνώριση των χαρακτήρων (OCR). Τα βασικά στάδια ενός συστήματος επεξεργασίας και κατανόησης εγγράφων (σχήμα 1.12) είναι τα ακόλουθα:

- **Προεπεξεργασία.** Περιλαμβάνει την δυαδική μετατροπή (μετατροπή της gray scale εικόνας σε ασπρόμαυρης), την βελτίωση της ποιότητας (εξάλειψη θορύβου, βελτίωση της ποιότητας του κειμένου) και την διόρθωση της στροφής της εικόνας (διόρθωση στροφής της εικόνας η οποία έχει προκύψει λόγω μη ευθυγραμμισμένης τοποθέτησης του εγγράφου στον σαρωτή).
- **Κατάτμηση.** Εντοπισμός των βασικών συστατικών του εγγράφου (εικόνες, γραφικά, τμήματα κειμένου, παράγραφοι, γραμμές κειμένου, λέξεις, γράμματα).
- **Κατανόηση δομής της σελίδας.** Χαρακτηρισμός των βασικών συστατικών του εγγράφου καθώς και εντοπισμός της ροής του κειμένου. Για παράδειγμα, αν το έγγραφο είναι εφημερίδα τα βασικά συστατικά μπορεί να χαρακτηρίζονται ως τίτλος, υπότιτλος, υπέρτιτλος, συγγραφέας, στήλη, εικόνα, λεζάντα κ.λ.π. ενώ αν το έγγραφο είναι σελίδα τεχνικού περιοδικού τα βασικά συστατικά μπορεί να χαρακτηρίζονται ως ονομασία περιοδικού, τίτλος, συγγραφέας, περίληψη, στήλες κειμένου κ.λ.π.
- **Αναγνώριση.** Περιλαμβάνει την εξαγωγή αξιόπιστων χαρακτηριστικών για κάθε χαρακτήρα και την κατάστρωση ενός ταξινομητή για την κατάταξη του κάθε χαρακτήρα σε γνωστή κλάση γράμματος. Συνήθως η αναγνώριση περιλαμβάνει και διόρθωση του τελικού αποτελέσματος με χρήση λεξικού.

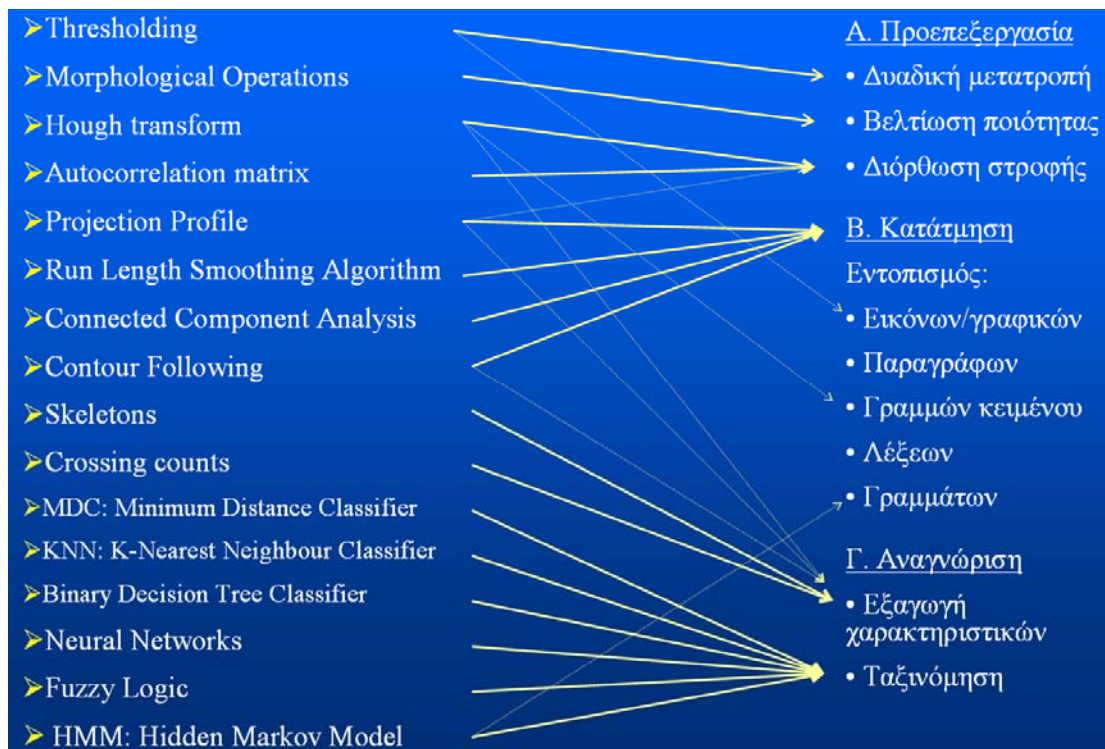
Για την σωστή ανάπτυξη ενός ολοκληρωμένο συστήματος επεξεργασίας και κατανόησης εγγράφων απαιτείται και ένα στάδιο αυτόματης αποτίμησης των διάφορων σταδίων της επεξεργασίας. Για το σκοπό δημιουργούμε ένα σύνολο εικόνων ελέγχου το οποίο συνοδεύεται από τα σωστά αποτελέσματα κατάτμησης και αναγνώρισης. Συγκρίνοντας με τα αποτελέσματα του συστήματος μπορούμε να αποτιμήσουμε αυτόματα το σύστημα εξάγοντας πληροφορίες για το ποσοστό και την ακρίβεια κατάτμησης, κατανόησης και αναγνώρισης.



