

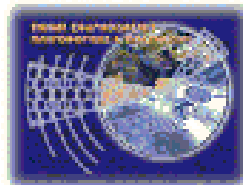


**ΤΕΧΝΟΛΟΓΙΚΟ  
ΕΚΠΑΙΔΕΥΤΙΚΟ  
ΙΔΡΥΜΑ ΚΡΗΤΗΣ**

# **Τεχνολογικό Εκπαιδευτικό Ίδρυμα Κρήτης**

**Σχολή Τεχνολογικών Εφαρμογών**

**Τμήμα Εφαρμοσμένης Πληροφορικής & Πολυμέσων**



**ΤΜΗΜΑ  
ΕΦΑΡΜΟΣΜΕΝΗΣ  
ΠΛΗΡΟΦΟΡΙΚΗΣ  
& ΠΟΛΥΜΕΣΩΝ**

**Πτυχιακή εργασία**

**Τίτλος: Εργαλεία Σημαιολογικού Σχολιασμού Εικόνων**

**(Image Annotation Tools)**

**Θεοδώρα Κουτσού (ΑΜ:2085)**

**Επιβλέπων καθηγητής: Τριανταφυλλίδης Γιώργος**

**Επιτροπή Αξιολόγησης:**

**Ημερομηνία Παρουσίασης: 4/3/2011**

## **Ευχαριστίες**

*Ολοκληρώνοντας την παρούσα πτυχιακή εργασία θα ήταν παράληψη να μην*

*ευχαριστήσω τον επιβλέποντα καθηγητή μου κ.Τριανταφύλλιδη Γιώργο για την εποπτεία*

*και την σωστή καθοδήγηση που μου παρείχε. Όπως επίσης στους γονείς μου, τους φίλους και όσους με υποστήριξαν στην προσπάθεια μου μέχρι σήμερα.*

## **Abstract**

The aim of this dissertation is to give a comprehension of various multimedia concepts, as well as, their applications. Concepts will be briefly analysed; our intention is to give comprehensive and a proper description of the most significant image animation tools. In further detail we will analyse software multimedia like the **“Image annotation”, “Anvil”, “Annotate Expert”, “FotoTagger”, “M-OntoMat-Annotizer”, “Annotea”, “Video Image Annotation”, “LabelMe”, “FotoNotes” and lastly the “IAT”**. This selection was based on their popularity, usefulness and according to people needs. Consequently, a reference to lots of image annotation tools is described to show that not only a simple user but also an expert could be able to use those tools. A brief but comprehensive presentation took place as the main purpose was not to indicate how each tool actually works but what their capabilities and characteristics are. Lastly, one of the image animation tools, “Label Me”, was chosen to be studied deeper through “Matlab”. “Label Me”, has various capabilities and it is an open – code tool that can be easily changed, hence configured.

## ΣΥΝΟΨΗ

Σκοπός της παρούσας πτυχιακής εργασίας είναι η κατανόηση της έννοιας των πολυμέσων καθώς επίσης και των εφαρμογών τους. Όλα αυτά βέβαια σε ένα γενικότερο πλαίσιο, καθώς για εμάς στόχος είναι η κατανόηση και η περιγραφή των πιο σημαντικών εργαλείων σημασιολογικού σχολιασμού εικόνων και πως λειτουργούν. Τα εργαλεία τα οποία παρουσιάζονται σε αυτή την εργασία είναι το Image annotation Tool, το Anvil, το Annotate Expert, το FotoTagger, το M-OntoMat-Annotizer, το Annotea, το Video Image Annotation Tool, το LabelMe, το FotoNotes και το IAT. Η επιλογή τους έγινε με βάση την δημοτικότητα, την χρηστικότητά τους αλλά και τους διαφορετικούς σκοπούς που μπορεί να εξυπηρετεί το καθένα. Έτσι γίνεται αναφορά σε αρκετά απλά εργαλεία σχολιασμού εικόνας τα οποία μπορεί να χρησιμοποιήσει ακόμα και ένας χρήστης ο οποίος δεν είναι εξοικειωμένος αλλά και σε πιο εξειδικευμένα εργαλεία. Η παρουσίαση κάθε εργαλείου γίνεται συνοπτικά αλλά περιεκτικά καθώς σκοπός μας δεν ήταν πώς λειτουργεί το κάθε εργαλείο αλλά ποιες οι δυνατότητές του και ποια τα βασικά χαρακτηριστικά του. Τέλος, ένα από τα εργαλεία σχολιασμού εικόνων, το Label Me επιλέχθηκε προκειμένου να μελετηθεί περαιτέρω στο matlab. Η επιλογή έγινε λόγω των δυνατοτήτων που μας προσφέρει το LabelMe καθώς και γιατί είναι ένα εργαλείο με ανοιχτό κώδικα ικανό να παραμετροποιηθεί.

**Πίνακας Περιεχομένων**

<b>Ευχαριστίες</b>	2
<b>Abstract</b>	3
<b>Σύνοψη</b>	4
<b>Πίνακας Περιεχομένων</b>	5
<b>Πίνακας Σχημάτων</b>	7
<b>Πίνακας Εικόνων</b>	7
<b>1 Εισαγωγή</b>	9
1.1 Ορισμός	9
1.2 Σημασία των πολυμέσων	11
1.3 Σημαντικά χαρακτηριστικά των πολυμέσων	12
1.4 Αυτόνομα και Δικτυωμένα Πολυμέσα	13
1.5 Ταξινόμηση των εφαρμογών πολυμέσων	14
<b>2 Εφαρμογή Εργαλείων Σχολιασμού Ανάγκη για Οργάνωση</b>	15
2.1 Λόγος χρήση και το πλαίσιο	16
2.2 Το πρόβλημα της διαχείρισης των εικόνων	16
2.3 Αυτόνομα- Μη Αυτόνομα Σχόλια Εικόνας	18
<b>3 Παραδείγματα Εργαλείων Σχολιασμού</b>	20
3.1 Εισαγωγή	20
3.2 Image Annotation Tool	21
3.3 Anvil	22
3.4 Annotate Expert	24
3.5 Foto Tagger	26
3.6 M-Onomat-Annotizer	28
3.7 Annotea	31
3.8 Video Image Annotation Tool	32
3.9 LabelMe	34
3.10 JFotoNotes	36
3.11 IAT	38
3.12 Συνοπτικός πίνακας	39
<b>4 Υλοποίηση στην MATLAB</b>	41
4.1 Εισαγωγή	41
4.2 Εγκατάσταση Εργαλειοθήκης	42
4.3 Περιγραφή Εργαλειοθήκης	43
4.4 Διαβάζοντας το ευρετήριο	46
4.5 Οπτικοποίηση	47
4.6 Ερωτήματα	49
4.7 Εξαγωγή πολυγώνων και τμημάτων	51
4.8 Συλλογή στατιστικών σχολιασμών	52
4.9 Αντιμετώπιση συνωνύμων και επισήμανση θορύβου	53
4.10 Συνάρτηση <code>Lmreplaceobjectname</code>	53
4.11 Σχολιάστε δικές σας εικόνες	54
4.12 Άλλες χρήσιμες εντολές	55
4.13 Κάνοντας τις εικόνες μας Label Me	56
<b>Παράρτημα</b>	62

1.	Χειρισμός Εικόνων	62
2.	Αναγνώριση Εικόνων	62
<b>5</b>	<b>Συμπεράσματα- Εφαρμογές</b>	<b>63</b>
5.1	Η εικόνα ως πολυμεσικό αντικείμενο	63
5.2	Ηλεκτρονικές βιβλιοθήκες	63
5.2.1	Υπηρεσίες Συνεργατικής ανάκτηση πληροφορίας	64
	<b>ΒΙΒΛΙΟΓΡΑΦΙΑ-ΠΗΓΕΣ</b>	<b>66</b>

**Πίνακας Σχημάτων**

Σχήμα 1.1 Ταξινόμηση ειδών πληροφορίας	11
Σχήμα 1.2 Αυτόνομα Δικτυωμένα Πολυμέσα	13

**Πίνακας Εικόνων**

Εικόνα 1 Στιγμιότυπο Image Annotation Tool	22
Εικόνα 2 Στιγμιότυπο ANVIL	23
Εικόνα 3 Στιγμιότυπο Annotate Expert	25
Εικόνα 4 Στιγμιότυπο Annotate Expert	26
Εικόνα 5 Στιγμιότυπο Foto Tagger	27
Εικόνα 6 Στιγμιότυπο Foto Tagger	28
Εικόνα 7 Στιγμιότυπο M-OntoMat-Annotizer	29
Εικόνα 8 Αρχιτεκτονική M-OntoMat-Annotizer	30
Εικόνα 9 Στιγμιότυπο M-OntoMat-Annotizer	31
Εικόνα 10 Στιγμιότυπο Annotea	32
Εικόνα 11 Στιγμιότυπο VIA	33
Εικόνα 12 Στιγμιότυπο VIA	34
Εικόνα 13 Στιγμιότυπο LabelMe	35
Εικόνα 14 Στιγμιότυπο LabelMe	36
Εικόνα 15 Στιγμιότυπο jFotoNotes	37
Εικόνα 16 Ιστοσελίδα Lable me	40
Εικόνα 17 Ιστοσελίδα Lable me	40
Εικόνα 18 Διαβάζει τις εικόνες διαδικτυακά	41
Εικόνα 19 Διαβάζει τις εικόνες διαδικτυακά μετά από χρόνο	42
Εικόνα 20 Μεταμορφώνοντας τη βάση δεδομένων μετά από χρόνο	43
Εικόνα 21 Μεταμορφώνοντας τη βάση δεδομένων με Lmistall	43
Εικόνα 22 Διαβάζοντας από το ευρετήριο	44
Εικόνα 23 Οπτικοποίηση πρώτης εικόνας	45
Εικόνα 24 D(1:6)	46
Εικόνα 25 Όλα τα αντικείμενα της πρώτης εικόνας	46
Εικόνα 26 Από τον D αναζητά όλα τα αντικείμενα με το car	47
Εικόνα 27 Αναζητώντας mouse+pad	48
Εικόνα 28 Αναζητώντας mouse- pad	48
Εικόνα 29 Αναζητώντας building, car+side ,tree	49
Εικόνα 30 Πολύγωνο	50
Εικόνα 31 Συλλογή στατιστικών σχολιασμών	51
Εικόνα 32 Εξαγωγή πολυγώνων και τμημάτων	52
Εικόνα 33 Συνάρτηση Lmobjectname	53

<b>Εικόνα34</b> Συνάρτηση Lmreplaceobjectname	54
<b>Εικόνα 35</b> Η αρχική σελίδα του Label me	56
<b>Εικόνα36</b> Βήματα πώς να σχολιάσουμε τις εικόνες	57
<b>Εικόνα 37</b> Βήματα πώς να σχολιάσουμε τις εικόνες	57
<b>Εικόνα38</b> Δικές μου σχολιασμένες εικόνες – τοπία	58
<b>Εικόνα 39</b> Δικές μου σχολιασμένες εικόνες - τοπία	58
<b>Εικόνα40</b> Δικές μου σχολιασμένες εικόνες – τοπία	59
<b>Εικόνα 41</b> Δικές μου σχολιασμένες εικόνες – τοπία	59
<b>Εικόνα 42</b> Δικές μου σχολιασμένες εικόνες – προσωπικές	60
<b>Εικόνα 43</b> Δικές μου σχολιασμένες εικόνες– προσωπικές	60
<b>Εικόνα44</b> Δικές μου σχολιασμένες εικόνες– προσωπικές	62
<b>Εικόνα 45</b> Δικές μου σχολιασμένες εικόνες– προσωπικές	62



## ΚΕΦΑΛΑΙΟ 1

### Εισαγωγή

Τα πολυμέσα είναι μία από τις πιο πολυσυζητημένες τεχνολογίες των αρχών της δεκαετίας του 90. Το ενδιαφέρον αυτό είναι απόλυτα δικαιολογημένο, αφού τα πολυμέσα αποτελούν το σημείο συνάντησης πέντε μεγάλων βιομηχανιών: της πληροφορικής, των τηλεπικοινωνιών, ηλεκτρονικών εκδόσεων, της βιομηχανίας audio και video καθώς και της βιομηχανίας της τηλεόρασης και του κινηματογράφου. Μια ανάλογη αναστάτωση επέφερε και η εμφάνιση της επιστήμης των δικτύων υπολογιστών στη δεκαετία του 70, φέρνοντας πιο κοντά την πληροφορική με τις τηλεπικοινωνίες. Αυτή η προσέγγιση οδήγησε σε προϊόντα που στόχευαν κυρίως στην αγορά των επιχειρήσεων. Τα πολυμέσα έκαναν κάτι περισσότερο, διεύρυναν την αγορά των προϊόντων των παραπάνω βιομηχανιών που πλέον στοχεύουν και στους καταναλωτές.

Η πληθώρα και οι ποικιλία των νέων προϊόντων καθώς και η προσπάθεια εκμετάλλευσης του ενδιαφέροντος που επέδειξε το αγοραστικό κοινό για την τεχνολογία των πολυμέσων συνετέλεσαν στην σύγχυση που υπάρχει ακόμα και σήμερα όσον αφορά στο τι είναι και τι δεν είναι ένα σύστημα πολυμέσων.

#### 1.1 Ορισμός

Ο αγγλικός όρος, που εδώ έχει αποδοθεί ως πολυμέσα, είναι multimedia. Ο όρος αυτός αποτελείται από δύο μέρη: το πρόθεμα multi και τη ρίζα media. Κατά συνέπεια ο ορισμός που προκύπτει είναι: : "πολλαπλοί μεσολαβητές" ή "πολλαπλάμέσα" και χρησιμοποιείται είτε ως ουσιαστικό είτε ως επίθετο.

Η πρώτη προσέγγιση του ορισμού δεν μας λέει και πολλά πράγματα. Μπορούμε όμως να τον βελτιώσουμε αναλογιζόμενοι τον τρόπο χρήσης των όρων multi, και media. Ο αγγλικός όρος media χρησιμοποιείται σε πολλούς οικονομικούς, τεχνικούς και επιστημονικούς τομείς με διαφορετικές σημασίες. Το κοινό σημείο αυτών των χρήσεων είναι ότι σχετίζονται πάντοτε με κάποιο είδος χειρισμού πληροφορίας:

1. Αποθήκευση και επεξεργασία στην πληροφορική
2. Παραγωγή στον χώρο των εκδόσεων
3. Διανομή στον χώρο των μαζικών μέσων επικοινωνίας
4. Μετάδοση στις τηλεπικοινωνίες
5. Αντίληψη κατά την αλληλεπίδραση του ανθρώπου με το περιβάλλον του.

Κατά συνέπεια μπορούμε να βελτιώσουμε τον ορισμό ως εξής: Πολυμέσα στον χώρο της τεχνολογίας πληροφορίας (information field) σημαίνει "πολλαπλοί μεσολαβητές" μεταξύ της πηγής και του

παραλήπτη της πληροφορίας ή “πολλαπλά μέσα” μέσω των οποίων η πληροφορία αποθηκεύεται, μεταδίδεται, παρουσιάζεται ή γίνεται αντιληπτή..

Σύμφωνα με αυτόν τον ορισμό, ένα σύστημα που συνδυάζει, για παράδειγμα, τον έλεγχο βιντεοκασέτας και οπτικών μέσων αποθήκευσης μπορεί να χαρακτηριστεί ως σύστημα πολυμέσων. Επίσης συστήματα πολυμέσων θα είναι η εφημερίδα, που συνδυάζει κείμενο και εικόνα, και η τηλεόραση, που συνδυάζει ήχο και κινούμενη εικόνα. Εδώ δεν αναφερόμαστε σε τόσο ευρύ φάσμα συστημάτων. Περιοριζόμαστε σε αυτά στα οποία η πληροφορία είναι ψηφιακή (ή ψηφιοποιημένα - digitized) και ελέγχεται από υπολογιστή. Ενδιαφερόμαστε δηλαδή για ψηφιακά πολυμέσα τα οποία και ορίζουμε ως εξής:

*Ψηφιακά πολυμέσα είναι ο τομέας που ασχολείται με την ελεγχόμενη από υπολογιστή ολοκλήρωση κειμένου, γραφικών, ακίνητης και κινούμενης εικόνας, animation, ήχου, και οποιουδήποτε άλλου μέσου ψηφιακής αναπαράστασης, αποθήκευσης, μετάδοσης και επεξεργασίας της πληροφορίας*

Διαβάζοντας αυτόν τον ορισμό, δημιουργείται το ερώτημα: ποιους και πόσους τύπους πληροφορίας πρέπει να συνδυάζει ένα σύστημα, για να μπορεί δίκαια να χαρακτηρίζεται ως σύστημα πολυμέσων; Όπως είναι φανερό, η απάντηση σε αυτό το ερώτημα δεν μπορεί να είναι αυστηρή, γιατί δεν έχει γίνει κάποια συμφωνία πάνω στον ορισμό των πολυμέσων. Όμως στην πράξη έχουν δημιουργηθεί κάποιοι de facto κανόνες που καθορίζουν τι πρέπει να περιλαμβάνει ένα σύστημα πολυμέσων ανάλογα με το είδος της εφαρμογής. Σαν κατευθυντήρια γραμμή μπορούμε να δώσουμε τον παρακάτω ορισμό: Στην πράξη, ο όρος πολυμέσα υπονοεί την ολοκλήρωση ενός τουλάχιστον “διακριτού” τύπου πληροφορίας και ενός “συνεχούς”.

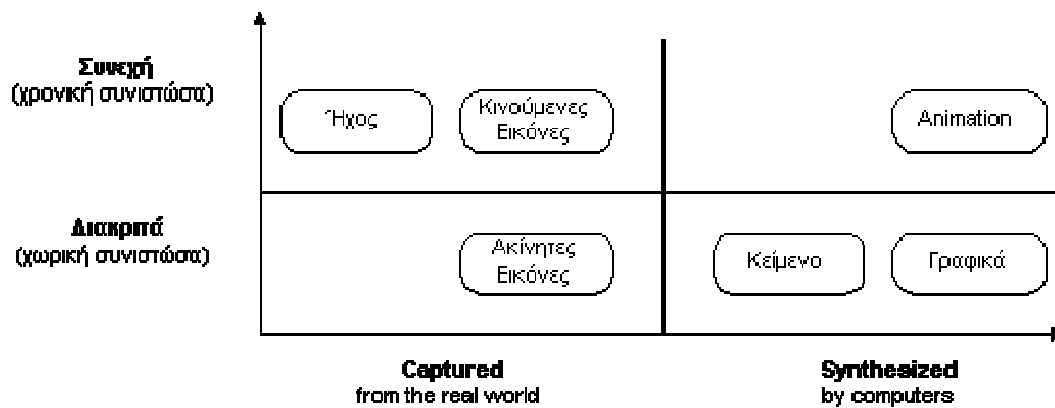
Στον παραπάνω ορισμό έχει γίνει διαχωρισμός των τύπων πληροφορίας σε διακριτούς και συνεχείς. Ένας άλλος διαχωρισμός είναι σε captured και synthesized μέσα. Ας δούμε τι σημαίνουν αυτοί οι διαχωρισμοί (Σχήμα 1.1):

## **1. Captured versus synthesized media**

Αυτός ο διαχωρισμός αναφέρεται στον τρόπο μεταφοράς της πληροφορίας στη μορφή που υπαγορεύει ο κάθε τύπος. Αν η πληροφορία συλλαμβάνεται απευθείας από τον πραγματικό κόσμο μιλάμε για captured media ενώ αν δημιουργείται από τον άνθρωπο μέσω κάποιων εργαλείων έχουμε τα συνθετικά μέσα. Για παράδειγμα, μια ψηφιακή φωτογραφική μηχανή ή ένας scanner μεταφέρει αυτόματα την εικόνα ενός αντικειμένου σε ψηφιακή μορφή κατάλληλη για χρήση στον υπολογιστή. Δηλαδή οι εικόνες είναι captured media. Το κείμενο, όταν αυτό πληκτρολογείται στον υπολογιστή είναι προφανώς συνθετικό μέσο. Αν όμως λαμβάνεται μέσω scanner και προγράμματος OCR πρέπει να θεωρηθεί ως captured.

## **2. Discrete versus continuous media**

Όταν ένας τύπος πληροφορίας έχει μόνο χωρική διάσταση ονομάζεται διακριτός. Αν υπάρχει και η συνιστώσα του χρόνου ονομάζεται συνεχές. Για παράδειγμα, οι εικόνες, το κείμενο και τα γραφικά είναι διακριτά, ενώ το βίντεο, ο ήχος και το animation είναι συνεχή.



Σχήμα 1.1. Ταξινόμηση ειδών πληροφορίας

Όλα αυτά τα μέσα που έχουν αναφερθεί ως τώρα απευθύνονται σχεδόν αποκλειστικά στην όραση και στην ακοή του ανθρώπου. Ένα σύστημα πολυμέσων δεν περιέχει απαραίτητα πληροφορίες για παραπάνω από μια αισθήσεις, παρόλο που κάτι τέτοιο είναι γενικά επιθυμητό.

## 1.2 Σημασία των πολυμέσων

Από την εποχή της εφεύρεσης του πρώτου μέσου μαζικής επικοινωνίας, του τυπογράφου, επικρατούσε η αντίληψη ότι η πληροφορία εκφράζεται και μεταδίδεται πιο εύκολα με τον γραπτό λόγο. Με την ανακάλυψη όμως του φωτογραφικού φιλμ και της τηλεόρασης στις αρχές του εικοστού αιώνα και πολύ περισσότερο, αργότερα των ηλεκτρονικών υπολογιστών και του διαδικτύου, ο ρόλος της εικόνας στην διαδικασία της επικοινωνίας απέκτησε καινούρια διάσταση. Ως αποτέλεσμα αυτών των τεχνολογικών επιτευγμάτων, κατέστη δυνατή η μαζική αναπαραγωγή και μεταφορά εικόνας και η ολοένα και μεγαλύτερη χρησιμοποίησή της στην μετάδοση πληροφορίας. Η στροφή που παρατηρήθηκε προς την επικοινωνία με χρήση οπτικού υλικού, οφείλεται κατά κύριο λόγο σε κάποια πλεονεκτήματα που κατέχει η εικόνα, έναντι των άλλων μέσων.

Η εικόνα, σε αντίθεση με τον γραπτό λόγο, δεν βασίζεται σε κάποιο κώδικα επικοινωνίας ή λεξιλόγιο αλλά η άντληση της πληροφορίας που περιέχεται σ' αυτήν μπορεί να γίνει αυτούσια, χωρίς περαιτέρω γνώση. Ένα κείμενο που περιγράφει στα γερμανικά τον τρόπο χειρισμού μιας συσκευής, ίσως να μην είναι κατανοητό από όλους αλλά μια σειρά από εικόνες σίγουρα είναι. Αποτελεί δηλαδή, ένα μέσο με περισσότερους πιθανούς αποδέκτες επιτυγχάνοντας με αυτόν τον τρόπο αποδοτικότερη μετάδοση της πληροφορίας. Πέρα από αυτό, η εικόνα ως μέσο επικοινωνίας μπορεί να εκφράσει και να μεταφέρει πληροφορία η οποία τις περισσότερες φορές δεν θα ήταν δυνατό να εκφραστεί διαφορετικά. Υπερτερεί του γραπτού λόγου, υπό την έννοια ότι διευρύνεται η πληροφορία που μπορεί να εξάγει ο αποδέκτης αυτής, κάνοντας την επικοινωνία αποτελεσματικότερη. Τα δύο αυτά χαρακτηριστικά της εικόνας, σε συνδυασμό με την ευκολία με την οποία μπορεί να τις επεξεργάζεται ο άνθρωπος, καθιέρωσαν στις μέρες μας την οπτική πληροφορία ως το κυρίαρχο μέσο επικοινωνίας.

Η καλπάζουσα ανάπτυξη που παρατηρείται σήμερα στην τεχνολογία των ηλεκτρονικών υπολογιστών, των δικτύων επικοινωνίας καθώς και των μέσων αποθήκευσης, δίνουν την δυνατότητα ολοένα ευκολότερης και μαζικότερης μεταφοράς εικόνας σε παγκόσμια κλίμακα. Η συνεχής αύξηση της υπολογιστικής δύναμης των επεξεργαστών και της ταχύτητας των δικτύων, σε συνδυασμό με την ανάπτυξη όλο και πιο αποτελεσματικών αλγόριθμων συμπίεσης, δίνουν πλέον την ευχέρεια στον χρήστη να αναπαράγει να επεξεργαστεί και να μεταδώσει οπτικό υλικό χωρίς περιορισμούς. Επιπλέον η εμφάνιση καινούριων εφαρμογών που κάνουν χρήση πολυμεσικού περιεχομένου (web-conferencing, peer-to-peer networks, ηλεκτρονικά παιχνίδια κ.α.) προσελκύουν όλο και περισσότερους χρήστες και αυξάνουν περαιτέρω τα πολυμεσικά δεδομένα που είναι διαθέσιμα.

Γίνεται επομένως σαφές, ότι ο όγκος της πληροφορίας υπό μορφή εικόνας που αναπαράγεται, μεταδίδεται και είναι αποθηκευμένος ανά πάσα στιγμή είναι τόσο μεγάλος, που η πρόσβαση σ' αυτό που μας ενδιαφέρει γίνεται δύσκολη υπόθεση. Το γεγονός αυτό μπορεί να γίνει καλύτερα αντιληπτό, αν λάβει κανείς υπόψη ότι το εβδομήντα τοις εκατό της πληροφορίας που υπάρχει στον παγκόσμιο ιστό είναι σε μορφή εικόνων. Η πληροφορία που αναζητούμε είναι διαθέσιμη αλλά δεν υπάρχει τρόπος να ανακτηθεί αποτελεσματικά, τουλάχιστον σε εύλογο χρονικό διάστημα. Είναι λοιπόν

αναγκαία η ανάπτυξη εκείνων των τεχνικών, που θα επιτρέψουν την εξαγωγή περιγραφών της πληροφορίας που περιέχεται σε μια εικόνα και στην συνέχεια, την αναζήτηση και ανάκτηση τους από μια βάση δεδομένων χωρίς την παρέμβαση του ανθρώπινου παράγοντα. Στα πλαίσια της επίτευξης αυτού του στόχου, όπως θα φανεί στην συνέχεια, έγιναν αρκετές προσπάθειες και αναπτύχθηκαν συστήματα που συντέειναν στην αποτελεσματικότερη ανάλυση και εν συνεχεία αναζήτηση των εικόνων.

### 1.3 Σημαντικά χαρακτηριστικά των πολυμέσων

Multimedia (πολυμεσικές) παρουσιάσεις μπορούν να προβληθούν μεμονωμένα ανά χρήστη, να τεθούν σε αναμονή, να μεταδοθούν ή να παιχτούν σε τοπικό επίπεδο με ένα media player. Μια τηλεοπτική παρουσίαση μπορεί να είναι ζωντανή είτε μαγνητοσκοπημένη παρουσίαση πολυμέσων. Εκπομπές και εγγραφές μπορούν να γίνουν με αναλογική ή ψηφιακή τεχνολογία ηλεκτρονικών μέσων. Ψηφιακά πολυμέσα μπορούν να ληφθούν ή να μεταδοθούν σε απευθείας σύνδεση. Streaming (συνεχής ροή) πολυμέσων μπορεί να είναι ζωντανά ή κατά παραγγελία.

Πολυμεσικά παιχνίδια και προσομοιώσεις μπορούν να χρησιμοποιηθούν σε ένα φυσικό περιβάλλον με ειδικά εφέ, με πολλούς χρήστες σε ένα online δίκτυο, ή σε τοπικό επίπεδο με μια σύνδεση του υπολογιστή και του συστήματος του παιχνιδιού, ή του προσομοιωτή.

Οι διάφορες μορφές τεχνολογικών ή ψηφιακών πολυμέσων μπορούν να αποσκοπούν στη βελτίωση της εμπειρίας του χρήστη, για παράδειγμα, για να καταστεί ευκολότερη και ταχύτερη η μετάδοση πληροφοριών. Ή στην ψυχαγωγία ή την τέχνη, να ξεπεράσει την καθημερινή εμπειρία. Ένα lasershow είναι μια ζωντανή παράσταση πολυμέσων.

Τα ενισχυμένα επίπεδα διαδραστικότητας κατέστησαν δυνατά με το συνδυασμό περιεχομένου πολλαπλών μορφών πολυμέσων. Τα Online multimedia (πολυμέσα σε απευθείας μετάδοση) οδηγούνται όλο και περισσότερο σε δεδομένα βασισμένα στα αντικείμενα, διευκολύνοντας τις εφαρμογές σε μία καινοτόμα συνεργασία με τον τελικό χρήστη εξατομικεύοντας το περιεχόμενό τους. Παραδείγματα αυτών ποικίλλουν από πολλαπλές μορφές περιεχομένου σε Ιστοσελίδες, όπως το φωτογραφικό υλικό και με τις δύο εικόνες (φωτογραφίες) και τον τίτλο (κείμενο) για την ενημέρωση του χρήστη, με τις προσομοιώσεις του οποίου συντελούνται, εκδηλώσεις, εικονογραφήσεις, κινούμενα σχέδια ή βίντεο τροποποιήσιμα, επιτρέποντας στα πολυμέσα να αποκτήσουν εμπειρία που πρέπει να τροποποιηθεί χωρίς επαναπρογραμματισμό. Εκτός από να δει και να ακούσει, η Haptic τεχνολογία επιτρέπει εικονικά αντικείμενα να αισθανθούν. Οι αναδυόμενες τεχνολογίες περιλαμβάνουν ψευδαισθήσεις της γεύσης και της όσφρησης οι οποίες μπορούν επίσης να ενισχύσουν την εμπειρία των πολυμέσων.

Με βάση τον παραπάνω ορισμό, προκύπτουν τέσσερα χαρακτηριστικά για τα συστήματα πολυμέσων που μας ενδιαφέρουν:

1. Πρέπει να ελέγχονται από υπολογιστή.

Δηλαδή η παρουσίαση της πληροφορίας γίνεται μέσω του υπολογιστή και ελέγχεται από αυτόν.

2. Είναι ολοκληρωμένα (integrated).

Η ολοκλήρωση υπονοεί ότι ο αριθμός των υποσυστημάτων είναι κατά το δυνατόν ελάχιστος και ενσωματωμένος στον υπολογιστή. Παράδειγμα ολοκλήρωσης αποτελεί ή οθόνη του υπολογιστή που χρησιμοποιείται για την απεικόνιση κειμένου, εικόνας και βίντεο.

3. Η πληροφορία πρέπει να είναι σε ψηφιακή μορφή.

Το χαρακτηριστικό αυτό είναι απόρροια της απαίτησης για έλεγχο και παρουσίαση μέσω υπολογιστή.

4. Το interface με το χρήστη πρέπει να επιτρέπει αλληλεπίδραση (interaction).

Αν και δεν περιλαμβάνεται ευθέως στον ορισμό, η δυνατότητα αυτή επιτρέπει την δημιουργία εφαρμογών με περισσότερες δυνατότητες από την απλή παρουσίαση της πληροφορίας (όπως γίνεται για παράδειγμα μέσω ενός video-player ή ενός CD-player) και είναι ιδιαίτερο χαρακτηριστικό των ελεγχόμενων μέσω υπολογιστή πολυμέσων.

## 1.4 Αυτόνομα και Δικτυωμένα Πολυμέσα

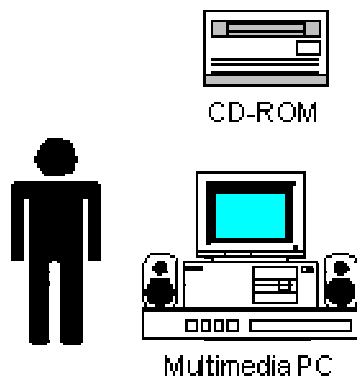
Ο όρος αυτόνομα ή τοπικά πολυμέσα αναφέρεται σε εφαρμογές που χρησιμοποιούν μόνο τον υπολογιστή στον οποίο τρέχουν. Κατά συνέπεια, ο υπολογιστής αυτός πρέπει να έχει όλες τις απαραίτητες υπό μονάδες όπως:

1. επεξεργαστή (όχι τερματικό δηλαδή)
2. ικανό υποσύστημα γραφικών και ήχου
3. ηχεία, μικρόφωνο
4. αρκετά αποθηκευτικά μέσα
5. κάποιες μορφές οπτικό δίσκο συνήθως CD-ROM

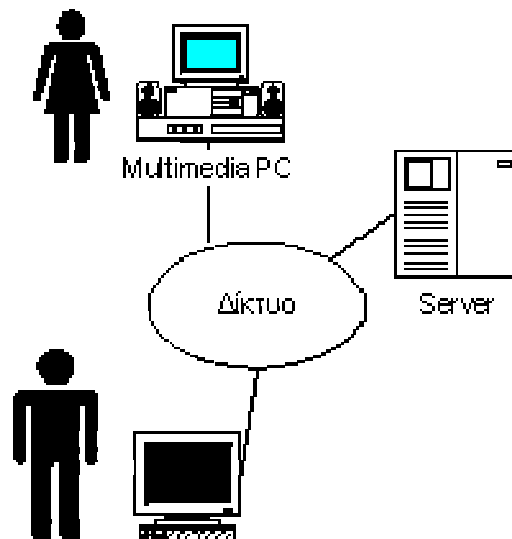
Πολλές όμως φορές είναι επιθυμητό οι εφαρμογές πολυμέσων να επικοινωνούν μέσω δικτύου με άλλους υπολογιστές για δύο λόγους:

- Την υποστήριξη εφαρμογών οι οποίες είναι εγγενώς δικτυακές. Παραδείγματα τέτοιων εφαρμογών είναι το ηλεκτρονικό ταχυδρομείο πολυμέσων και η τηλεδιάσκεψη.
- Την υλοποίηση του μοντέλου πελάτη-εξυπηρετητή(client-server). Πολλές φορές αν και μια εφαρμογή πολυμέσων μπορεί κάλλιστα να υλοποιηθεί σε έναν υπολογιστή μόνο, για λόγους οικονομίας του υλικού, είναι επιθυμητό να μπορεί να αξιοποιεί και υποσυστήματα που ανήκουν σε άλλους υπολογιστές. Χαρακτηριστική περίπτωση είναι η ύπαρξη ενός υπολογιστή με μεγάλα αποθηκευτικά μέσα (εξυπηρετητής) προσπελάσιμα μέσω δικτύου και από άλλους υπολογιστές με περιορισμένες δυνατότητες αποθήκευσης (πελάτες).

### Αυτόνομα Πολυμέσα



### Δικτυωμένα Πολυμέσα



Σχήμα 1.2. Αυτόνομα και Δικτυωμένα Πολυμέσα

## 1.5 Ταξινόμηση Των Εφαρμογών Πολυμέσων

### Ένας Χρήστης

Μέχρι πριν από μερικά χρόνια, σχεδόν όλες οι εφαρμογές των πολυμέσων αφορούσαν έναν μόνο χρήστη (αυτόνομα πολυμέσα). Οι πρώτες από αυτές ήταν εξειδικευμένα συστήματα εκπαίδευσης, παροχής πληροφοριών ή πωλήσεων (CBT-Computer Based Training, POI-Points Of Information, POS-Points Of Sales). Στη συνέχεια, με την αύξηση των δυνατοτήτων των προσωπικών υπολογιστών και την καθιέρωση του CD-ROM ως φθηνή και πρακτική πλατφόρμα διανομής, η χρήση αυτών των εφαρμογών γενικεύτηκε. Οι περισσότερες από τις εφαρμογές των αυτόνομων πολυμέσων ουσιαστικά δεν είναι καινούργιες. Από παλιά υπήρχαν παρόμοια συστήματα τα οποία όμως δεν ήταν ούτε αρκετά ελκυστικά ούτε και αποτελεσματικά ώστε να συγκεντρώσουν το ενδιαφέρον. Με τον εμπλουτισμό τους με τις δυνατότητες των πολυμέσων, τους δόθηκε μια άλλη διάσταση και η χρήση τους γενικεύτηκε σε πολλά πεδία, αλλά στην ουσία δεν γεννήθηκαν νέα είδη εφαρμογών. Η προσπάθεια κατηγοριοποίησης αυτών, καταλήγει σε διαχωρισμό τους με βάση το πεδίο εφαρμογής τους, αφού ουσιαστικά όλες αποτελούν interactive εφαρμογές που δίνουν την δυνατότητα στο χρήστη να χειριστεί αποτελεσματικά μεγάλους όγκους και πολλά είδη πληροφορίας.

### Πολλοί Χρήστες

Ο χώρος των δικτυωμένων πολυμέσων αποτελεί σήμερα τον κλάδο της επιστήμης των υπολογιστών που έχει συγκεντρώσει το μεγαλύτερο ενδιαφέρον. Το πιο σημαντικό στοιχείο είναι ότι το ενδιαφέρον προέρχεται τόσο από την ερευνητική κοινότητα όσο και από τις επιχειρήσεις και το αγοραστικό κοινό. Υπάρχει βέβαια ακόμα μια ρευστότητα στον χώρο, που οφείλεται στους τεχνολογικούς φραγμούς της υπάρχουσας δικτυακής υποδομής, και στην έλλειψη προτύπων, χωρίς όμως αυτό να έχει εμποδίσει την εμφάνιση των πρώτων δειγμάτων μιας πληθώρας νέων εφαρμογών. Αυτές οι εφαρμογές χωρίζονται σε δύο βασικές κατηγορίες

#### 1. People-to-People Multimedia Applications

Οι εφαρμογές αυτές έχουν ως στόχο τη διευκόλυνση της επικοινωνίας μεταξύ δύο οι περισσότερων προσώπων. Η φύση αυτής της επικοινωνίας μπορεί να έχει διάφορες μορφές: από προσωπική συνομιλία κοινωνικού χαρακτήρα μέχρι πολύπλοκη επικοινωνία ομάδων εργασίας. Οι εφαρμογές αυτής της κατηγορίας μπορούν επιπλέον να διαιρεθούν στις εξής υποκατηγορίες:

- ιδιωτικές έναντι επαγγελματικών εφαρμογών
- εφαρμογές επικοινωνίας μεταξύ δύο ατόμων (interpersonal applications) έναντι εφαρμογών επικοινωνίας μεταξύ ομάδων ατόμων (group-oriented applications)
- εφαρμογές πραγματικού χρόνου (real-time applications) έναντι ασύγχρονων εφαρμογών (asynchronous applications)

#### 2. People-to-Systems Multimedia Applications

Οι εφαρμογές αυτές δίνουν τη δυνατότητα σε μεμονωμένους χρήστες να προσπελάσουν πληροφορίες που βρίσκονται σε κάποιο απομακρυσμένο σύστημα. Συνήθως, το απομακρυσμένο σύστημα είναι κάποιος εξυπηρετητής πολυμέσων (multimedia information server ).

### *Σύγχρονες Εφαρμογές People-to-People*

- Εφαρμογές στις οποίες μόνο δύο άτομα παίρνουν μέρος (interpersonal applications)
- Εφαρμογές διανομής πληροφορίας (person-to-group applications), στις οποίες η πληροφορία πηγάζει από ένα μόνο σημείο ενώ οι παραλήπτες είναι πολλοί. Επιπλέον οι εφαρμογές αυτές επιτρέπουν μόνο μια κατεύθυνση επικοινωνίας. Χαρακτηριστικό παράδειγμα είναι τα σεμινάρια: η εικόνα του ομιλητή μεταδίδεται σε πολλούς ανθρώπους, οι οποίοι βρίσκονται μακριά από τη αίθουσα.
- Τηλεδιάσκεψη ομάδων (group teleconferencing). Το μοντέλο αυτό επιτρέπει επικοινωνία δύο κατευθύνσεων (bi-directional communication) μεταξύ δύο ή περισσότερων ομάδων ατόμων. Συνήθως, η επικοινωνία γίνεται μέσω ομιλίας (audio teleconferencing) ή ομιλίας και εικόνας (audio-video teleconferencing). Μια άλλη μορφή επικοινωνίας είναι η εργασία σε ένα μοιραζόμενο χώρο εργασίας υπολογιστή (shared computer workspace). Όταν συνδυάζονται διάφοροι τρόποι επικοινωνίας, τότε μιλάμε για multimedia teleconference. Περισσότερες λεπτομέρειες για αυτές τις εφαρμογές θα δώσουμε παρακάτω.

### ***Ασύγχρονες Εφαρμογές People-to-People***

- Ηλεκτρονικό ταχυδρομείο πολυμέσων (Multimedia electronic mail). Η λειτουργία του είναι παρόμοια με το συμβατικό ηλεκτρονικό ταχυδρομείο, μόνο που τα έγγραφα που ανταλλάσσονται μπορούν να περιέχουν και μορφοποιημένο κείμενο, εικόνες, ήχο, video καθώς και συνδέσμους προς εξωτερικά αρχεία.
- Ασύγχρονη διάσκεψη μέσω υπολογιστών με πολυμέσα (multimedia asynchronous computer conferencing). Ο όρος διάσκεψη μέσω υπολογιστών αναφέρεται σε εφαρμογές, που εμφανίστηκαν τη δεκαετία του 80, που επιτρέπουν την ασύγχρονη συνδιάλεξη ατόμων μέσω πινάκων ανακοινώσεων (bulletin boards). Η προσθήκη των πολυμέσων επιτρέπει την αποστολή ανακοινώσεων που περιέχουν και άλλα είδη πληροφορίας εκτός από απλό κείμενο.

### ***Εφαρμογές People-to-System***

- Interactive εφαρμογές. Το όνομα αυτό έχει στόχο να τονίσει το γεγονός ότι ο χρήστης έχει το πρώτο λόγο στη επικοινωνία και όχι ο εξυπηρετητής. Κατά συνέπεια, η επικοινωνία γίνεται κατόπιν αιτήσεως του χρήστη την στιγμή που αυτός επιθυμεί. Συνήθως, οι εφαρμογές αυτές έχουν ως στόχο την ανεύρεση κάποιας πληροφορίας (information retrieval applications). Σε άλλες περιπτώσεις, η αλληλεπίδραση του χρήστη με το σύστημα έχει ως στόχο κάποιου είδους δόσοληψία (transaction oriented applications).
- Εφαρμογές διανομής (distribution applications). Την κατηγορία αυτή την συναντήσαμε και παραπάνω. Η διαφορά εδώ είναι ότι η πληροφορία που διανέμεται δεν έχει ως στόχο την επικοινωνία μεταξύ ατόμων. Η διανομή μπορεί να περιορίζεται σε ομάδες ατόμων που έχουν λάβει κάποιας μορφής εξουσιοδότηση ή σε όλους όσους διαθέτουν τον κατάλληλο εξοπλισμό.

## **ΚΕΦΑΛΑΙΟ 2**

### **Εφαρμογή Εργαλείων Σχολιασμού**

#### **Ανάγκη για Οργάνωση**

##### **2.1 Λόγος χρήση και το πλαίσιο**

Από τη στιγμή που τα μέσα είναι ο πληθυντικός του μέσου, ο όρος "multimedia"(πολυμέσα) είναι ένας πλεονασμός αν το "multi" χρησιμοποιείται για να περιγράψει πολλές εμφανίσεις μόνο μιας μορφής πολυμέσων, όπως μια συλλογή από CD ήχου. Για το λόγο αυτό είναι σημαντικό ότι η λέξη "πολυμέσα", χρησιμοποιείται αποκλειστικά για να περιγράψει πολλαπλές μορφές των μέσων ενημέρωσης και περιεχομένου.

Ο όρος "πολυμέσα" είναι επίσης ασαφής. Στατικά περιεχόμενα (όπως ένα βιβλίο από χαρτί) μπορούν να θεωρούνται πολυμέσα εάν περιέχουν εικόνες και κείμενα ή μπορούν να θεωρηθούν διαδραστικά εάν ο χρήστης αλληλεπιδρά με την αλλαγή σελίδας κατά βούληση. Τα βιβλία μπορούν επίσης να θεωρηθούν μη γραμμικά αν οι σελίδες είναι προσπελάσιμες μη διαδοχικά. Ο όρος «βίντεο», εάν δεν χρησιμοποιείται αποκλειστικά για να περιγράψει κινούμενη φωτογραφία, είναι ασαφής στην ορολογία των πολυμέσων. Το βίντεο χρησιμοποιείται συχνά για να περιγράψει τη μορφή του αρχείου, τη μορφή παράδοσης, ή τη μορφή παρουσίασης αντί για "πλάνο" που χρησιμοποιείται για τη διάκριση κινούμενης φωτογραφίας από το "animation" των παρεχόμενων εικόνων κίνησης. Πολλαπλές μορφές περιεχομένου πληροφοριών δεν θεωρούνται συχνά σύγχρονες μορφές παρουσίασης όπως ήχου ή βίντεο. Ομοίως, ενιαίες μορφές περιεχομένου πληροφοριών με ενιαίες μεθόδους επεξεργασίας πληροφοριών (π.χ. μη διαδραστικός ήχος) ονομάζονται συχνά πολυμέσα, ίσως για να διακρίνουν τα στατικά μέσα από τα ενεργά μέσα.

##### **2.2 Το πρόβλημα της διαχείρισης των εικόνων**

Ήδη από τα τέλη τις δεκαετίας του εβδομήντα, έγινε ορατή η ανάγκη για την ανάπτυξη συστημάτων και μεθοδολογιών, μέσω της εξαγωγής σημασιολογικής περιγραφής για το περιεχόμενο των εικόνων, θα επιτρέπουν την δεικτοδότηση τους με τέτοιο τρόπο ώστε να μπορούν στην συνέχεια να ανακτηθούν αποτελεσματικά. Ένα σχολιασμός είναι δυνατόν να αποτελείται από γραφικά, κείμενο, ήχο και URL's. Ο επεξεργαστής σχολιασμών δίνει την δυνατότητα στο χρήστη να δημιουργήσει και να τροποποιήσει γραφικά και κείμενο, ενώ μπορεί ταυτοχρόνως να ακούσει τον ήχο και να



παρουσιάσει τα URL's σε ένα νέο φυλλομετρητή. Χρησιμοποιώντας κανείς τον επεξεργαστή σχολιασμών μπορεί να γράψει ένα εισάγει κείμενα και εικόνες από το δίκτυο και στη συνέχεια να τα επεξεργαστεί. Φυσικά πάντα μπορεί να δημιουργήσει και δικά του. Οι πρώτες προσπάθειες προς αυτήν την κατεύθυνση, βασίζονταν σε κλασσικές μεθόδους δεικτοδότησης και αναζήτησης των συμβατικών βάσεων δεδομένων.