Σχήμα 1.12. Επεξεργασία και κατανόηση εγγράφων

Για την ολοκληρωμένη περιγραφή του εγγράφου ένα σύστημα επεξεργασίας και κατανόησης εγγράφων μπορεί να περιέχει και ένα σύστημα κατανόησης των εικόνων και των γραφικών (κατηγοριοποίηση σε πορτρέτα, τοπία, σχήματα, σχέδια κ.λ.π.) καθώς και ένα σύστημα επεξεργασίας της φυσικής γλώσσας (NLP, Natural Language Processing) ώστε να γίνεται σημασιολογική ανάλυση της αναγνωρισμένης πληροφορίας με σκοπό την καλύτερη δεικτοδότηση και αναζήτηση της.

Για την υλοποίηση όλων των παραπάνω σταδίων χρησιμοποιούνται διάφορες τεχνικές επεξεργασίας εικόνας και αναγνώρισης προτύπων (σχήμα 1.13). Όπως φαίνεται και από το σχήμα, μερικές από τις τεχνικές αυτές μπορεί να φανούν χρήσιμες σε παραπάνω από μία φάσεις της επεξεργασίας και αναγνώρισης. Για παράδειγμα, ο μετασχηματισμός Hough της εικόνας μπορεί να εφαρμοστεί για την εύρεση της στροφής του εγγράφου, για τον εντοπισμό των γραμμών του κειμένου και της εξαγωγή χαρακτηριστικών.



Σχήμα 1.13. Βασικές τεχνικές επεξεργασίας εικόνας και αναγνώρισης προτύπων.

## 1.13 Ερευνητικές ομάδες σχετικές με επεξεργασία και κατανόηση εγγράφων



<http://www.cenparmi.concordia.ca/>

**CENPARMI:** Centre for Pattern Recognition and Machine Intelligence at Concordia University (Montréal, Québec, Canada). Pattern recognition and machine intelligence (PARMI) has been a very active field in the past 20 years, particularly in the application area. PARM techniques have been used to recognize characters, time-varying signals, voice, satellite pictures, objects, fingerprints, and weather patterns, in many scientific, business, and military applications.



<http://www.cedar.buffalo.edu/>

**CEDAR:** Center of Excellence for Document Analysis and Recognition (NY, USA). Research at CEDAR focuses on the theory and applications of pattern recognition, machine learning, and information

retrieval. Over the years, the applications explored have included document analysis and recognition, object recognition in images in three-dimensional scenes, forensic document examination, textual information retrieval, biometrics, and bio-informatics.