Η πιο απλή από αυτές τις προσεγγίσεις ήταν η **αναζήτηση με βάση λέξεις-κλειδιά (keyword-based retrieval)**. Σύμφωνα με αυτήν την μέθοδο, σε κάθε εικόνα της πολυμεσικής βάσης δεδομένων δίνονται κάποιες λέξεις-κλειδιά ή σχόλια υπό την μορφή κειμένου που χαρακτηρίζουν κατά κάποιον τρόπο το περιεχόμενο της. Οι εικόνες ανακτώνται στην συνέχεια από την βάση δεδομένων με ερωτήματα (queries) που κάνουν χρήση αυτών των λέξεων-κλειδιών, όπως ακριβώς συμβαίνει και με την αναζήτηση κειμένων. Αν και η μέθοδος αυτή ήταν εύκολο να υλοποιηθεί καθώς τα συστήματα διαχείρισης αυτών των βάσεων ήταν ήδη διαθέσιμα, σύντομα έγινε αντιληπτό ότι η λύση αυτή συνοδεύεται από κάποια βασικά μειονεκτήματα τα οποία την καθιστούν την αναποτελεσματική. Το πιο σημαντικό από αυτά, είναι η ανάγκη ορισμού από τον άνθρωπο κάποιων λέξεων-κλειδιών ή περιγραφών σε μορφή κειμένου που χαρακτηρίζουν την πολυμεσική οντότητα (manual annotation), διαδικασία ιδιαίτερα χρονοβόρα. Επιπλέον, οι λέξεις αυτές εισάγουν αβεβαιότητα ως προς το συνολικό περιεχόμενο της εικόνας, από την στιγμή που υπόκεινται στην ανθρώπινη υποκειμενικότητα.

Σε μετέπειτα στάδιο, οι μέθοδοι που στηρίζονταν σε λέξεις-κλειδιά εγκαταλείφθηκαν και οι προσπάθειες για καλύτερη διαχείριση των εικόνων στράφηκαν προς την αυτόματη εξαγωγή περιγραφών για το περιεχόμενο της εικόνας και την αξιοποίηση του ίδιου του περιεχομένου της εικόνας κατά την διαδικασία της δεικτοδότησης. Στην **αναζήτηση με βάση-το-περιεχόμενο (content-based image retrieval)** εξάγεται από την εικόνα ένα σύνολο από χαρακτηριστικά χαμηλού επιπέδου που αφορούν συνήθως, τα χαρακτηριστικά του χρώματος, της υφής και του σχήματος του περιεχομένου της εικόνας. Με βάση αυτά τα χαρακτηριστικά γίνεται η δεικτοδότηση των εικόνων σε μια πολυμεσική βάση δεδομένων. Παράλληλα για κάθε χαρακτηριστικό ορίζεται μια τέτοια συνάρτηση απόστασης, που προσεγγίζει κατά τον βέλτιστο τρόπο την ομοιότητα δύο εικόνων όπως την αντιλαμβάνεται ο άνθρωπος. Στην συνέχεια ο χρήστης μπορεί να αναζητήσει εικόνες με βάση οποιοδήποτε από τα πιο πάνω χαρακτηριστικά (ή συνδυασμό αυτών). Έτσι, μπορεί να αναζητήσει εικόνες με βάση κριτήρια της μορφής «βρες όλες τις εικόνες με κατανομή χρώματος 60% μπλε, 40% πράσινο» ή και πιο σύνθετες. Στην μέθοδο αυτή, γίνεται η σιωπηρή υπόθεση ότι είναι εύκολο για τον χρήστη να αντιστοιχίσει σε χαρακτηριστικά χαμηλού επιπέδου (χρώμα, υφή, σχήμα) τα αντικείμενα που περιέχει η εικόνα που τον ενδιαφέρει. Αυτό όμως δεν είναι πάντοτε δυνατόν και η αναζήτηση αυτόματα οδηγείται προς λάθος κατεύθυνση.

Για την αποφυγή αυτού του προβλήματος, αναπτύχθηκαν μέθοδοι που βασίζονταν στην **αναζήτηση με παράδειγμα (query-by-example)**. Πρόκειται για βελτίωση της προηγούμενης μεθόδου, όπου η αναζήτηση γίνεται παρέχοντας μια εικόνα στο σύστημα και αυτό επιστρέφει εικόνες παρόμοιες με αυτήν. Μια παρόμοια μέθοδος βασίζεται στην αναζήτηση με σκίτσο (query-by-sketch), όπου ο χρήστης παρέχει στο σύστημα ένα πρόχειρο σχέδιο για την μορφή της εικόνας/αντικειμένου που ψάχνει. Επιπλέον, σε αυτά τα συστήματα ενσωματώνεται η δυνατότητα ανατροφοδότησης, από την πλευρά του χρήστη (relevance feedback). Η πληροφορία που δίνεται από τον χρήστη πίσω στο σύστημα αφορά την σχετικότητα των αποτελεσμάτων δίνοντας έτσι την δυνατότητα στο σύστημα να προσαρμόζεται στις ανάγκες του χρήστη ενώ εξαλείφει την επίδραση εσφαλμένων χαρακτηριστικών που πιθανόν να δόθηκαν αρχικά. Αυτή η διαδικασία γίνεται επαναληπτικά, κατευθύνοντας έτσι την αναζήτηση προς εκείνο που επιθυμεί ο χρήστης.

Όλες αυτές οι προσεγγίσεις, συνέβαλαν στην αποδοτικότερη αναζήτηση και σε κάποιο βαθμό κατάφεραν να αναπαραστήσουν το περιεχόμενο των εικόνων. Η περιγραφή όμως αυτή, απείχε ακόμη πολύ από τον τρόπο με τον οποίο ο άνθρωπος αντιλαμβάνεται μια εικόνα.

Προέκυψε λοιπόν η ανάγκη για ανάπτυξη μεθόδων ανάλυσης, όπου από τα χαρακτηριστικά χαμηλού επιπέδου θα μπορούσαμε να περάσουμε σε έννοιες πιο κατανοητές από τον άνθρωπο. Η ανάλυση αυτή αναφέρεται ως **σημασιολογική ανάλυση των πολυμέσων (semantic multimedia analysis)** και αποσκοπεί στην γεφύρωση του χάσματος (semantic gap) μεταξύ της αναπαράστασης της

εικόνας με χαρακτηριστικά χαμηλού επιπέδου και της αυτόματης αντιστοίχισης αυτών σε έννοιες με υψηλό σημασιολογικό περιεχόμενο που περιγράφουν την εικόνα με ένα τρόπο πιο κατανοητό στον άνθρωπο. Η εξαγωγή περιγραφών για το περιεχόμενο της εικόνας γίνεται με την χρήση των χαρακτηριστικών χαμηλού επιπέδου των αντικειμένων, καθώς και με **κάποια γνώση εκ των προτέρων (a priori knowledge)**. Με αυτήν, εννοούμε την συλλογή ποιοτικών και ποσοτικών χαρακτηριστικών χαμηλού επιπέδου για τα αντικείμενα του πεδίου ενδιαφέροντος καθώς και τις σχέσεις μεταξύ τους. Τα δεδομένα αυτά διαμορφώνονται μετά από θεωρητική μελέτη της υπό εξέταση θεματικής ενότητας και ποικίλουν ανάλογα με τον σκοπό που προορίζονται. Έχουν προταθεί μια σειρά από μεθόδους με τις οποίες μπορεί να γίνει χρήση της γνώσης εκ των προτέρων και αυτές θα μπορούσαν να χωριστούν σε δύο κατηγορίες, ανάλογα με τον τρόπο τον οποίο αναπαρίσταται αυτή η γνώση.

Στην πρώτη προσέγγιση, οι μέθοδοι που αναπτύχθηκαν βασίζονται σε τεχνικές **μηχανικής μάθησης (machine learning)**, όπου σε αυτές έπειτα από μια διαδικασία εκμάθησης με χρήση αυτή της γνώσης, το σύστημα καταφέρνει να αναπαραστήσει τα χαρακτηριστικά και τις σχέσεις των αντικειμένων του πεδίου ενδιαφέροντος. Οι πιο συχνά χρησιμοποιούμενες τεχνικές στηρίζονται σε κρυφά μαρκοβιανά μοντέλα (Hidden Markov Models), σε νευρωνικά δίκτυα (Neural Networks) και σε ασαφή συστήματα (Fuzzy Systems). Οι μέθοδοι αυτής της κατηγορίας πλεονεκτούν σε εκείνες τις περιπτώσεις που, η διαθέσιμη γνώση είναι μεγάλης έκτασης και πολύπλοκη ώστε να μπορέσει να αναπαρασταθεί ρητά. Όμως συχνά δεν είναι αποδοτική, καθώς για την επέκταση της εφαρμογής ή την χρησιμοποίηση της σε άλλο θεματικό πεδίο, απαιτείται συνεχής προσαρμογή και επανεκπαίδευση του συστήματος.

Με βάση την δεύτερη προσέγγιση, η γνώση εκ των προτέρων αναπαρίσταται με την **μορφή μοντέλων (model based)** των αντικειμένων, των γεγονότων και των σχέσεων μεταξύ τους σε μια κατάλληλη δομή. Σε αυτήν την προσέγγιση, η διαθέσιμη γνώση αναπαρίσταται ρητά χρησιμοποιώντας κάποια γλώσσα αναπαράστασης γνώσης η οποία μπορεί να ποικίλει από ad hoc μέχρι δομημένη αναπαράσταση, π.χ. Description Logics, αντικειμενοστραφή μοντέλα κ.τ.λ. Μέσω αυτής της αναπαράστασης, τα χαρακτηριστικά χαμηλού επιπέδου αποκτούν νόημα με την αντιστοίχιση τους σε έννοιες υψηλού σημασιολογικού περιεχομένου. Αυτή η προσέγγιση έχει το πλεονέκτημα ότι μας δίνει την ευχέρεια να έχουμε πρόσβαση σ' αυτήν την γνώση, με αποτέλεσμα να μπορούμε εύκολα να την εμπλουτίζουμε ή να την προσαρμόζουμε ανά περίπτωση.

## 2.3 Αυτοματοποιημένα-Μη Αυτοματοποιημένα Σχόλια Εικόνας

Υπάρχουν αυτοποιημένα σχόλια εικόνας δηλαδή αυτόματη τοποθέτηση πινακίδων (tagging) εικόνας όπου η διαδικασία με την οποία ένα σύστημα ηλεκτρονικού υπολογιστή εκχωρεί αυτόματα τα μεταδεδομένα με την μορφή λεζάντες (captioning) ή λέξεων κλειδιά (keywords) σε μια ψηφιακή εικόνα. Αυτή η εφαρμογή όρασης του υπολογιστή που χρησιμοποιείται στην ανάκτηση εικόνας για συστήματα είναι για να οργανώσει και να εντοπίσει τις εικόνες που παρουσιάζουν ενδιαφέρον από μια βάση δεδομένων. Τα πλεονεκτήματα του αυτόματου σχολιασμού εικόνας σε αντίθεση με την ανάκτηση βάση του περιεχομένου εικόνας είναι ότι τα ερωτήματα μπορούν να οριστούν με πιο φυσικό τρόπο ο οποίος ορίζεται από τον χρήστη. Ορισμένες μηχανές σχολιασμού είναι online. Οι παραδοσιακές μέθοδοι ανάκτησης εικόνων, όπως αυτές που χρησιμοποιούνται από τις βιβλιοθήκες βασίζονται στο χειροκίνητο σχολιασμό εικόνας, η οποία είναι δαπανηρή και χρονοβόρα, ιδίως λόγω της μεγάλης και συνεχώς αυξανόμενης βάσης δεδομένων εικόνων που υπάρχουν.

Ως εκ τούτου, ο χρήστης-σχολιαστής θα πρέπει να υποστηριχθεί σε αυτό το έργο όσο το δυνατόν περισσότερο. Παρά τις βελτιώσεις στον τομέα, ο αυτόματος σχολιασμός των εικόνων δεν είναι εφικτός αυτή τη στιγμή. Αυτό οφείλεται στο γεγονός ότι αυτό που απεικονίζεται σε μια εικόνα είναι πολύ υποκειμενικό. Αυτό καθιστά ένα καλό σημείο εκκίνησης για ημι-αυτοματοποιημένο σχολιασμό. Το σύστημα που προτείνετε βοηθά το χρήστη να προσθέσει χωρικές πληροφορίες στην εικόνα σχολιασμού. Για το σκοπό αυτό, χρησιμοποιούμε μια συλλογή από πίνακες ζωγραφικής σχολιασμένους με τα αντικείμενα που είναι ορατά σε αυτούς. Για την εξαγωγή κάθε χρώματος ζωγραφικής και χαρακτηριστικών υφής έχουν χρησιμοποιηθεί φίλτρα Gabor. Η στρατηγική του ημι-

αυτοματοποιημένου σχολιασμού εικόνας είναι καλύτερη από σημειώσεις του χρήστη όσον αφορά την αποτελεσματικότητα και καλύτερη από ό, τι ο αυτόματος σχολιασμός όσον αφορά την ακρίβεια.

## ΚΕΦΑΛΑΙΟ 3

### 3.1 Εισαγωγή

Χιλιάδες αντικείμενα καταλαμβάνουν τον οπτικό κόσμο στον οποίο ζούμε. Ο Biederman εκτιμά ότι οι άνθρωποι μπορούν να αναγνωρίσουν περίπου 30000 κατηγορίες αντικειμένων entry-level. Πρόσφατη εργασία στην μηχανική όραση έδειξε εντυπωσιακά αποτελέσματα για την ανίχνευση και αναγνώριση μερικών διαφορετικών κατηγοριών αντικειμένων. Ωστόσο, το μέγεθος και το περιεχόμενο των υφιστάμενων σειρών δεδομένων, μεταξύ άλλων παραγόντων, περιορίζει τις τρέχουσες μεθόδους από την κλιμάκωση σε χιλιάδες κατηγορίες αντικειμένων. Η έρευνα στον τομέα της ανίχνευσης και αναγνώρισης αντικειμένων θα επωφελούνταν από τις μεγάλες εικόνες και τις συλλογές βίντεο με ετικέτα ground truth (επίγεια αλήθειας) συνδέοντας πολλές διαφορετικές κατηγορίες αντικειμένων σε ολοκληρωμένες σκηνές. Για κάθε αντικείμενο το οποίο απεικονίζετε σε μία εικόνα, οι ετικέτες πρέπει να παρέχουν πληροφορίες σχετικά με την ταυτότητα, το σχήμα και τη θέση του αντικειμένου πιθανόν και άλλες ιδιότητες, όπως η στάση.

Κατ'αναλογία προς την ομιλία και τις γλωσσικές κοινότητες, η ιστορία έχει δείξει ότι η απόδοση αυξάνεται εντυπωσιακά όταν παρέχονται περισσότερα επονομαζόμενα δεδομένα κατάρτισης. Κάποιος μπορεί να υποστηρίξει ότι αυτός είναι ένας περιορισμός των τρεχουσών τεχνικών εκμάθησης με συνέπεια το πρόσφατο ενδιαφέρον για τις προσεγγίσεις στην εκμάθηση του Bayes και την εκμάθηση πολλαπλών στόχων (multi-task). Παρ'όλα αυτά, ακόμη και αν μπορούμε να μάθουμε μόνο από έναν μικρό αριθμό παραδειγμάτων, υπάρχουν ακόμη πολλά μαθήματα για να μάθουμε.

Σύνολα δεδομένων μεγάλης εικόνας με ετικέτες ground truth (επίγεια αλήθειας) είναι χρήσιμα για την επιβλεπόμενη μάθηση των κατηγοριών αντικειμένου. Πολλοί αλγόριθμοι έχουν αναπτυχθεί για σύνολα δεδομένων εικόνας, όπου όλα τα παραδείγματα κατάρτισης έχουν το αντικείμενο του ενδιαφέροντος καλά ευθυγραμμισμένο με τα άλλα παραδείγματα. Αλγόριθμοι που εκμεταλλεύονται το πλαίσιο για την αναγνώριση αντικειμένων θα επωφελούνταν από σύνολα δεδομένων με πολλές επονομαζόμενες κατηγορίες αντικειμένου που ενσωματώνονται στις σύνθετες σκηνές. Τέτοια σύνολα δεδομένων πρέπει να περιέχουν μια ευρεία ποικιλία των περιβαλλόντων με τα σχολιασμένα αντικείμενα που επανεμφανίζονται στις ίδιες εικόνες. Κατά τη σύγκριση διαφορετικών αλγορίθμων για την ανίχνευση και αναγνώριση αντικειμένων τα σχολιασμένα δεδομένα είναι απαραίτητη η ποσοτική μέτρηση των επιδόσεών τους. Ακόμα και όταν οι αλγόριθμοι δεν χρειάζονται επίβλεψη είναι αναγκαίο ένα ποσοτικό πλαίσιο.

Η δημιουργία ενός μεγάλου συνόλου δεδομένων από σχολιασμένες εικόνες με πολλά αντικείμενα είναι ένα δαπανηρό και χρονοβόρο εγχείρημα. Παραδοσιακά τα σύνολα δεδομένων δημιουργούνται από μία και μόνο ερευνητική ομάδα η οποία είναι προσαρμοσμένη στην επίλυση συγκεκριμένων προβλημάτων. Κατά συνέπεια πολλά τρέχοντα διαθέσιμα σύνολα δεδομένων περιέχουν μόνο ένα μικρό αριθμό κατηγοριών όπως πρόσωπα, πεζούς και αυτοκίνητα. Αξιοσημείωτη εξαίρεση είναι το Caltech σύνολο δεδομένων, με 101 κατηγορίες αντικειμένων (το οποίο πρόσφατα

επεκτάθηκε σε 256 κατηγορίες αντικειμένων), η συλλογή Pascal και η βάση δεδομένων CBCL-σκηνών δρόμου.

Ευχής έργο θα ήταν η συλλογή μεγάλων συνόλων δεδομένων από σχολιασμένες εικόνες. Για την επίτευξη αυτού του στόχου μελετούμε μεθόδους συλλογής δεδομένων που βασίζονται στο Διαδίκτυο. Τέτοια εργαλεία σχολιασμού παρέχουν τη δυνατότητα δημιουργίας μεγάλων συνόλων δεδομένων σχολιασμού, στηριζόμενα στη συλλογική προσπάθεια μεγάλου πλήθους χρηστών. Πρόσφατα τέτοιες προσπάθειες είχαν μεγάλη επιτυχία. Η Open Mind Initiative στοχεύει στη συλλογή μεγάλων συνόλων δεδομένων από τους χρήστες του Διαδικτύου έτσι ώστε να αναπτυχθούν ευφυείς αλγόριθμοι. Πιο συγκεκριμένα γεγονότα κοινής λογικής καταγράφονται (π.χ. το κόκκινο είναι βασικό χρώμα), με πάνω από 700K γεγονότα καταγεγραμμένα μέχρι σήμερα.

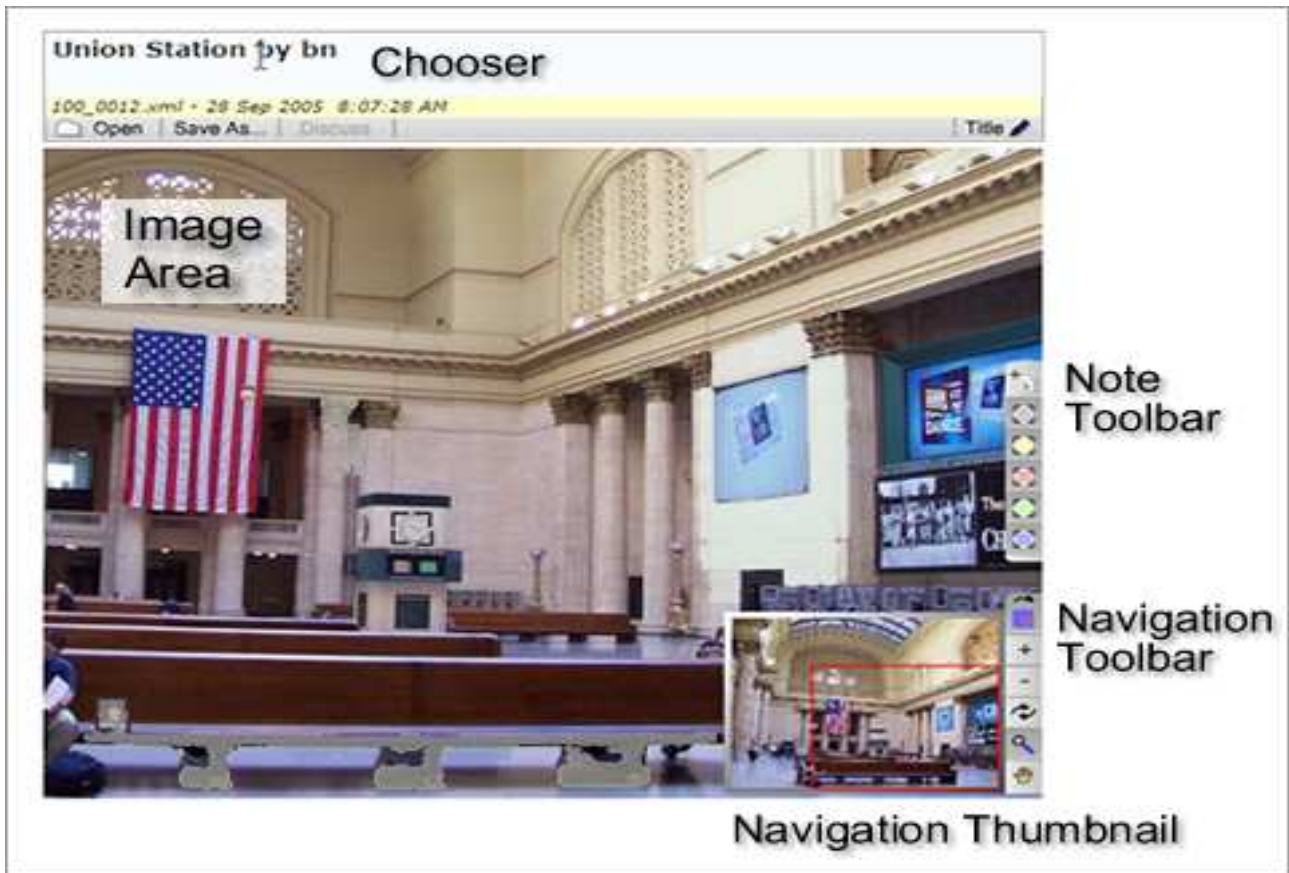
Σε αυτό το κεφάλαιο θα γίνει η παρουσίαση των πιο σημαντικών εργαλείων σημασιολογικού σχολιασμού εικόνων, όπως το Image annotation Tool, το Anvil, το Annotate Expert, το FotoTagger, το M-OntoMat-Annotizer, το Annotea, το Video Image Annotation Tool, το Label Me, το FotoNotes και το IAT.