<http://www.cse.salford.ac.uk/prima/>



<http://www-dii.ing.unisi.it/research/CS/docproc/home.html>

**PRIMA:** Pattern Recognition and Image Analysis group (Salford, UK). Document Image Analysis / Understanding, Layout Analysis, WWW Document Image Analysis, Performance Evaluation, Drawing / Chart Recognition, Character Recognition

**DANTE:** Document ANalysis and TEchnologies research group (Italy). Form understanding, Invoice understanding, Form registration, Logo recognition, Document model, Word recognition, OCR Optical Character Recognition, Handwritten character recognition.



<http://www.loria.fr/equipes/read/>

**READ:** REcognition of writing and Analysis of Documents (Loria Lab, France). The scientific objective of the READ team is to study the writing modeling in its various aspects: printed, handwritten text and composite document. For handwriting, particular investigations are made in stochastic modeling by the use of Markov chains and fields. For printed documents, the efforts are related to structure modeling as well as to analysis strategies and content comprehension. (Document Analysis and retroconversion)



<http://lamp.cfar.umd.edu/>

**LAMP:** Language and Media Lab (Maryland, USA). Duplicate Document Image Detection, Functional Document Description, Intelligent Document Image Retrieval, Page Decomposition and Structural Analysis, Signature Verification, Zone Classification (Decision Tree), Arabic OCR Evaluation, Automatic Ground-Truth Generation, Document Image Enhancement, Media Adaptation, Recognition of people in low quality video



<http://dsp.di.uoa.gr/>

**SIPL:** Signal & Image Processing Laboratory (Univ. of Athens). Pattern Recognition, Image Restoration and Compression, Image Segmentation, Character Recognition (OCR) - Byzantine Music Recognition.



<http://www.iit.demokritos.gr/~bgat/>

**CIL:** Computational Intelligence Laboratory (NCSR "DEMOKRITOS"). Image Processing, Pattern Recognition, Document Image Analysis – OCR, Processing and recognition of Historical Documents, Web Document Analysis



**AIA:** Artificial Intelligence Information Analysis (Aristotle Univ. - Thessaloniki). Digital Image Processing and Analysis, video processing, watermarking, morphological filters, Neural networks.

<http://poseidon.csd.auth.gr/>

**IPML**

<http://ipml.ee.duth.gr/~papamark/>

**IPML:** Image Processing and Multimedia Lab (Univ. of Thrace). Image Processing, Computer Vision, Document Processing and Analysis, Digital Filter Design, Optimization Hough Transform, Neural Networks, Color Quantization, Image Retrieval, Fuzzy Clustering. OCR, segmentation, skew correction, morphological operations.

## 1.14 Σχετικά περιοδικά και συνέδρια

- **Pattern Recognition:** The Journal of the Pattern Recognition Society. Publisher: Pergamon - Elsevier.  

- **Pattern Recognition Letters** Publisher: Elsevier.  

- **International Journal on Document Analysis and Recognition (IJAR)** Publisher: Springer.  

- **IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)** Publisher: IEEE Computer Society.  

- **Pattern Analysis & Applications(PAA)** Publisher: Springer.  

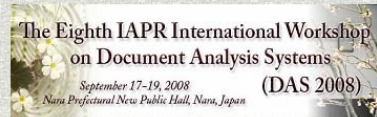

- **CVPR 2008:** IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 24-26 June 2008, Anchorage, Alaska. (submission due 3/12/2007)



- **ICFHR 2008:** 11th International Conference on Frontiers in Handwriting Recognition. 19-21 August 2008, Montreal, Canada. (submission due 31/1/2008)



- **DAS 2008:** 8th International Workshop on Document Analysis Systems. 17-19 September 2008, Nara, Japan. (submission due 1/3/2008)



- **ICIP 2008:** 2008 IEEE International Conference on Image Processing. 12-15 October 2008, San Diego, CA.



- **ICPR 2008:** 19th International Conference on Pattern Recognition. 8-11 December 2008, Tampa, Florida, USA. (submission due 8/4/2008)



- **ICDAR 2009:** 10th International Conference on Document Analysis and Recognition. 26-29 July 2009, Barcelona, Spain. (submission due 12/1/2009)