## 3.2 Image Annotation Tool

Το Image Annotation Tool (Εργαλείο Σχολιασμού Εικόνας) μας επιτρέπει να δούμε μία μεγέθυση υψηλής ανάλυσης, της εικόνας και να προσθέσουμε "sticky notes" (επικόλουμενες σημειώσεις) για τον σχολιασμό των χαρακτηριστικών της.

Υπάρχουν πέντε τομείς του εργαλείου της διεπαφής του χρήστη:

1. Χρησιμοποιήστε τον **Επιλογέα (Chooser)** για να ανοίξετε ένα έγγραφο σχολιασμού εικόνας, να αλλάξετε τον τίτλο ενός εγγράφου, ή να αποθηκεύσετε ένα αντίγραφο.
2. Η **Περιοχή Εικόνας (Image Area)** εμφανίζει μια μεγέθυση της εικόνας και δείκτες σχολιασμού που συνδέουν τις σημειώσεις με τα χαρακτηριστικά της εικόνας.
3. Χρησιμοποιήστε την **Εργαλειοθήκη Σημειώσεων (Note Toolbar)** για να δημιουργήσετε νέες σημειώσεις και να ελέγξετε την ορατότητα των σημειώσεων ανά χρώμα.
4. Χρησιμοποιήστε τη **Εργαλειοθήκη Πλοήγησης (Navigation Toolbar)** για μεγέθυση, σμίκρυνση και αλλαγή του ορατού τμήματος της εικόνας.
5. Η **Thumbnail** δείχνει τι μέρος μιας εικόνας που προβάλλεται. Η **Thumbnail** μπορεί να κρυφτεί ή να εμφανιστεί.



Εικόνα 1

Το Image Annotation Tool μας επιτρέπει τη δημιουργία ή την επεξεργασία εγγράφων σχολιασμού. Σκεφτείτε ένα έγγραφο σχολιασμού ως ένα είδος επεξεργασίας κειμένου το οποίο περιέχει μία συλλογή από σημειώσεις για μία εικόνα. Τα έγγραφα σχολιασμού είναι απλά αρχεία τύπου XML ,τα οποία μπορούν να μετακινηθούν από μέρος σε μέρος. Τα έγγραφα αυτά δεν περιλαμβάνουν τις εικόνες. Πολλοί άνθρωποι μπορούν να δημιουργήσουν ένα έγγραφο σχολιασμού με σημειώσεις συνημμένες σε μία μοναδική εικόνα. Όλα αυτά που θα είναι στο έγγραφο σχολιασμού είναι σημειώσεις και μία διεύθυνση URL επισημαίνοντας τη θέση της εικόνας.

### 3.3 ANVIL

Το ANVIL είναι ένα δωρεάν εργαλείο σχολιασμού βίντεο το οποίο μπορούμε να κατεβάσουμε κατόπιν αιτήσεως. Προσφέρει σχόλια πολλαπλών επιπέδων και βασίζεται σε σύστημα κωδικοποίησης καθορισμένο από το χρήστη .Κατά τη διάρκεια της κωδικοποίησης ο χρήστης μπορεί να δει χρωματισμένα στοιχεία σε πολλαπλά κομμάτια στη διάρκεια του χρόνου. Κάποια ιδιαίτερα χαρακτηριστικά είναι οι σύνδεσμοι μεταξύ των διαφόρων επιπέδων, μη χρονικά αντικείμενα, κομμάτια χρονικών σημείων, η συμφωνημένη ανάλυση κωδικοποίησης και ένα εργαλείο για τη διαχείριση όλου του σώματος των αρχείων σχολιασμών. Αρχικά αναπτύχθηκε για την έρευνα χειρονομιών, το 2000, το ANVIL τώρα χρησιμοποιείται σε πολλούς τομείς έρευνας, συμπεριλαμβανομένης της αλληλεπίδρασης ανθρώπου-υπολογιστή, της γλωσσολογίας, της ηθολογίας, της ανθρωπολογίας, της ψυχοθεραπείας, ενσωματωμένων πρακτόρων, ζωτικότητας των υπολογιστών και της ωκεανογραφίας.

Το ANVIL μπορεί να εισάγει δεδομένα από φωνητικά εργαλεία όπως το Praat που επιτρέπουν ακριβή και άνετη καταγραφή ομιλίας. Το Anvil μπορεί να εμφανίσει κυματομορφές και περιγράμματα τόνων ήχων. Τα αρχεία δεδομένων του Anvil βασίζονται σε XML. Οι εξαγόμενοι πίνακες μπορούν να χρησιμοποιηθούν για την ανάλυση των στατιστικών εργαλείων όπως το SPSS. Η επερχόμενη έκδοση θα είναι επίσης σε θέση να εισάγει αρχεία ELAN. Το ANVIL είναι γραμμένο σε Java και τρέχει σε Windows, Macintosh και πλατφόρμες Unix. Η τελευταία έκδοση είναι ANVIL 5, είναι ακόμη δοκιμαστική, αλλά πολύ σταθερή.



Εικόνα 2

## Χαρακτηριστικά

ANVIL5 θα κυκλοφορήσει στα τέλη του 2010. Ωστόσο η Beta έκδοση (δοκιμαστική έκδοση) είναι σταθερή και χρησιμοποιείται από τους περισσότερους χρήστες. Παρακάτω είναι μια λίστα με νέα χαρακτηριστικά που η beta έκδοση περιλαμβάνει:

1. **Installer:** για πλατφόρμες Windows και Mac, ειδικοί εγκαταστάτες (installers) κάνουν την εγκατάσταση εύκολη.
2. **Integrated online manual** (Ολοκληρωμένο ηλεκτρονικό εγχειρίδιο) : η τεκμηρίωση είναι πλέον διαθέσιμη στο ίδιο το εργαλείο πιέζοντας το πράσινο κουμπί βοήθειας
3. **Specification editor** (Προδιαγραφές επεξεργαστή): προδιαγραφές (σύστημα κωδικοποίησης), μπορούν να δημιουργηθούν και να αλλάξουν με ένα επεξεργαστή γραφικών.
4. **3D Motion capture player** (τριδιάστατη απαθανάτιση κινούμενου «παίχτη» ) : αρχεία καταγραφής της κίνησης σε μορφή BVH μπορεί να εμφανιστούν με ένα 3D σκελετό και να συγχρονιστούν με αναπαραγωγή βίντεο.
5. **Subdivision track type** (Υποδιαίρεση Τύπου τροχιάς): αυτό είναι ισοδύναμο με τον τύπο υποδιαίρεση ELAN του. Σε ένα κομμάτι υποδιαίρεση του κωδικοποιητή χωρίζει ένα στοιχείο για την τροχιά αναφοράς σε μικρότερα κομμάτια.

6. **Time point track type** (Χρονικό σημείο τύπου τροχιάς) : το «Πρωταρχικό Σημείο» τύπου τροχιάς επιτρέπει τον σχολιασμό μόνο για χρονικά σημεία αντί διαστήματα. Αυτό ισοδυναμεί με τον τύπο σημείου του Praat.
7. **ELAN Import** (Εισαγωγή ELAN) : ELAN αρχεία μπορούν να διαβαστούν σε Anvil. Διάφορες στρατηγικές μετατροπής εφαρμόζονται για την προσαρμογή των διαφορών μεταξύ των εργαλείων (σημειώστε ότι τύπος υποδιαίρεση ELAN έχει επίσης προστεθεί στο Anvil).
8. **Histograms** (Ιστογράμματα) : οι χρήστες μπορούν γρήγορα να εξετάσουμε την κατανομή των ετικετών σε μια συγκεκριμένη διαδρομή, αλλά και διά μέσου σχολιασμών.
9. **Analysis** (Ανάλυση) : Το Anvil 5 προσφέρει διάφορες χρήσιμες δυνατότητες ανάλυσης που μπορούν να εφαρμοστούν σε ένα μοναδικό αρχείο Anvil ή σε ολόκληρη την εργασία
  - **Coding agreement (kappa)** (συμφωνία Κωδικοποίησης (κάππα)) : αυτόματος υπολογισμός της συμφωνίας μεταξύ δύο σχολιασμών του ίδιου κομματιού (δύο Anvil αρχεία.). Υπολογίζει το Κάπα Κοέν, διορθώνει το Κάπα και δημιουργεί ενιαίο ποσοστό. Επίσης, επιτρέπει να δούμε μόνο την συμφωνία κατακερματισμού.
  - **Transition diagrams** (Διαγράμματα Μετάβασης): οπτικοποίηση των πιθανοτήτων η ετικέτα A να μεταβαίνει στην ετικέτα B σε ένα συγκεκριμένο κομμάτι, εμφανίζεται ως μια γραφική παράσταση, όπου το πάχος του βέλους δείχνει τις πιθανότητες μετάβασης.
  - **Association analysis** ( Ανάλυση Συνάφειας) : Επιτρέπει τον υπολογισμό ενός chi square test, για περιστατικά που ξαναεμφανίζονται σε ετικέτες δύο διαφορετικών κομματιών. Ο χρήστης καθορίζει τη χρονική συσχέτιση που οι οντότητες υπό σύγκριση πρέπει να πληρούν.
- 10 **Improved search interface** (Βελτιωμένη Διεπαφή αναζήτησης): Αναζήτηση κειμένου εμφάνισης Google είναι προεπιλεγμένη μαζί με αναζήτηση συνήθους έκφρασης και την παλαιότερη αναζήτηση λίστας συμπερίληψης/απόκλισης
- 11 **Comparing tracks** (Σύγκριση Ιχνών) Κάποιος μπορεί να εισάγει μία διαδρομή από ένα διαφορετικό σχολιασμό (δεξί κλικ στο όνομα της διαδρομής).

### 3.4 Annotate Expert

Το Annotate Expert είναι ένα δωρεάν λογισμικό που παρέχεται από την Nikonians μας επιτρέπει να βάλουμε ετικέτα στις εικόνες, λέξεις-κλειδιά και σχόλια που περιορίζονται σε συγκεκριμένα τμήματα της εικόνας, **για να τονίσει τμήματα της** με την τοποθέτηση "**εικονικών συγκομιδών**" και να μοιράσουμε τις φωτογραφίες μας με άλλους, πολύ εύκολα.. Σκεφτείτε το γράφημα σε έναν υαλοπίνακα που βρίσκεται στην κορυφή της εικόνας σας. Τα σχόλια αποθηκεύονται με ένα μη-φθαρτό τρόπο στην κεφαλίδα (**EXIF**) ως μέρος του αρχείου της εικόνας σε μορφή XMP, **χωρίς να αλλάζει την ίδια την εικόνα.** Τα σχόλια που γίνονται στις εικόνες είναι ορατά μόνο εάν ανοίξουμε το αρχείο στο Annotate, εκτός αν εξάγουμε τα σχόλια και τα ορίσουμε ως διαθέσιμα σε όλους. Το πρόγραμμα μπορεί να χρησιμοποιηθεί για εκπαιδευτικούς σκοπούς, επιτρέποντας σε ένα ή περισσότερους φωτογράφους να δώσουν κριτικές σε μια φωτογραφία, συμπεριλαμβανομένων των προτεινόμενων βελτιώσεων τους, όπως τα σχόλια απευθείας πάνω στην εικόνα. Ένα άλλο σενάριο είναι όταν μια ομάδα εργάζεται σε μια συλλογική προσπάθεια με εικόνες, όπως σε ένα πρόγραμμα βιβλίο-εκδόσεις, όπου προσθέτουν σημειώσεις και σχόλια για τις εικόνες χωρίς την ανάγκη δημιουργίας e-mails επιπλέον PDF ή με τα σχόλια.

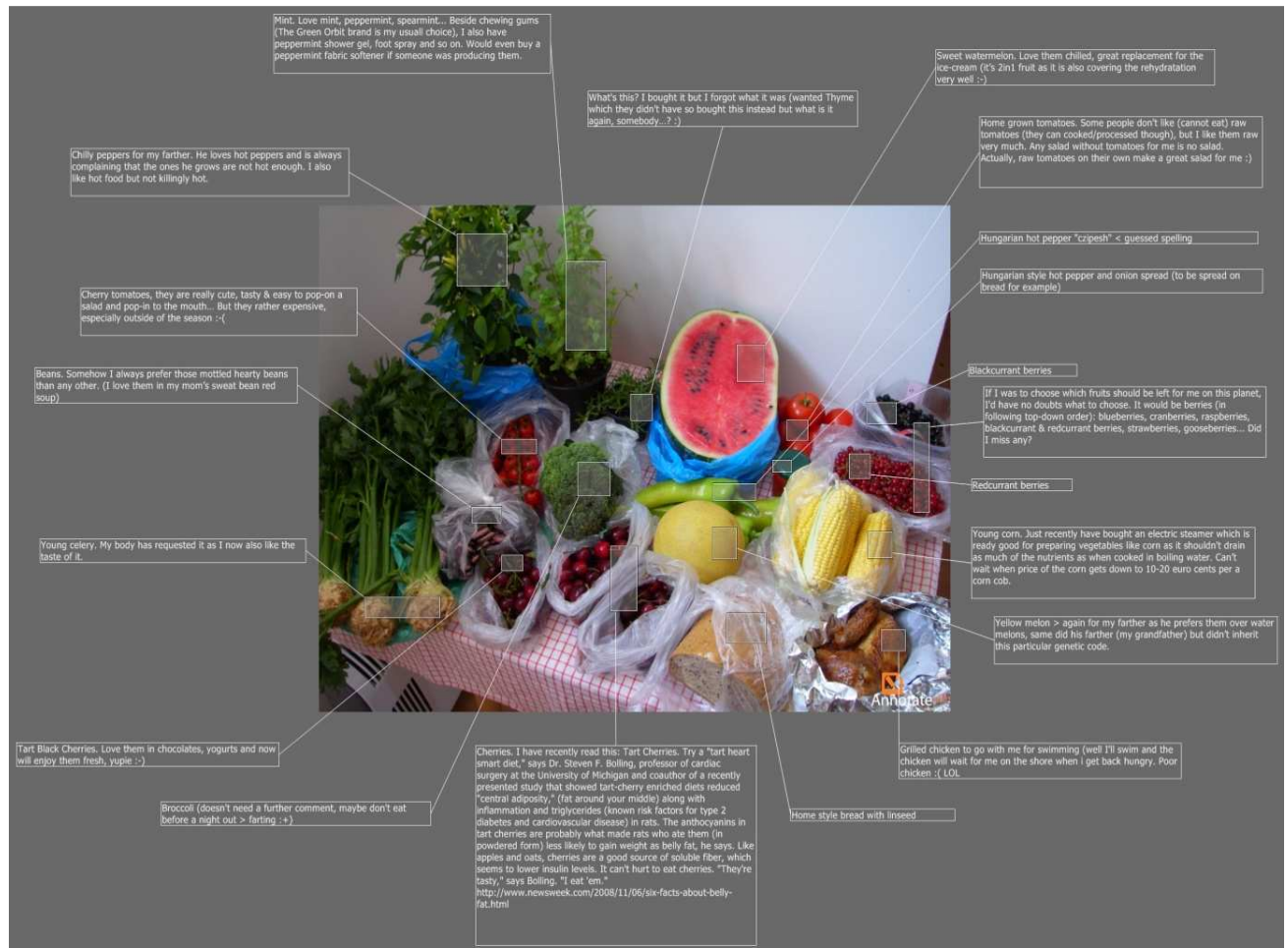




Εικόνα 3

## Χαρακτηριστικά

1. **Ενσωματωμένα σχόλια** : Τα ενσωματωμένα σχόλια δεν χάνονται ούτε αποσυνδέονται από την εικόνα. Όποτε η εικόνα αντιγράφετε ή αποστέλλετε τα σχόλια την «ακολουθούν» έτσι ώστε εάν ανοιχθεί με το Annotate τα σχόλια τα εμφανίζονται.
2. **Σχόλια ήχου**: Είναι δυνατή η επισύναψη ηχητικών σχολίων σε μία εικόνα.
3. **IPTC ολοκλήρωση λέξεων-κλειδιά**: Οι λέξεις κλειδιά ενημερώνονται με τις ετικέτες σας και διαβάζονται από τις υπάρχουσες λέξεις κλειδιά κατά το άνοιγμα μίας εικόνας για τον πρώτο σχολιασμό.
4. **Επίπεδα σχολίων**: Επιτρέπει την εύκολη κατηγοριοποίηση των σχολίων κατά σκοπό όπως μετρήσεις, υποδείξεις ποιότητας.
5. **Ανιχνευσιμότητα** : Για να επιτρέπετε η ανιχνευσιμότητα κάθε σχόλιο αποθηκεύεται κατά χρήστη και επιπρόσθετα με χρονικό προσδιορισμό στο ιστορικό καταγραφής.
6. **Έξυπνη αντιγραφή**: Κατά το άνοιγμα μίας μη σχολιασμένης εικόνας το Annotate Expert μπορεί να μεταβιβάσει τα σχόλια μίας παρόμοιας εικόνας ως πρότυπο κείμενο.



Εικόνα 4

## Τεχνικές Απαιτήσεις

Το Annotate Expert είναι διαθέσιμο για MS Windows XP, Vista, Windows 7 και Mac OS X υποστηρίζει εικόνες τύπου jpg, png, tif, bmp, psd ενώ επιτρέπει το σχολιασμό μίας εικόνας μέχρι και με 60.000 χαρακτήρες.

## 3.5 FotoTagger

Το Foto Tagger είναι ένα δωρεάν εργαλείο σχολιασμού εικόνων, το οποίο δημιουργήθηκε από την Cogitum LC. Με το Foto Tagger είναι δυνατή η επισήμανση των εικόνων με τη προσθήκη εύκολα μεταφερόμενων σημειώσεων σε ψηφιακές εικόνες.

Το Foto Tagger επιτρέπει την οργάνωση εικόνων σε άλμπουμ που ορίζονται από το χρήστη. Η προσθήκη εικόνων σε άλμπουμ είναι εύκολη, μπορούμε να μετακινήσουμε και να δημοσιεύσουμε άλμπουμ σχεδόν στιγμιαία. Επίσης προσθέτει επεξηγηματικές σημειώσεις και υπερσυνδέσμους σε σχετικές πηγές στο Διαδίκτυο, μεταφορτώνει σχολιασμένες εικόνες στο λογαριασμό μας στο Flickr μέσα σε δευτερόλεπτα ή ακόμα δημοσιεύει σχολιασμένες εικόνες απευθείας σε μπλοκ. Οι σχολιασμοί δεν αλλάζουν την αρχική εικόνα έτσι η μετακίνηση και το μοίρασμα αρχείων διατηρεί τους σχολιασμούς ανέπαφους. Το Foto Tagger δουλεύει με JPEG.



Εικόνα 5

## Χαρακτηριστικά

Το Foto Tagger αντιμετωπίζει τρεις κύριες προκλήσεις στη διαχείριση ψηφιακών εικόνων παρέχοντας τις παρακάτω δυνατότητες :

1. Διατηρεί τις λεπτομέρειες της εικόνας. Κάθε χρήστης έχει σίγουρα την εμπειρία να ξεχάσει ονόματα των ανθρώπων και αντικειμένων που παρουσιάζονται στις φωτογραφίες. Οι προσπάθειες να συμπεριληφθούν οι σχετικές πληροφορίες ως μέρος ενός ονόματος αρχείου ή η προσθήκη λέξεων κλειδιών είναι συχνά ακατάλληλα και ανεπαρκής όταν πρόκειται για την ταυτοποίηση προσώπων ή αντικειμένων. Από την άλλη πλευρά τα ονόματα και τα σχόλια που τοποθετούνται απευθείας πάνω σε ψηφιακές εικόνες είναι χρήσιμα, μόνο όταν ο χρήστης έχει τη δυνατότητα εύκολα να κρύβει τους σχολιασμούς από την αρχική εικόνα. Αυτό είναι ακριβώς που κάνει το Foto Tagger.

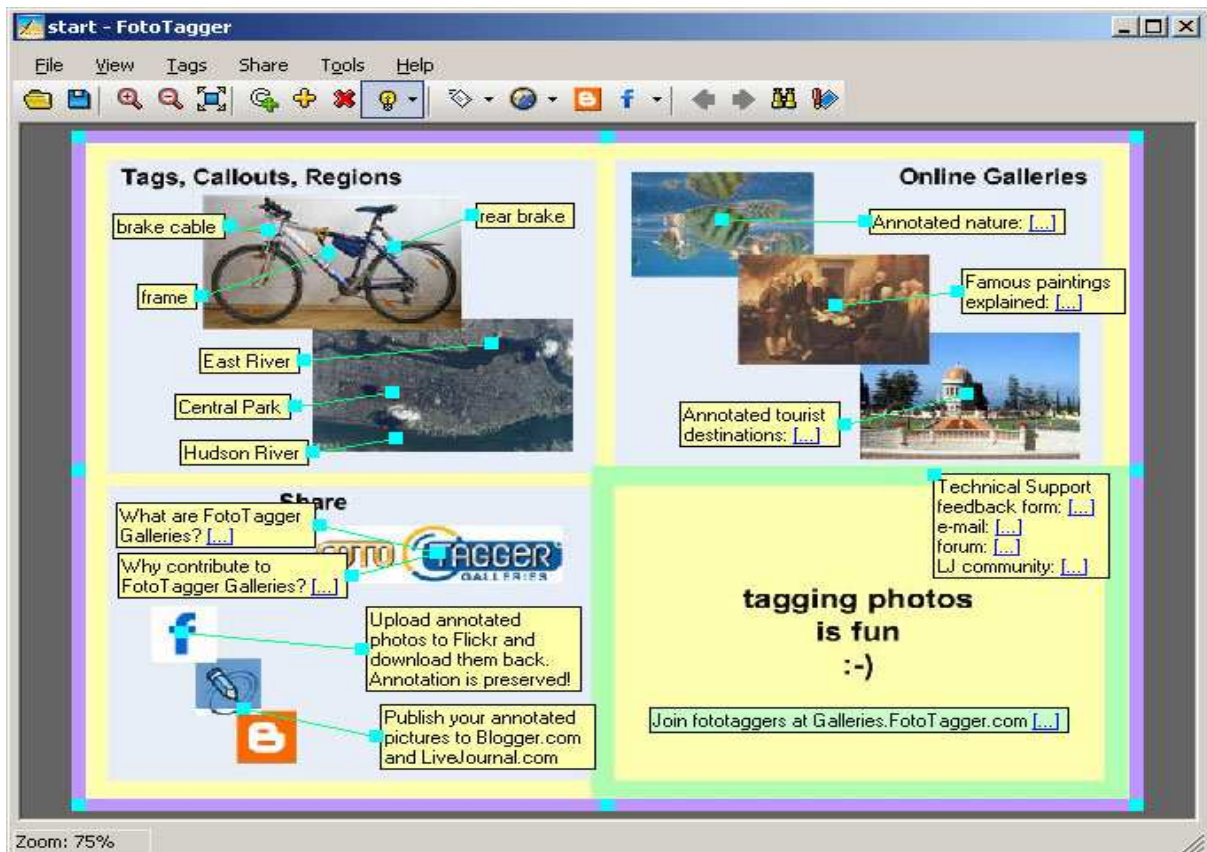
Σήμερα το Foto Tagger χρησιμοποιείται ήδη για το σχολιασμό φωτογραφιών από ταξίδια, οικογενειακές εικόνες και ζωγραφίες.

2. Εύρεση απαιτούμενων εικόνων άμεσα: Για τους περισσότερους χρήστες το χάσιμο χρόνου στη προσπάθεια να βρουν μία φωτογραφία ενός συγκεκριμένου προσώπου ή αντικειμένου είναι σύνθηρες φαινόμενο. Το Foto Tagger παρέχει τη δυνατότητα εύρεσης της επιθυμητής εικόνας από την αναζήτηση σημειώσεων σε μορφή κειμένου. Η δυνατότητα αυτή επιτρέπει στο χρήστη να εκτελέσει μία σημαντική έρευνα και να



βασιστεί σε ότι παρουσιάζεται σε μία φωτογραφία και όχι σε τυπικές παραμέτρους , όπως ένα όνομα αρχείου ή μια ημερομηνία .

3. Κοινή χρήση σχολιασμένων εικόνων με τη χρήση πρότυπων μπλοκ: Πολλοί χρήστες συχνά επιθυμούν να μοιραστούν εικόνες μαζί με κάποιες παρατηρήσεις και επεξηγήσεις με μορφή κειμένου. Η πραγματική αξία ανταλλαγής φωτογραφιών πολλές φορές δεν μετριέται μόνο από το πλήθος των φωτογραφιών. Σε ορισμένες περιπτώσεις καθορίζεται από μερικές φωτογραφίες με λογικό σχολιασμό. Αυτός είναι ο λόγος που το FotoTagger υποστηρίζει τη δημοσίευση σε μπλοκ ως μία μορφή κοινής χρήσης φωτογραφιών.

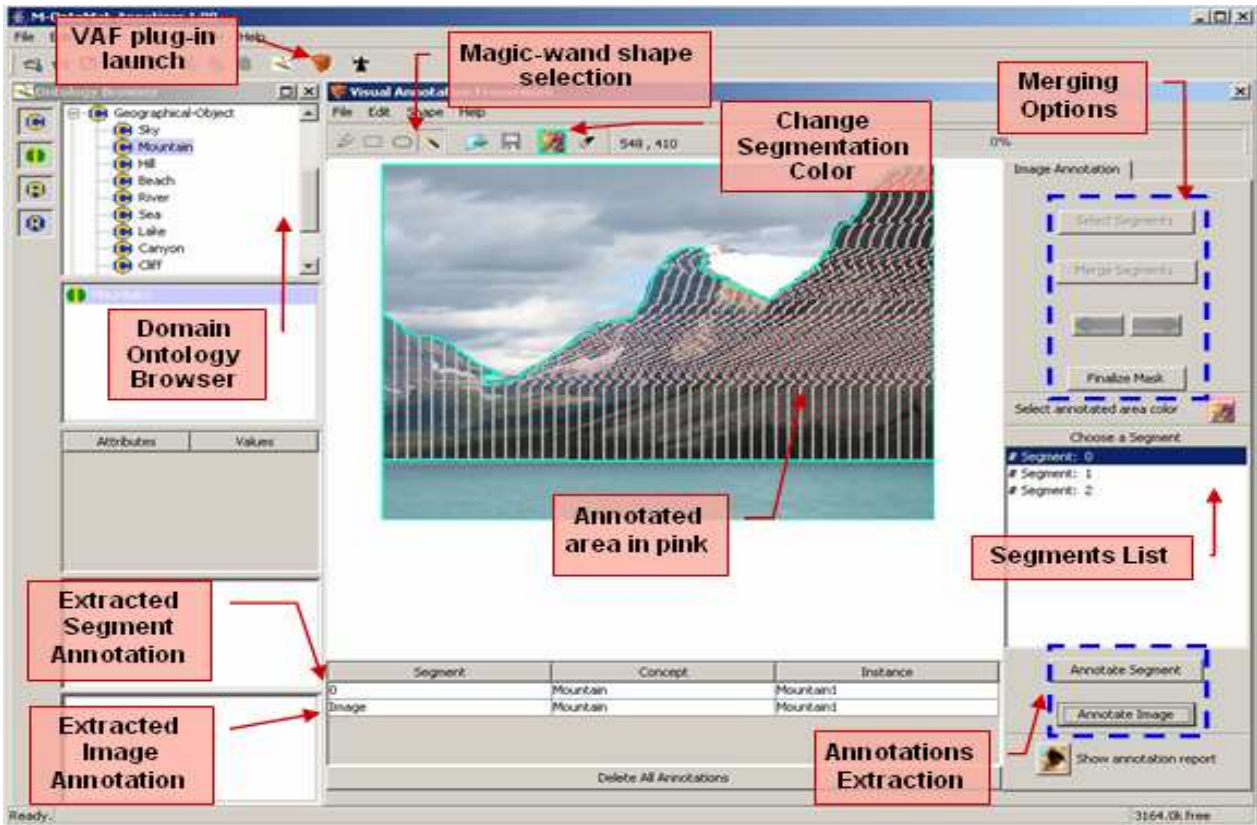


Εικόνα 6

### 3.6 M- OntoMat-Annotizer

Ο σχολιασμός των πολυμεσικών εγγράφων συνήθως ακολουθεί δύο διαφορετικές κατευθύνσεις. Είτε προηγούμενες προσεγγίσεις οι οποίες εστιάζουν σε χαμηλού επιπέδου περιγραφές, όπως το κυρίαρχο χρώμα, είτε εστιάζουν στις εμπεριέχουσες διαστάσεις και τα αντίστοιχα σχόλια όπως άτομα ή οχήματα. Το M- OntoMat-Annotizer επιτρέπει τη σύνδεση χαμηλού επιπέδου MPEG-7 οπτικών περιγραφών με τις συμβατικές σημασιολογικές οντολογίες Διαδικτύου και των σχολιασμών. Το M- OntoMat-Annotizer είναι ένα διαδικτυακό, διαδραστικό και φιλικό προς το χρήστη εργαλείο σχολιασμού. Υποστηρίζει το χρήστη στο έργο του να δημιουργήσει και να διατηρήσει OWL-σημάνσεις βασισμένες στις οντολογίες πχ να δημιουργήσει OWL-στιγμιότυπα ,γνωρίσματα και σχέσεις. Περιλαμβάνει και φυλλομετρητή οντολογιών για την εξερεύνηση των

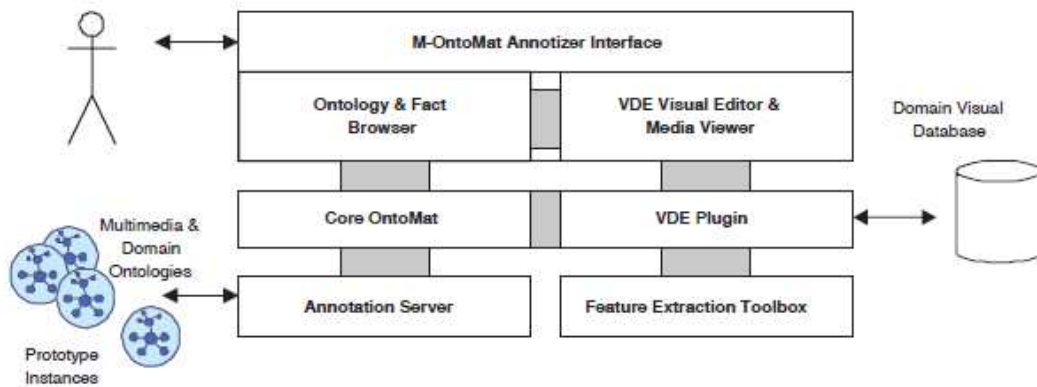
οντολογιών και των στιγμιότυπων και έναν HTML φυλλομετρητή ο οποίος παρουσιάζει τα σχολιασμένα τμήματα του κειμένου. Είναι βασισμένο στη JAVA και παρέχει μία πρόσθετη διεπαφή για επεκτάσεις. Ο χρήστης είναι ανεξάρτητος σχολιαστής π.χ. άτομα τα οποία θέλουν να εμπλουτίσουν τις ιστοσελίδες τους με OWL δεδομένα. Αντί ενός χειροκίνητου σχολιασμού με κειμενογράφο το M- OntoMat-Annotizer επιτρέπει στο σχολιαστή να επισημάνει σχετικά κομμάτια της ιστοσελίδας και να δημιουργήσει νέα στιγμιότυπα μέσω αλληλεπιδράσεων τύπου drag and drop (σύρατε αποθέσατε).



Εικόνα 7

## Χαρακτηριστικά του M- OntoMat-Annotizer

Για να αξιοποιηθούν οι υποδομές των οντολογιών που προαναφέρθηκαν και για τον εμπλουτισμό τους με πολυμεσικές περιγραφές ,το M- OntoMat-Annotizer επεκτείνει το CREAM (CREATING Metadata for the Semantic Web δηλαδή Δημιουργία Μεταδεδομένων για τον Σημασιολογικό Ιστό ).Η εικόνα 8 δείχνει την ολοκληρωμένη αρχιτεκτονική των ενοτήτων οι οποίες εξηγούνται παρακάτω λεπτομερώς.



Εικόνα 8

**Core OntoMat-Annotize** Το M- OntoMat-Annotizer υποστηρίζει δύο βασικές εφαρμογές (i) χρησιμοποιείται σαν εργαλείο σχολιασμού ιστοσελίδων και (ii) λειτουργεί ως βάση ενός περιβάλλοντος εφαρμοσμένης μηχανικής οντολογίας. Επιπλέον παρέχοντας μία ευέλικτη επιπρόσθετη διεπαφή προσφέρει τη δυνατότητα υλοποίησης νέων στοιχείων και την επεκτείνει τη βασική λειτουργικότητα του M- OntoMat-Annotizer.

**Annotation Server** Ο εξυπηρετητής σχολιασμού(Annotation Server) λειτουργεί στο παρασκήνιο και αποθηκεύει τις οντότητες της βάσης των γνώσεων, διατηρεί τις κοινές παραπομπές και είναι υπεύθυνος για τη διατήρηση και συνολική ακεραιότητα των αποθηκευμένων οντοτήτων.

**Domain Visual Database** (Τομέας Οπτικών Βάσεων Δεδομένων) Δεδομένου ότι η εύκολη πρόσβαση περιεχομένου είναι κρίσιμη για τις διαδικασίες ανάλυσης σχολιασμών και περιεχομένου, μία οπτική βάση δεδομένων που περιέχει περιεχόμενο σχετικό με τον τομέα που εξετάζεται και αναλύεται είναι πάντα απαραίτητη.

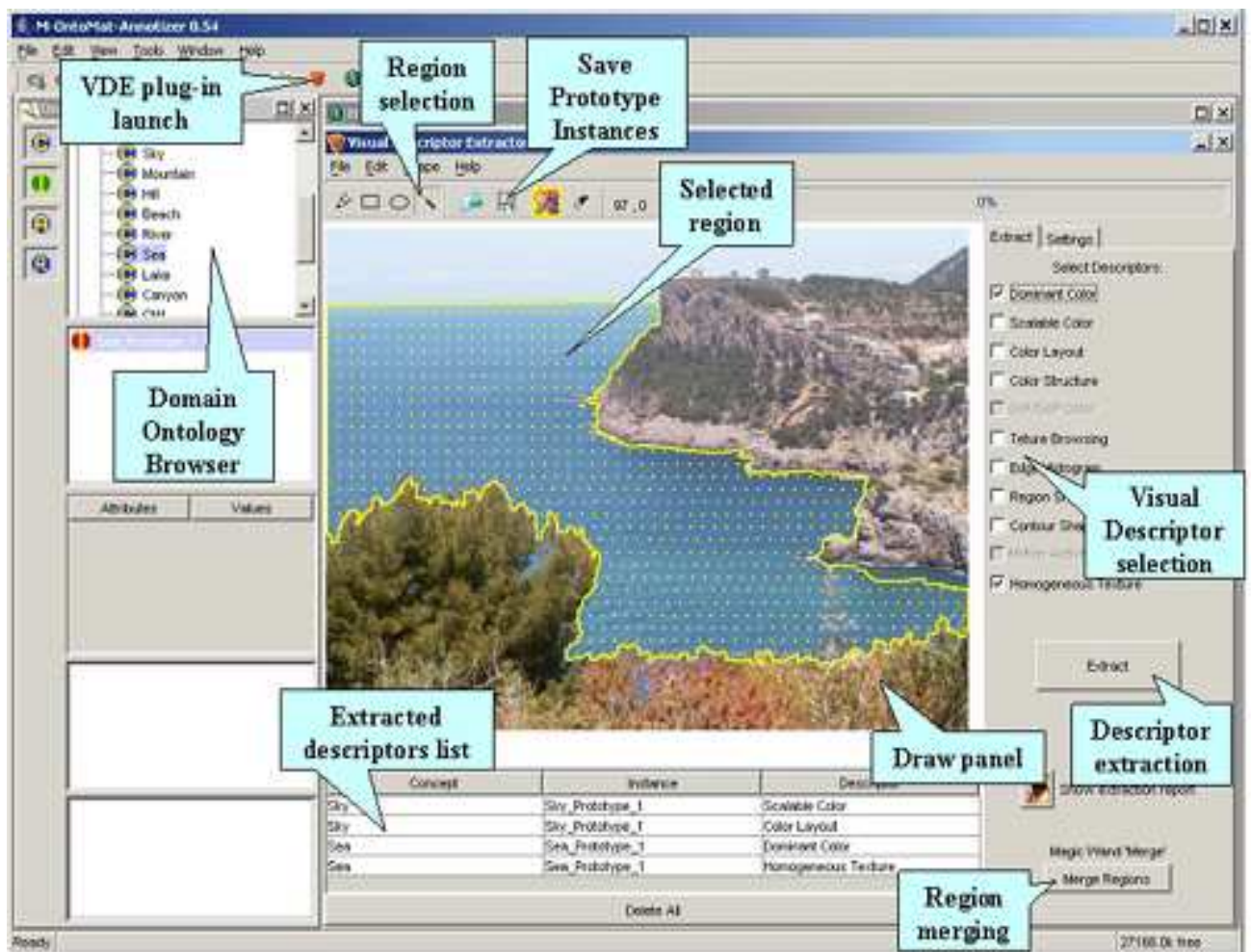
**Feature Extraction Toolbox** Η πραγματική εξαγωγή των οπτικών περιγραφών πραγματοποιείται με τη χρήση μίας εργαλειοθήκης εξαγωγής χαρακτηριστικών, η οποία ονομάζεται aceToolbox, μίας εργαλειοθήκης προεπεξεργασίας περιεχομένου και εξαγωγής χαρακτηριστικών γνωρισμάτων ανεπτυγμένης μέσα σε ένα aceMedia έργο. Η aceToolbox αποθηκεύει τους εξαγόμενους περιγραφείς MPEG-7 σε XML μορφή.

**VDE Visual Editor and Media Viewer** Ο VDE Visual Editor and Media Viewer παρουσιάζει μία γραφική διεπαφή για την φόρτωση και επεξεργασία του οπτικού περιεχομένου (εικόνες και βίντεο), οπτική εξαγωγή χαρακτηριστικών και σύνδεση με το τομέα εννοιών της οντολογίας. Η διεπαφή, συνεργάζεται άσπογα με τις κοινές διεπαφές OntoMat. Συνήθως, ο χρήστης πρέπει να εξάγει τα χαρακτηριστικά (πολυμέσα περιγραφείς) ενός συγκεκριμένου αντικειμένου στο εσωτερικό της εικόνας/πλαισίου. Για το λόγο αυτό, η VDE εφαρμογή επιτρέπει στο χρήστη να σχεδιάζει περιοχές ενδιαφέροντος μέσα στην εικόνα/πλαίσιο και να εφαρμόζει τη διαδικασία εξαγωγής πολυμεσικών περιγραφών μόνο στην ειδικά επιλεγμένη περιοχή. Εναλλακτικά το M- OntoMat-Annotizer υποστηρίζει επίσης αυτόματη κατάτμηση της εικόνας/πλαισίου κάθε φορά που μία νέα εικόνα/πλαίσιο φορτώνετε αυτόματα υποδιαιρείτε σε τμήματα χρήστη μπορεί τότε να επιλέξει ένα



επιθυμητό τμήμα ή ακόμα και να συγχωνεύσει δύο ή περισσότερα τμήματα και να προχωρήσει στην εξαγωγή. Επιλέγοντας μία συγκεκριμένη έννοια στον φυλλομετρητή οντολογιών OntoMat και επιλέγοντας ένα τμήμα ενδιαφέροντος ο χρήστης μπορεί να εξάγει και να συνδέσει έννοιες με τις κατάλληλες περιπτώσεις πρωτοτύπων μέσω των σχετικών λειτουργιών του VDE.

**VDE Plug-in** Το εργαλείο εξαγωγής οπτικών περιγραφών (VDE) υλοποιείται ως πρόσθετη λειτουργία του M- OntoMat-Annotizer και είναι βασικό συστατικό για την επέκταση των δυνατοτήτων του και την υποστήριξη αρχικοποίησης οντολογιών χαμηλού επιπέδου πολυμεσικών χαρακτηριστικών. Η πρόσθετη λειτουργία VDE διαχειρίζεται τη συνολική διαδικασία εξαγωγής και σύνδεσης χαμηλού επιπέδου χαρακτηριστικών με την επικοινωνία με τα άλλα συστατικά.

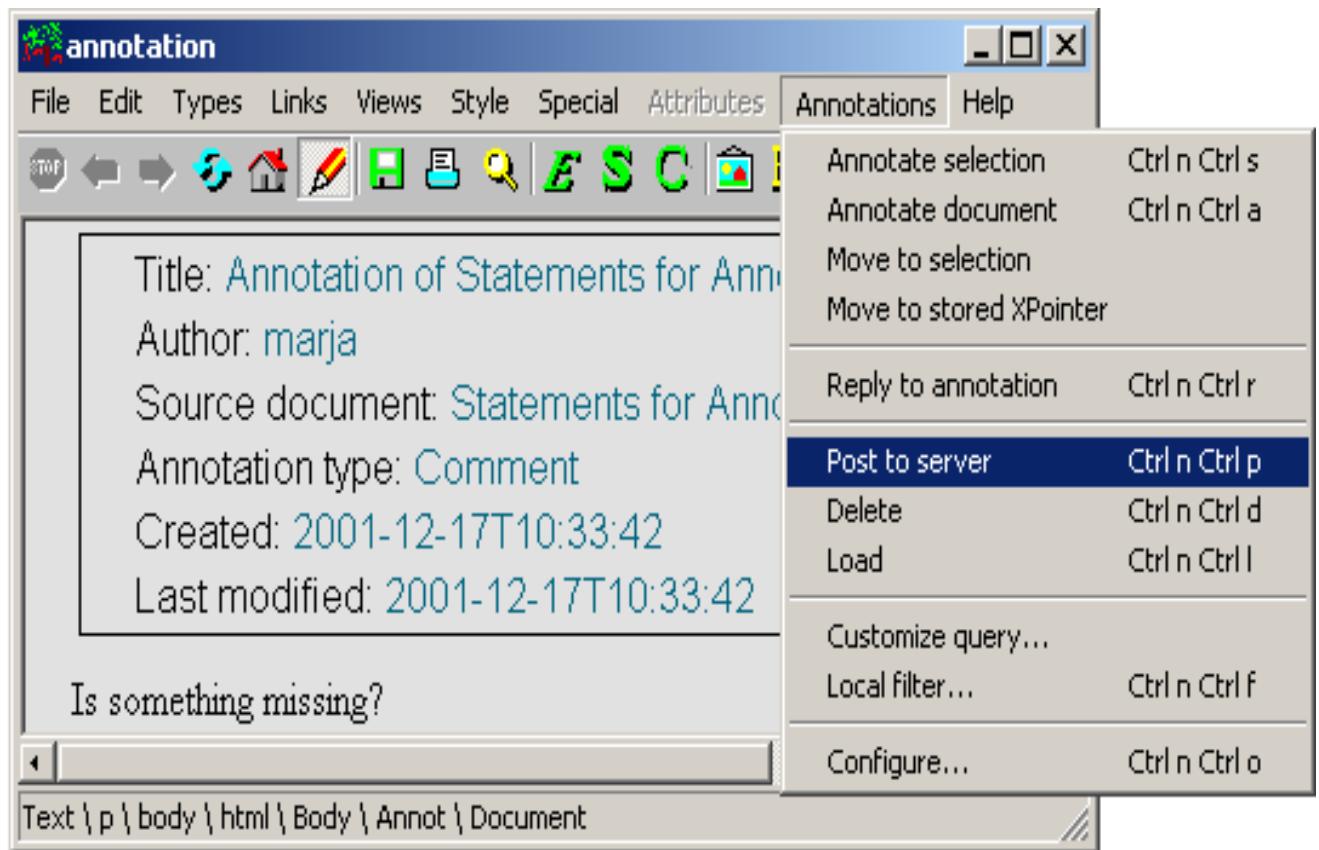


Εικόνα 9

### 3.7 Annotea

Το Annotea είναι ένα πρόγραμμα της LEAD που ενισχύει τη συνεργασία του W3C περιβάλλοντος με τους κοινόχρηστους σχολιασμούς. Με τον όρο σχολιασμοί εννοούμε σχόλια ,

σημειώσεις, επεξηγήσεις, ή άλλου τύπου εξωτερικές παρατηρήσεις οι οποίες μπορούν να επισυνάπτονται σε οποιοδήποτε έγγραφο του Διαδικτύου ή σε ένα επιλεγμένο μέρος του εγγράφου χωρίς στη πραγματικότητα να χρειάζεται να «αγγίζουμε» το έγγραφο. Όταν ο χρήστης παίρνει το έγγραφο μπορεί να φορτώσει τα σχόλια επισυνάπτοντάς τα σε αυτό από ένα επιλεγμένο διακομιστή σχολίων ή από περισσότερους διακομιστές και να δει τη πιστεύουν οι ομότιμοί του. Το Annotea είναι ανοιχτό, χρησιμοποιεί και βοηθάει να προωθηθούν τα πρότυπα του W3C όταν είναι δυνατό. Για παράδειγμα, χρησιμοποιούμε ένα σχήμα σχολιασμών βασισμένο στο RDF για να περιγράψουμε σχολιασμούς μεταδεδομένα και XPointer για να τοποθετήσουμε σχόλια στο σχολιαζόμενο έγγραφο. Το Annotea είναι μέρος των προσπαθειών της Semantic Web. Οι σχολιασμοί αποθηκεύονται σε διακομιστές σχολίων σαν μεταδεδομένα και παρουσιάζονται στο χρήστη από ένα πελάτη ικανό να κατανοήσει αυτά τα μεταδεδομένα και με τη δυνατότητα αλληλεπίδρασης με ένα διακομιστή σχολίων με το πρωτόκολλο υπηρεσιών HTTP. Η πρώτη εφαρμογή πελάτη του Annotea είναι ο επεξεργαστής/φυλλομετρητής Amaya του W3C. Επίσης τίποτα δεν εμποδίζει τους άλλους πελάτες από την υλοποίηση αυτών των δυνατοτήτων.

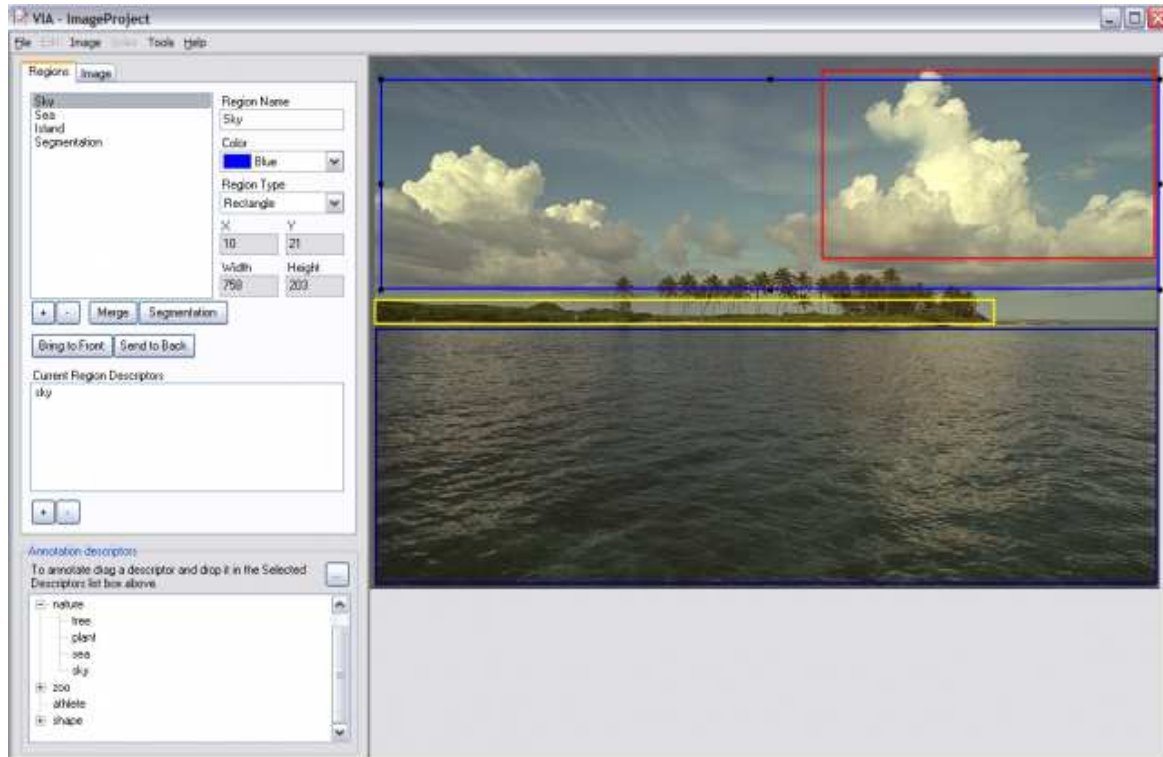


Εικόνα 10

### 3.8 Video Image Annotation Tool (VIA)

Το Video Image Annotation Tool (VIA) είναι μία εφαρμογή των Windows με την οποία ο χρήστης μπορεί χειροκίνητα να σχολιάσει βίντεο και εικόνες. Προσφέρει μία φιλική προς το χρήστη διεπαφή για την ακριβή και απαιτητικό ζωντανό "καρέ-καρέ" σχολιασμό ενός βίντεο αλλά και μίας εικόνας. Η ανάπτυξη αυτού του εργαλείου έχει υποστηριχθεί από την Ontology Evolution with Multimedia Information Extraction (BOEMIE) και την aceMedia.





Εικόνα 11

## Κύρια Χαρακτηριστικά

Επίπεδο σχολιασμού:

1. Εικόνα /πλαίσιο
2. Βίντεο
3. Στιγμιότυπο (σκηνή)
4. Περιοχή (οριοθέτηση πλαισίου: θέση/μέγεθος ανά πλαίσιο ή εικόνα)

Μέθοδοι;

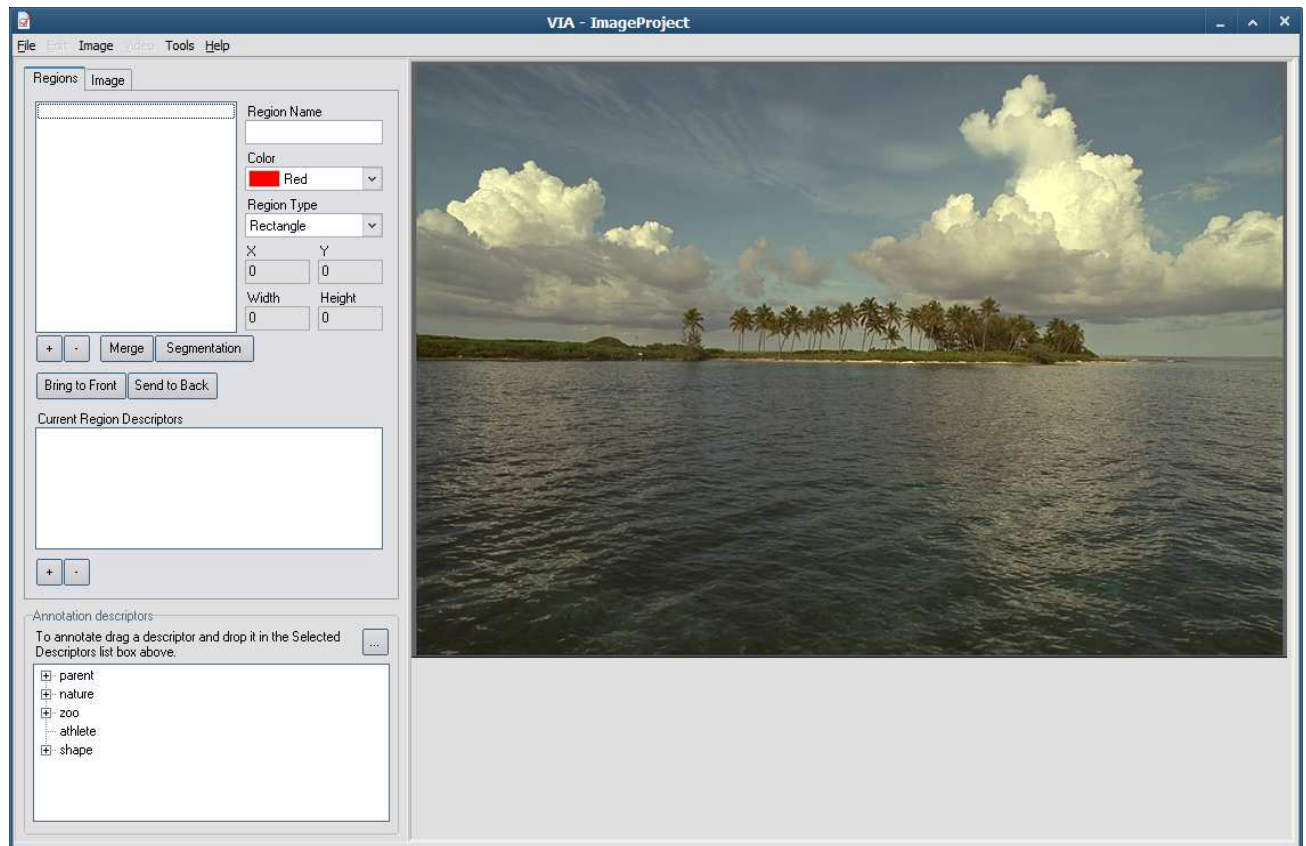
1. Ζωντανός σχολιασμός βίντεο (κατά την αναπαραγωγή)
2. Σχολιασμός καρτέ-καρτέ
3. Σχολιασμός εικόνας

Τύποι σχολίων:

1. Σχολιασμός ελεύθερου κειμένου

Διάφορα:

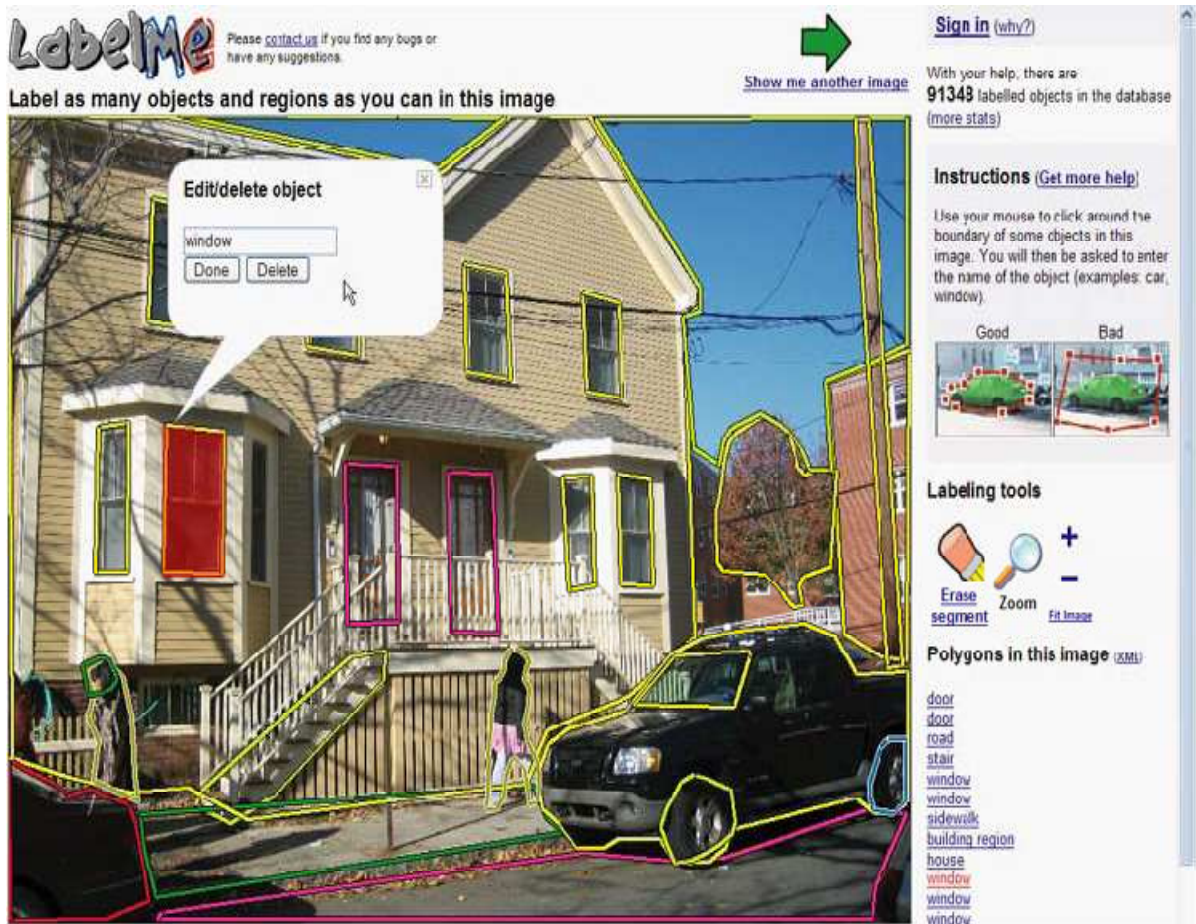
1. Υποστηρίζει MPEG 1,2 για βίντεο (ανεξάρτητα από την κωδικοποίηση του συστήματος)
2. Ακριβής αναπαραγωγή πλαισίου βίντεο και πλοήγηση
3. Φόρτωση αποθήκευση έργου διαφόρων μορφών (κείμενο, XML)



Εικόνα 12

### 3.9 LabelMe

Το LabelMe είναι ένα διαδικτυακό εργαλείο σχολιασμού εικόνων το οποίο επιτρέπει στους ερευνητές να επισημαίνουν εικόνες και να μοιράζονται τους σχολιασμούς τους με την υπόλοιπη κοινότητα. Επί του παρόντος οι υπολογιστές δυσκολεύονται να αναγνωρίσουν αντικείμενα μέσα σε εικόνες. Αν και υπάρχουν πρακτικές λύσεις για μερικές απλές κλάσεις, όπως ανθρώπινα πρόσωπα ή αυτοκίνητα το γενικότερο πρόβλημα της αναγνώρισης όλων των διαφορετικών κλάσεων αντικειμένων στο κόσμο (π.χ. κιθάρες, μπουκάλια, τηλέφωνα) παραμένει άλυτο. Οι ερευνητές τεχνητής όρασης πλέον ερευνούν μεθόδους με τις οποίες θα αναγνωρίζονται και θα εντοπίζονται χιλιάδες διαφορετικές κατηγορίες αντικειμένων σε σύνθετες σκηνές. Ένα βασικό στοιχείο αυτών των αλγορίθμων είναι το στοιχείο που χρησιμοποιείται για να εκπαιδεύσει το πρότυπο του υπολογιστή για κάθε αντικείμενο. Ο στόχος του LabelMe είναι να παρέχει ένα εργαλείο σχολιασμού σε απευθείας σύνδεση για να δημιουργηθεί μία μεγάλη βάση δεδομένων σχολιασμένων εικόνων με τη συλλογή συνεισφορών από πολλούς ανθρώπους.



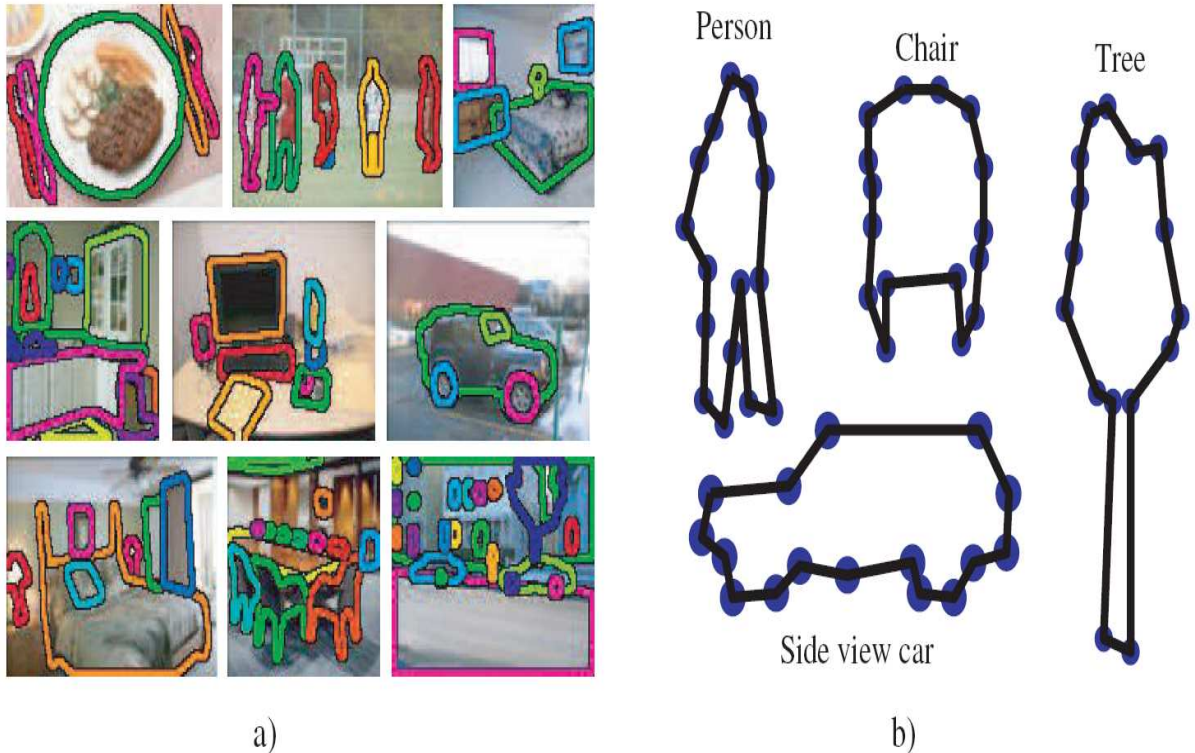
Εικόνα 13

## Χαρακτηριστικά του LabelMe

1. Έχει σχεδιαστεί για την αναγνώριση κλάσεων αντικειμένων σε αντίθεση με τη αναγνώριση περιπτώσεων. Για την αναγνώριση μίας κλάσης αντικειμένου από τη μία χρειάζεται πολλαπλές εικόνες διαφορετικών περιπτώσεων της ίδιας κλάσης και από την άλλη διαφορετικές συνθήκες παρατήρησης. Παρόλα αυτά πολλές βάσεις δεδομένων περιέχουν μόνο διαφορετικές περιπτώσεις σε κανονική στάση.
2. Είναι σχεδιασμένο για την εκμάθηση των αντικειμένων τα οποία είναι ενσωματωμένα σε μία σκηνή. Πολλές βάσεις δεδομένων αποτελούνται από μικρές κομμένες εικόνες περιπτώσεων αντικειμένων. Αυτές είναι κατάλληλες για την εκπαίδευση ανιχνευτών αντικειμένων βασισμένους στα patches (συμπληρωματικός κώδικας τροποποίησης προγράμματος), αλλά δεν μπορεί να χρησιμοποιηθεί για την εκπαίδευση ανιχνευτών που αξιοποιούν τα συμφραζόμενα παραδείγματα.
3. Υψηλής ποιότητας επισήμανση. Πολλές βάσεις δεδομένων απλά παρέχουν λεζάντες, οι οποίες προσδιορίζουν ότι το αντικείμενο βρίσκεται κάπου μέσα στην εικόνα. Ωστόσο, πιο λεπτομερείς πληροφορίες όπως οριοθετημένα πλαίσια, πολύγωνα ή μάσκες κατάτμησης είναι τρομερά χρήσιμες.
4. Πολλές διαφορετικές εικόνες. Σε πολλές εφαρμογές είναι χρήσιμο να ποικίλουν οι τύποι σκηνών (π.χ. φύση, δρόμος και σκηνές γραφείου), αποστάσεις (π.χ. τοπίου και κοντινές λήψεις) βαθμός ακαταστασίας κτλ.



5. Πολλές εικόνες χωρίς πνευματικά δικαιώματα. Για τη βάση δεδομένων του LabelMe η πλειοψηφία των εικόνων έχουν ληφθεί από μία ποικιλία ψηφιακών φωτογραφικών μηχανών χειρός. Επίσης υπάρχουν πολλές ακολουθίες βίντεο τα οποία έχουν ληφθεί μέσω διαδικτυακών καμερών.
6. Ανοιχτή και δυναμική. Η βάση δεδομένων του LabelMe είναι σχεδιασμένη ώστε να επιτρέπει στις συλλογές ετικετών να μοιραστούν άμεσα μέσω του διαδικτύου και να μεγαλώνουν με τη πάροδο του χρόνου.



Εικόνα 14

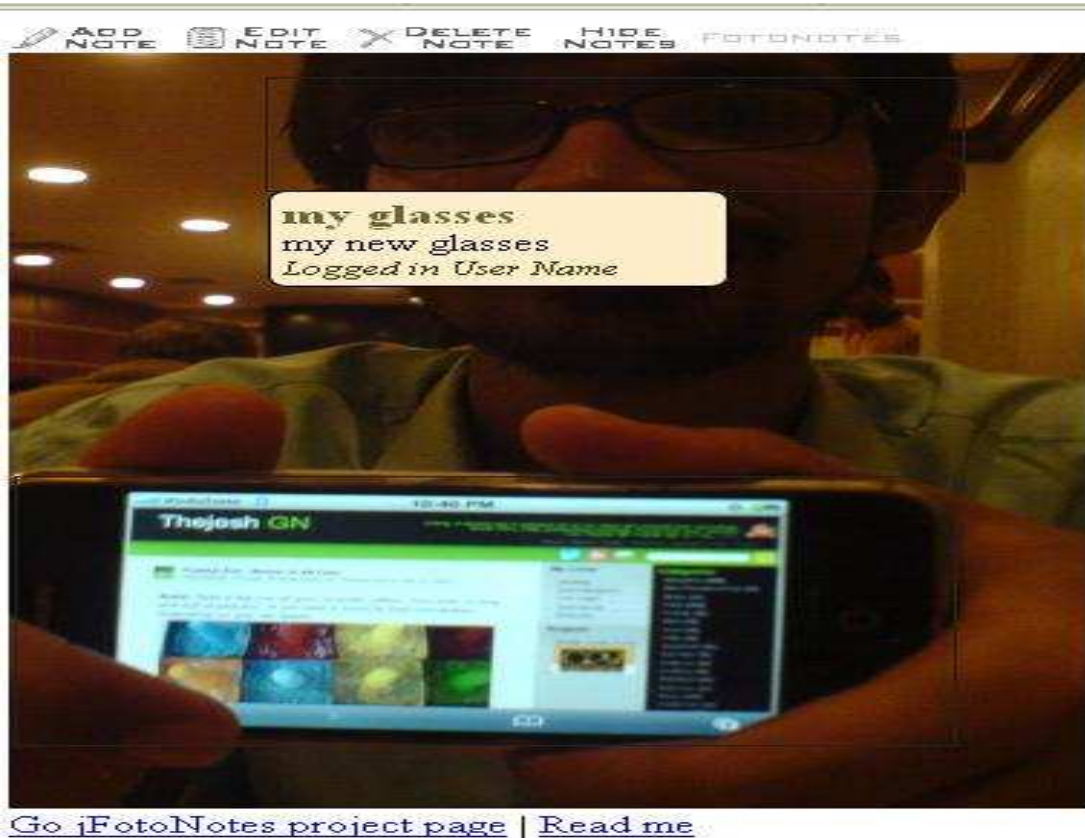
### 3.10 jFotoNotes

Το JFotoNotes είναι η εφαρμογή του FotoNotes σε Java. Το jFotoNotes λαμβάνει τη κάθε εικόνα ως μία συλλογή αντικειμένων παρέχοντας στο καθένα διαφορετικό νόημα και ερμηνεία. Αντί να εστιάζει απευθείας στα χαρακτηριστικά όπου μία φωτογραφία έχει, το jFotoNotes διαιρεί την εικόνα σε διάφορα ορατά αντικείμενα και στη συνέχεια αντιστοιχίζει τα σχόλια σε κάθε αντικείμενο. Κάθε διαφορετικό αντικείμενο μπορεί να περιλαμβάνει ανθρώπους, ιστορίες ,δεδομένα , τον συντάκτη, μη ορατά μεταδεδομένα, ή άλλες πληροφορίες που περιλαμβάνονται για την εικόνα. Ως σύνολο αυτό οι σχολιασμοί ενεργούν για να πουν μία ιστορία που περιγράφει την εικόνα. Επιπλέον τα ξεχωριστά αντικείμενα που επινοούνται από την μεγαλύτερη φωτογραφία βοηθούν στην επέκταση

του αρχείου όλων των αντικειμένων. Το JFotoNotes επίσης επιτρέπει στο χρήστη να ψάξει με βάση συγκεκριμένα αντικείμενα ή σχολιασμούς αντί ενός ευρέος φάσματος χαρακτηριστικών τα οποία μπορούν ή δεν μπορούν να υποστηριχθούν από ένα συγκεκριμένο αντικείμενο. Η αποσύνθεση των αντικειμένων από τη μεγαλύτερη πηγή ενισχύει την αποδοτικότητα των σημασιολογικών αναζητήσεων.

## Χαρακτηριστικά

1. Είναι κωδικοποιημένο σε mysql αλλά ανά πάσα στιγμή ο χρήστης μπορεί να το αλλάξει. Αρκεί να αλλάξει το `jFotoNotes.data.FotoNotesDataImpl`. Όπως και να έχει όμως ακόμα και αν γίνει χρήση της mysql θα πρέπει να αλλάξει η εφαρμογή `getConnection`.
2. Για την mysql θα χρειάζεται ο οδηγός προγράμματος (driver) για jdbc.
3. Συμπεριλαμβάνετε μία δοκιμαστική διάταξη με κάποια δοκιμαστικά δεδομένα εισάγοντας το `fotonotes.sql` στη βάση δεδομένων του FotoNotes.



Εικόνα 15

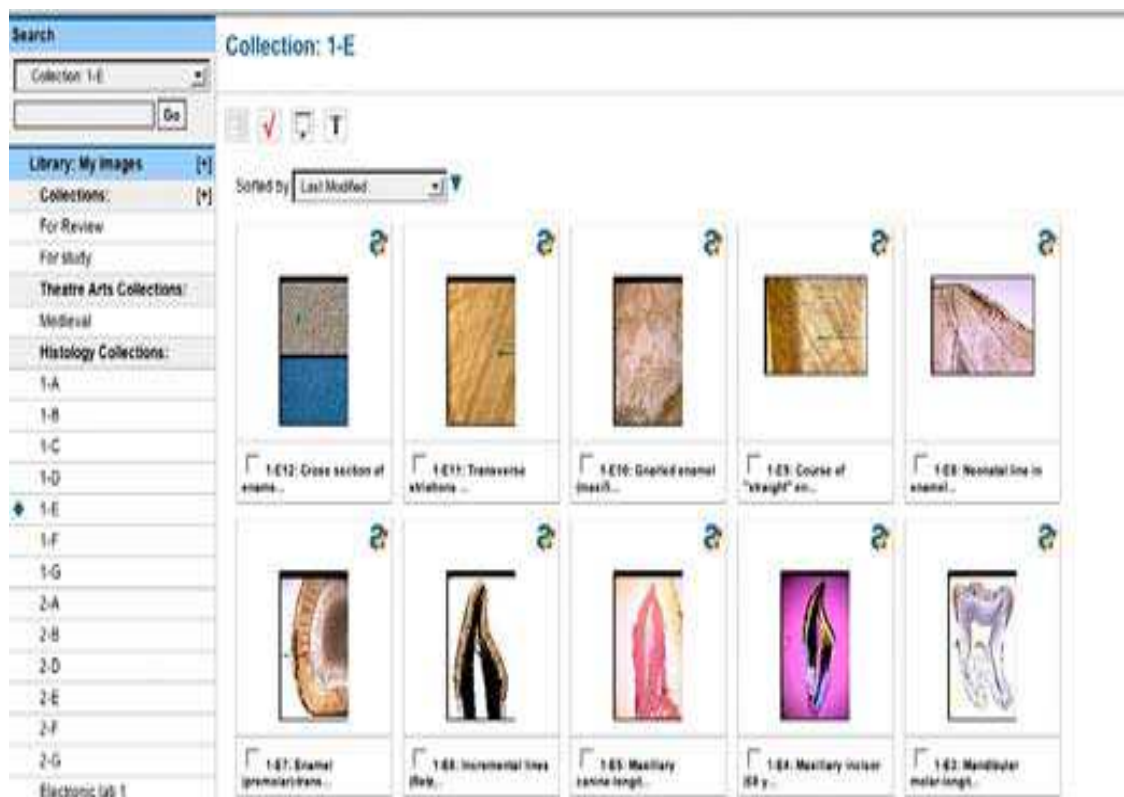
### 3.11 IAT

Στις μέρες μας η ανάπτυξη των ψηφιακών τεχνολογιών και η χρήση του διαδικτύου αποτελούν σημαντικά εργαλεία της επιστημονικής κοινότητας μέσω των οποίων οι επιστήμονες μπορούν να έχουν πρόσβαση σε τεράστιες βάσεις δεδομένων. Σε ότι αφορά την εκπαίδευση οι επιστήμονες πλέον μπορούν εύκολα να αναζητήσουν πληροφορίες. Ένα Εργαλείο Σχολιασμού Εικόνων το οποίο όμως επικεντρώνεται στον Εκπαιδευτικό τομέα είναι το IAT του Κέντρου Εκπαίδευσης και Μάθησης Νέων Μέσων του Κέντρου Κολούμπια.

Το IAT καθιστά εύκολη την φόρτωση το σχολιασμό και την κοινοποίηση εικόνων on-line παρέχοντας έναν εύκολο τρόπο για τις σχολές και τους φοιτητές να μελετήσουν και να επανεξετάσουν εικόνες. Το IAT αναπτύχθηκε για μια σειρά μαθημάτων της Σχολής Οδοντιατρικής και Στοματικής Χειρουργικής αλλά μπορεί να είναι ένα πολύτιμο εργαλείο για οποιαδήποτε επιστήμη που απαιτεί τη στενή μελέτη των εικόνων.

Με αυτό το Εργαλείο Σχολιασμού Εικόνων είναι δυνατή η φόρτωση εικόνων σε προσωπικές βιβλιοθήκες των σχολών αλλά και των φοιτητών. Οι εικόνες μπορούν να οργανωθούν σε συλλογές βασισμένες σε διάφορα θέματα ή σκοπούς. Επιπλέον, ένας εκπαιδευτικός μπορεί να δημιουργήσει τις κοινές βιβλιοθήκες για να καταστήσει τις συγκεκριμένες συλλογές των εικόνων προσιτές σε όλα τα μέλη της τάξης. Οι εικόνες μπορούν να καθοριστούν ως ελεύθερες για συγκεκριμένο χρονικό διάστημα έτσι ώστε ο εκπαιδευτικός να μπορεί να ελέγχει πότε αυτές θα είναι διαθέσιμες στους φοιτητές.

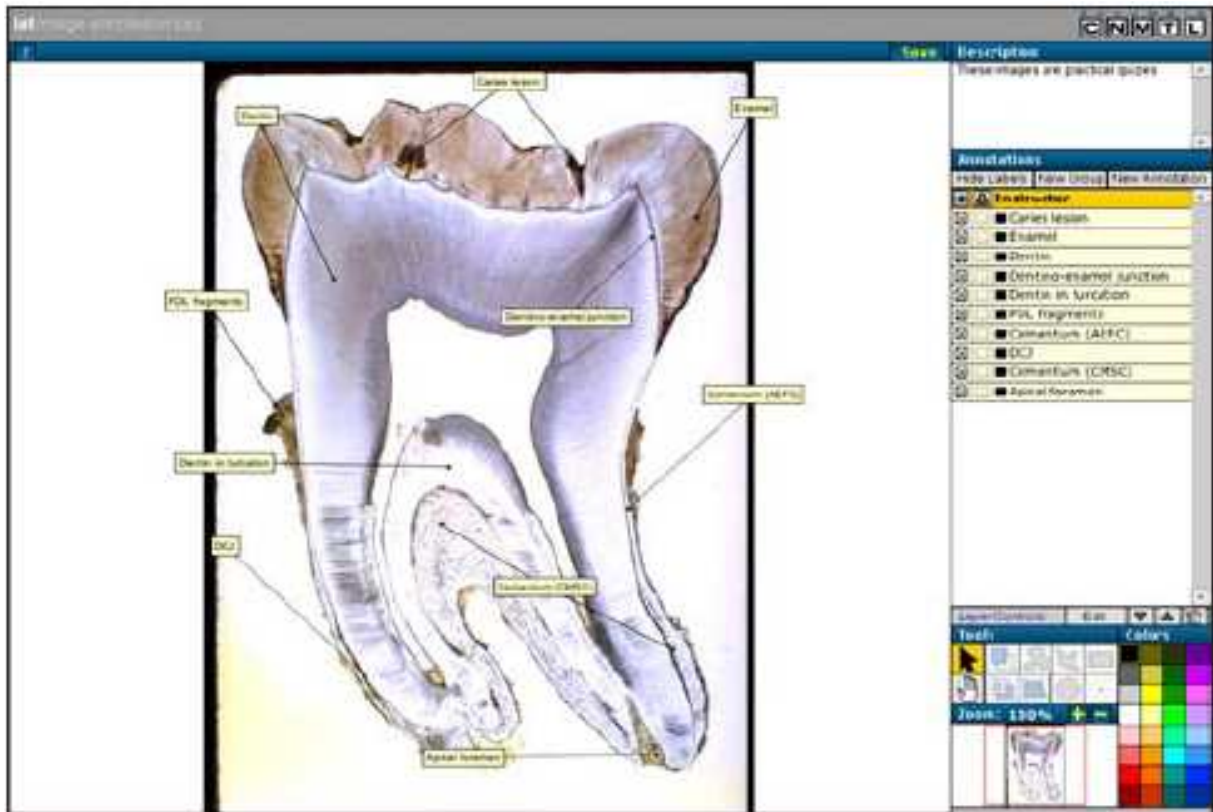
Η καρδιά του IAT είναι το ίδιο το εργαλείο σχολιασμών, το οποίο διαθέτει μια διαισθητική διεπαφή για τη δημιουργία και επεξεργασία ετικετών χρωματικής κωδικοποίησης οι οποίες μπορούν να επισυναφθούν σε συγκεκριμένες περιοχές μιας εικόνας. Κάθε ετικέτα είναι προσαρτημένη σε ένα στρώμα το οποίο μπορεί να οργανωθεί σε ομάδες. Οι φοιτητές μπορούν να αλλάξουν ακόμα και να σβήσουν τους σχολιασμούς για να αξιολογήσουν πόσο καλά έχουν μάθει τις σχετικές έννοιες και πληροφορίες, καθώς και οι εκπαιδευτές να εξετάσουν την εργασία των φοιτητών στο IAT προκειμένου να εκτιμηθεί η κατανόηση από τους φοιτητές.



Εικόνα16

## Χαρακτηριστικά του IAT

1. Δυνατότητα ζουμ και κύλισης πάνω στην εικόνα για προβολή λεπτομερειών
2. Αποθήκευση σημειώσεων και περιγραφών για κάθε εικόνα
3. Οι σχολιασμοί μπορούν να συνδεθούν με εξωτερικές ηλεκτρονικές πηγές
4. Δυνατότητα στον εκπαιδευτή να εξετάσει τους σχολιασμούς των σπουδαστών
5. Προβολή συλλογών εικόνων σε μορφή slide show.



Εικόνα17

### 3.12 Παρακάτω ακόλουθη ένας συνοπτικός πίνακας με τα annotation tools.



ΟΝΟΜΑ	ΔΗΜΙΟΥΡΓΟ	ΜΟΡΦΗ	ΙΣΤΟΣΕΛΙΔΑ	ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ
<b>Image annotatio n Tool</b>	Northwestern University	XML	<a href="http://dewey.at.northwestern.edu/ppad2/documents/head2/">http://dewey.at.northwestern.edu/ppad2/documents/head2/</a>	-επιτρέπει μεγέθυνση υψηλής ανάλυσης εικόνας, πρόσθετη sticky notes -το περιεχόμενο είναι σε αυτοέλεγχο
<b>Anvil</b>	Michael Kipp Quan Nguyen	XML	<a href="http://www.anvil-software.de/">http://www.anvil-software.de/</a>	-installer -integrated online manual -specification editon -3D motion capture player -sub division track type -time point track type -elan import
<b>Annotate Expert</b>	Nikonians	XMP	<a href="http://www.nikonians.org/html/resources/software/an">http://www.nikonians.org/html/resources/software/an</a>	ενσωματωμένα σχόλια - σχόλια ήχου -ανυχνευσιτότητα - έξυπνη αντιγραφή - ipct ολοκλήρωση λέξεων-κλειδιά -επίπεδα σχολίων
<b>Foto Tagger</b>	Cogitum LC	σε HTML	<a href="http://www.cogitum.com/">http://www.cogitum.com/</a>	διατηρεί λεπτομέρειες στην εικόνα -εύρεση απαιτούμενων εικόνων άμεσα -κοινή χρήση σχολιασμένων εικόνων
<b>M- OntoMat - Annotize</b>	European UnionSixth Framework Programme	XML	<a href="http://www.acedia.org/aceMedia/results/software/mat-ontomat-">http://www.acedia.org/aceMedia/results/software/mat-ontomat-</a>	-core onto mat annotate -annotation server -domain visual data base - featute extraction toolbox -vde visual editor &media viewer -vde plug-in
<b>Annotea</b>	aceMedia	XMP	<a href="http://www.w3.org/2001/Annotea/">http://www.w3.org/2001/Annotea/</a>	-abstract hosting -βελτιστοποίηση βάσης δεδομένων -bookmark server - σχολιασμένο parser σχήμα
<b>VIA</b>	MIT	XML	<a href="http://via-tool.sourceforge.net/">http://via-tool.sourceforge.net/</a>	επίπεδο σχολιασμού - μέθοδοι - τύποι σχολίων
<b>LabelMe</b>	MIT	XML	<a href="http://labelme.csail.mit.edu/">http://labelme.csail.mit.edu/</a>	-αναγνώριση κλάσεων αντικειμένων -σχεδιασμένο για την εκμάθηση των αντικειμένων
<b>FotoNotes</b>	Greg Elin	XML	<a href="http://fotonotes.net/">http://fotonotes.net/</a>	-κωδικοποιημένο mysql - για την mysql χρειάζεται driver για jdbc -εισάγοντας footnotes.sql στην βάση δεδομένων
<b>IAT</b>	Columbia Center	XML	<a href="http://ccnmtl.columbia.edu/projects/iat/">http://ccnmtl.columbia.edu/projects/iat/</a>	<ol style="list-style-type: none"> <li>1. Δυνατότητα ζουμ</li> <li>2. Αποθήκευση σημειώσεων και περιγραφών για κάθε εικόνα</li> <li>3. Οι σχολιασμοί μπορούν να συνδεθούν με εξωτερικές ηλεκτρονικές πηγές</li> <li>4. Δυνατότητα στον εκπαιδευτή να εξετάσει τους σχολιασμούς των σπουδαστών</li> <li>5. Προβολή συλλογών εικόνων σε μορφή slide show.</li> </ol>



## ΚΕΦΑΛΑΙΟ 4

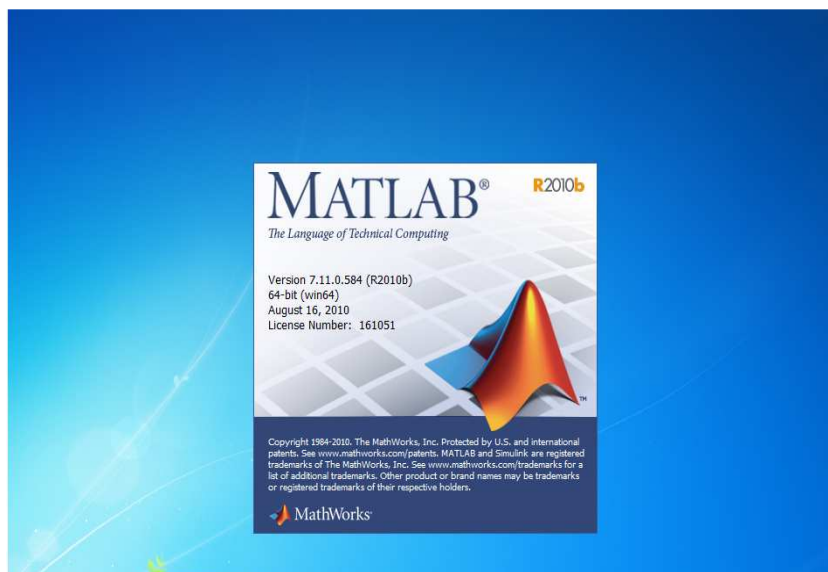
### 4.1 Εισαγωγή

Σε αυτό το κεφάλαιο περιγράφεται το «πείραμα» που πραγματοποιήσαμε για να αξιολογήσουμε την αποτελεσματικότητα του εργαλείου σχολιασμού LABEL ME χρησιμοποιώντας το πρόγραμμα MATLAB. Η επιλογή του LABEL ME έγινε πρώτον γιατί είναι ένα πρόγραμμα με ανοιχτό κώδικα ο οποίος μας επιτρέπει να κάνουμε αλλαγές, αλλά και γιατί μας δίνει το ίδιο τη δυνατότητα να προσθέσουμε την εργαλειοθήκη του στο MATLAB. Εν κατακλείδι είναι ένα κατανοητό, εύχρηστο και παραμετροποιήσιμο Εργαλείο Σχολιασμού.

Όπως θα δούμε και παρακάτω το LABEL ME παρέχει δύο τρόπους εργασίας με το σύνολο των δεδομένων:

1. Σε απευθείας σύνδεση με τη βάση δεδομένων του LABEL ME χωρίς να χρειάζεται να κατεβάσουμε όλη τη βάση δεδομένων στον υπολογιστή μας. Μπορεί με αυτό το τρόπο να κερδίζουμε χώρο στον υπολογιστή μας αλλά σε περίπτωση συχνής χρήσης του LABEL ME εύκολα διαπιστώνουμε ότι δεν μας ωφελεί καθώς είναι πιο αργός.
2. Αποθήκευση της βάσης δεδομένων στον υπολογιστή μας. Με τον τρόπο αυτό έχουμε γρηγορότερη πρόσβαση στη βάση δεδομένων μας αλλά χρειάζεται πολύς χώρος και χρόνος μέχρι να κατεβάζουμε την βάση αλλά και συχνές ανανεώσεις της βάσης μας προκειμένου να είναι συνεχώς ενημερωμένη.

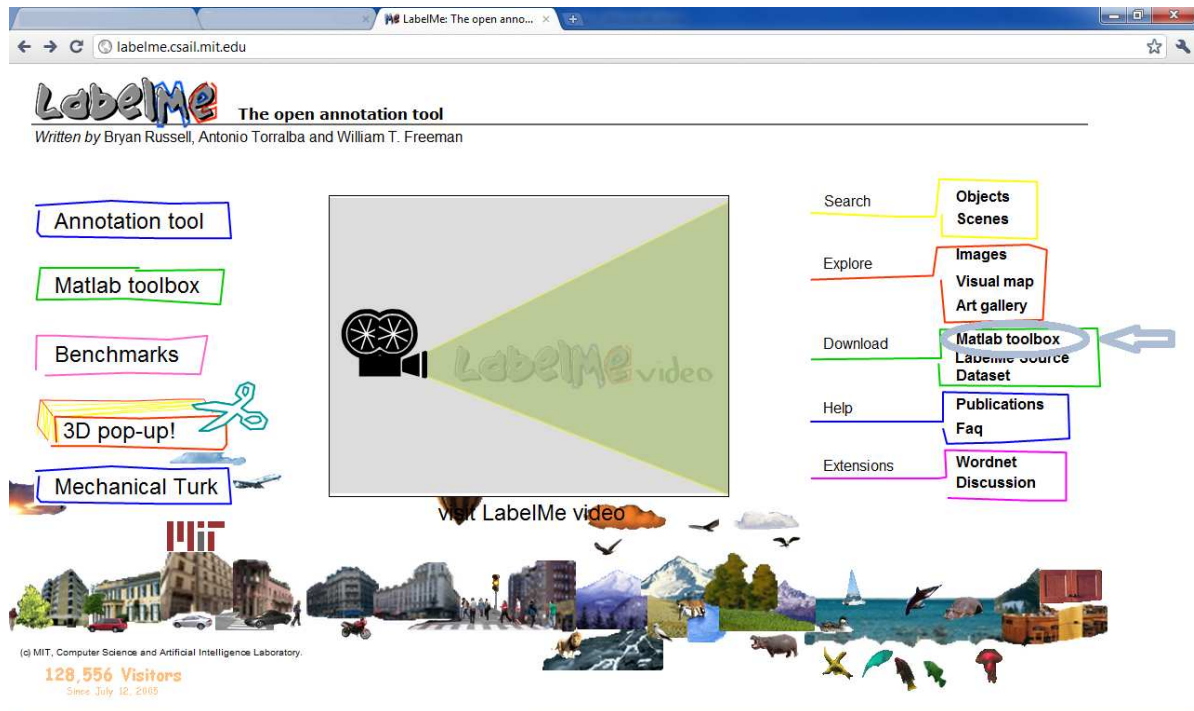
Η βάση δεδομένων του LABELME μας προσφέρει μία πληθώρα εικόνων από διάφορα άλμπουμ, μας παρέχει όμως και τη δυνατότητα να φορτώσουμε και δικές μας εικόνες. Σε αυτό το κεφάλαιο βλέπουμε πως οπτικοποιούμε τις επισημάνσεις για μια εικόνα, να αναζητάμε εικόνες από το άλμπουμ μας θέτοντας διάφορα ερωτήματα, να εξάγουμε πολύγωνα και τμήματα, να αντικαταστήσουμε μερικά ονόματα αντικειμένων κ.α. Τέλος βλέπουμε πώς να σχολιάζουμε τις δικές μας εικόνες.



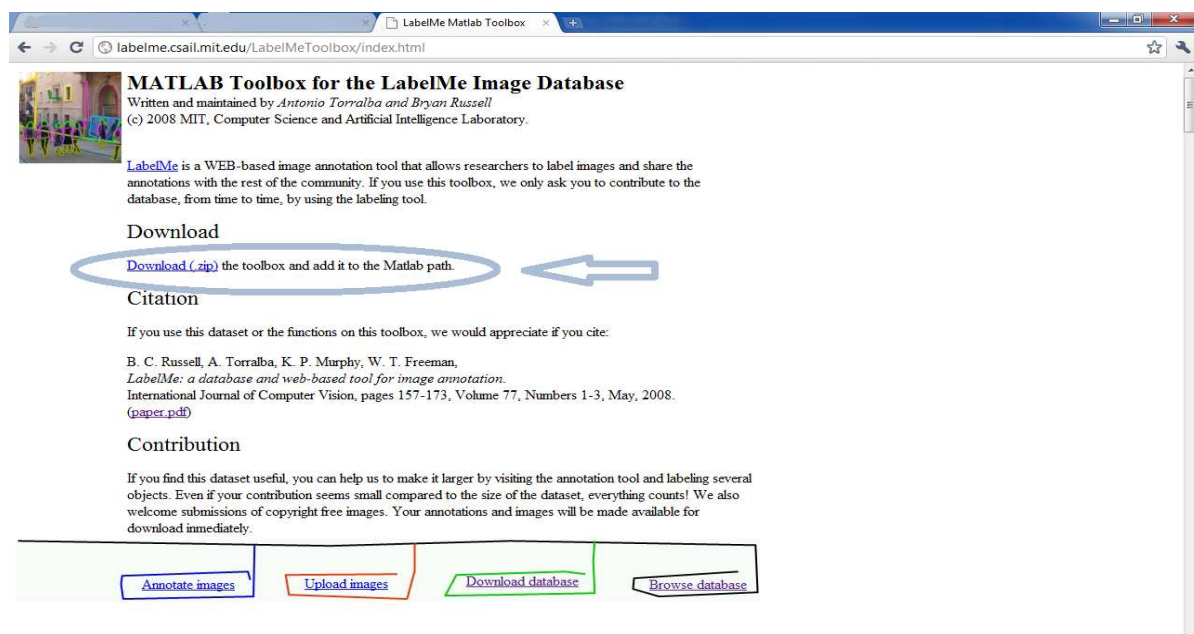
## 4.2 Εγκατάσταση Εργαλειοθήκη

Η εγκατάσταση της εργαλειοθήκης του LabelMe στο MATLAB είναι πολύ εύκολη παρακάτω παρουσιάζονται τα βήματα που απαιτούνται.

Βήμα 1. Από την αρχική σελίδα του LabelMe (<http://labelme.csail.mit.edu/>) πηγαίνουμε στο download=> MATLAB TOOLBOX , εκεί υπάρχει ένα zip αρχείο το οποίο και κατεβάζουμε.



Εικόνα 18



Εικόνα 19

Βήμα 2. Αυτό το zip αρχείο το αποσυμπιέζουμε και το αποθηκεύουμε στον φάκελο C:\Program Files\MATLAB\R2010b\toolbox

Βήμα 3. Κάθε φορά που ξεκινάμε την εφαρμογή του MATLAB θα εκτελούμε την παρακάτω εντολή, ώστε να μπορούμε να χρησιμοποιήσουμε τις συναρτήσεις του της εργαλειοθήκης του LabelMe. `Addpath(genpath('/desired/path/to/ LabelMeToolbox'))`;

Στην περίπτωση μας η εντολή είναι η εξής:

```
Addpath(genpath('C:\ProgramFiles\MATLAB\R2010b\toolbox\LabelMeToolbox'));
```

## 4.3 Περιγραφή εργαλειοθήκης

### Μια γρήγορη ματιά στα σύνολα δεδομένων

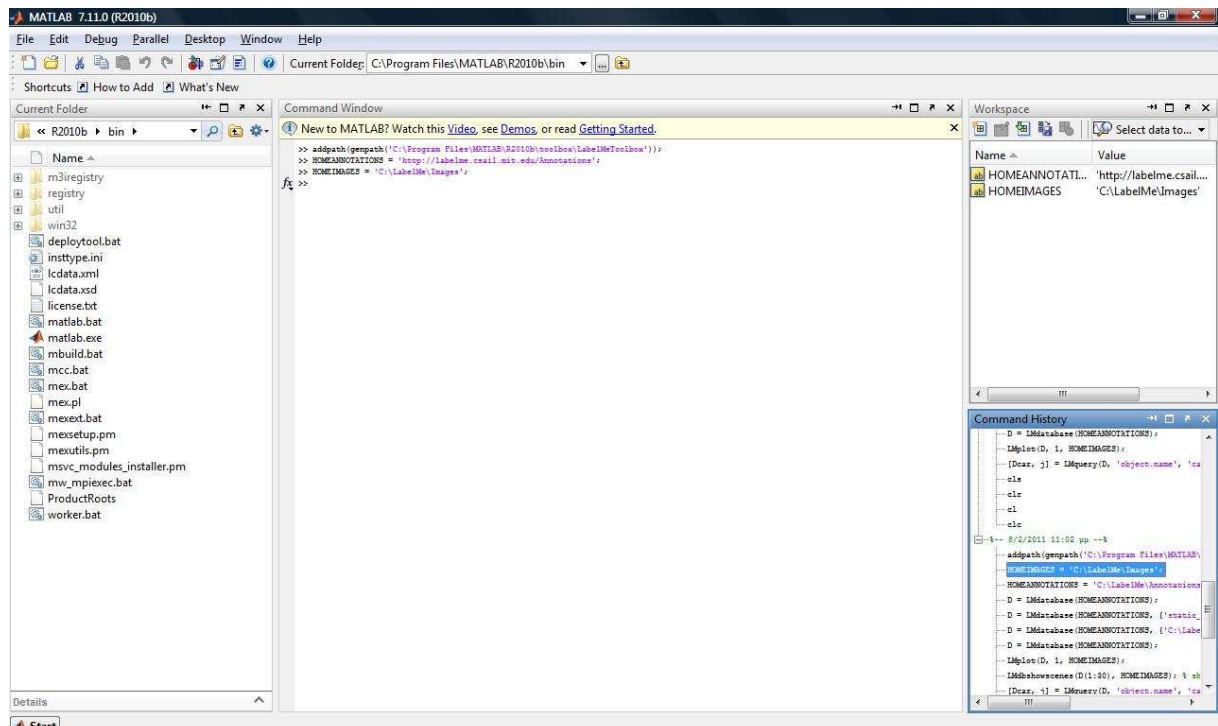
Η εργαλειοθήκη επιτρέπει να χρησιμοποιήσουμε ένα σύνολο δεδομένων διαδικτυακά χωρίς να χρειάζεται να το μεταφορτώσουμε. Εκτελούμε τις παρακάτω εντολές για να απεικονίσουμε το περιεχόμενο ενός φακέλου της συλλογής:

```
HOMEANNOTATIONS = 'http://labelme.csail.mit.edu/Annotations';
```

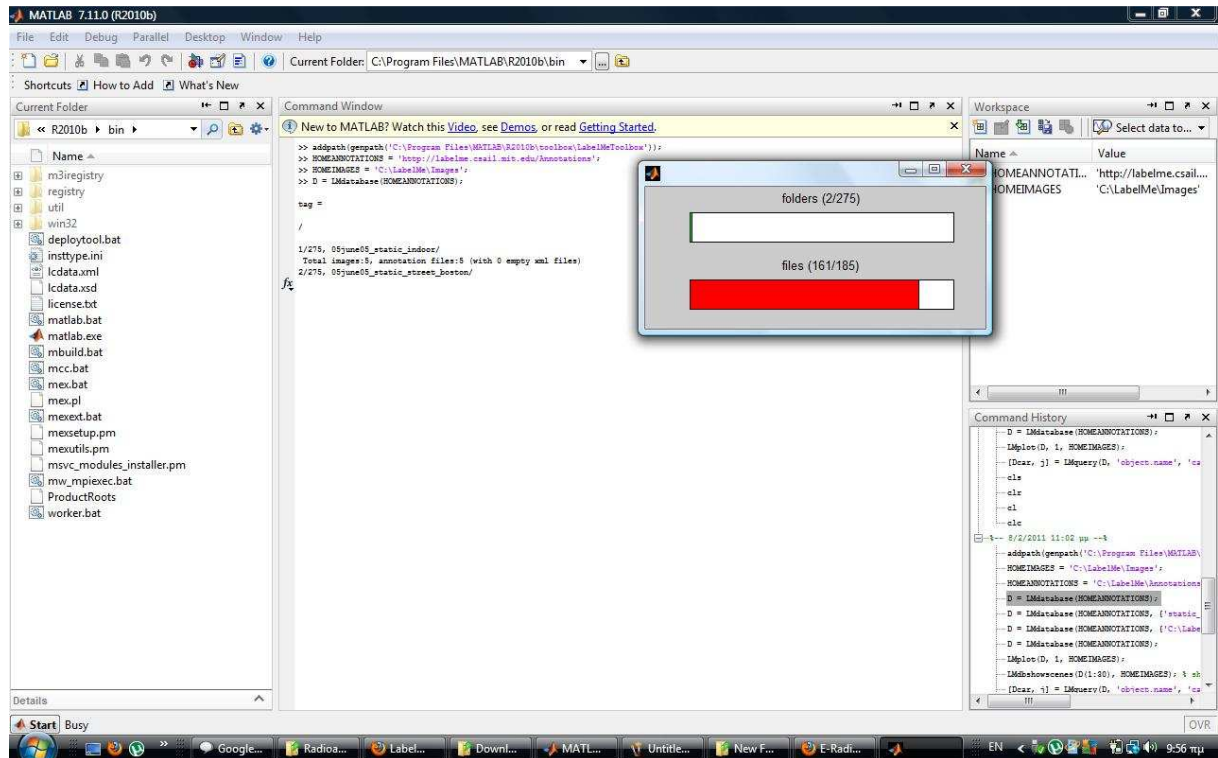
```
HOMEIMAGES = 'http://labelme.csail.mit.edu/Images';
```

```
D = LMdatabase(HOMEANNOTATIONS, {'static_street_statacenter_cambridge_outdoor_2005'});
```

```
LMdbshowscenes(D, HOMEIMAGES);
```



Εικόνα20



Εικόνα21

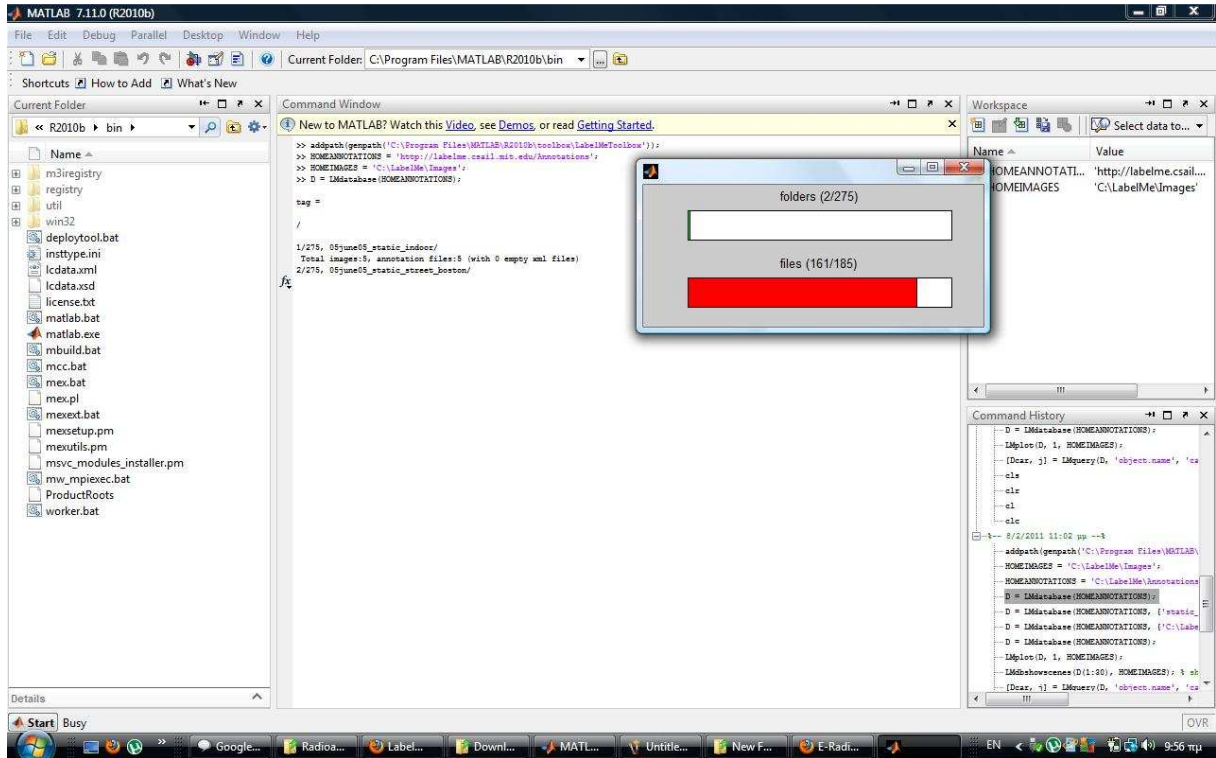
Το παράδειγμα διαβάζει τις εικόνες διαδικτυακά. Εγκαθιστώντας ένα τοπικό αντίγραφο στην βάση δεδομένων μπορούμε να έχουμε ταχύτερη πρόσβαση στις εικόνες και στις επισημάνσεις μειώνοντας το φόρτο εργασίας στο server μας.

## Μεταφορτώνοντας τη Βάση Δεδομένων του LabelMe

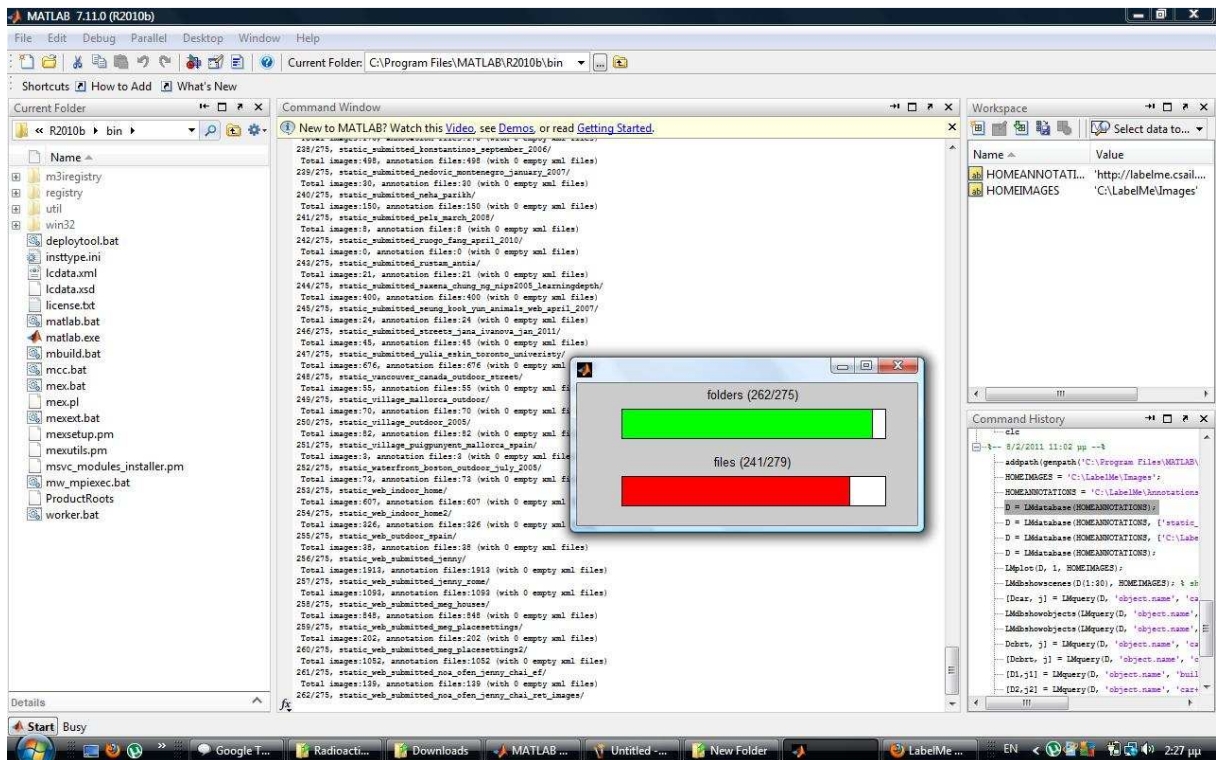
Για να μεταφορτώσουμε εικόνες και επισημάνσεις μπορούμε να χρησιμοποιήσουμε την συνάρτηση Lminstall :

```
HOMEIMAGES = '/desired/path/to/Images';
HOMEANNOTATIONS = '/desired/path/to/Annotations';
Lminstall (HOMEIMAGES, HOMEANNOTATIONS);
```





Εικόνα 22



Εικόνα 23

Ορίζουμε τις μεταβλητές HOMEIMAGES και HOMEANNOTATIONS για να δείξουμε στο τοπικό μας μονοπάτι .Η μεταφόρτωση ολόκληρης της βάσης δεδομένων του LabelMe μπορεί να είναι αργή.

## 4.4 Διαβάζοντας το ευρετήριο.

Τα σχολιασμένα αρχεία είναι τύπου XML. Η συνάρτηση LMdatabase.m διαβάζει το XML αρχείο και δημιουργεί ένα πίνακα δομών του Matlab ο οποίος θα χρησιμοποιηθεί για να εκτελέσει επερωτήσεις και να εξάγει καταμήψεις από τις εικόνες .Για τη δημιουργία ευρετηρίου για ολόκληρη τη βάση δεδομένων εκτελούμε :

```
D = LMdatabase(HOMEANNOTATIONS);
```

Όπου D είναι ένας πίνακας με τόσες καταχωρίσεις όσες είναι οι σχολιασμένες εικόνες. Για μία εικόνα n μερικά από τα πεδία είναι:

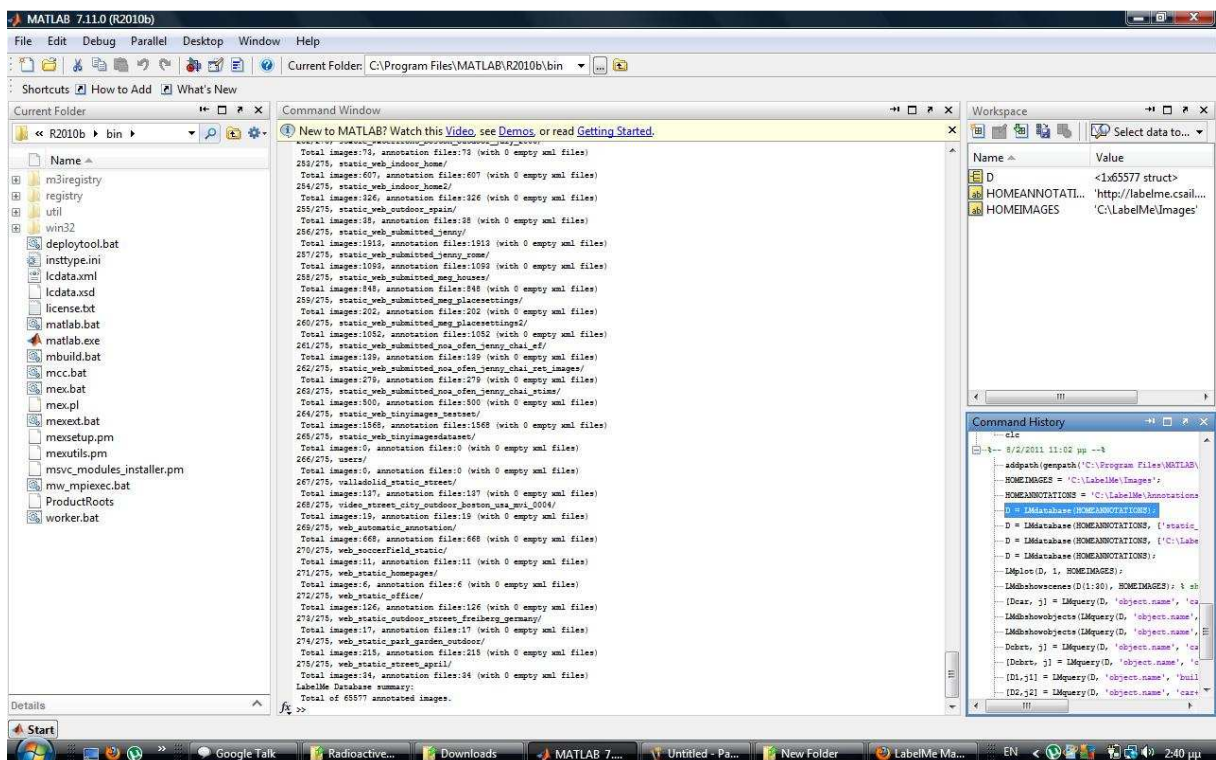
D(n).annotation.folder

D(n).annotation.filename

D(n).annotation.object(m).name

D(n).annotation.object(m).polygon

Όπου n και m είναι η εικόνα και οι δείκτες αντικειμένου αντίστοιχα.

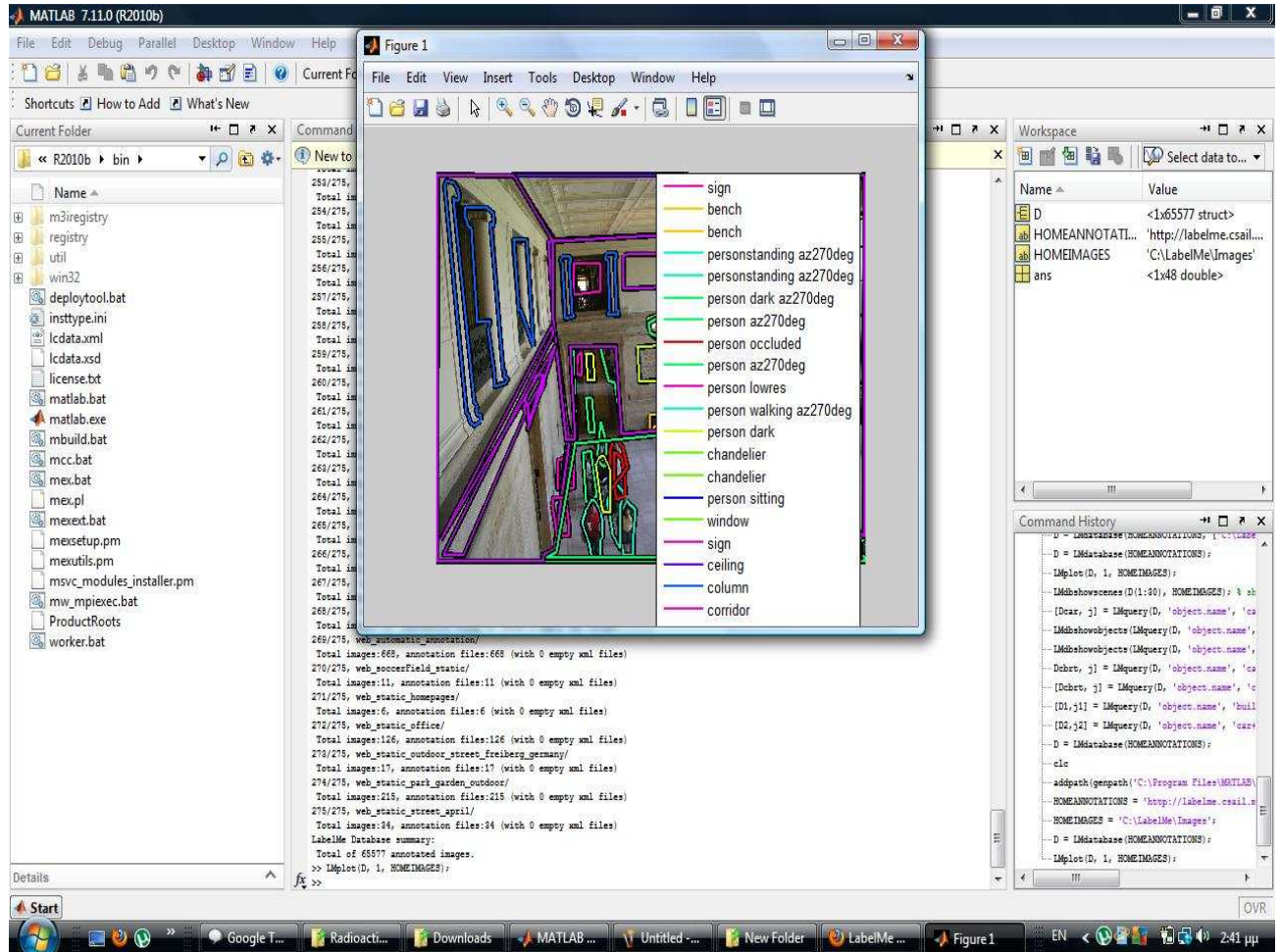


Εικόνα 24

## 4.5 Οπτικοποίηση

Μόλις δημιουργήσουμε ένα LabelMe ευρετήριο D, μπορούμε να οπτικοποιήσουμε τις επισημάνσεις για μία εικόνα με τη συνάρτηση Lmplot :

```
Lmplot(D, 1, HOMEIMAGES);
```



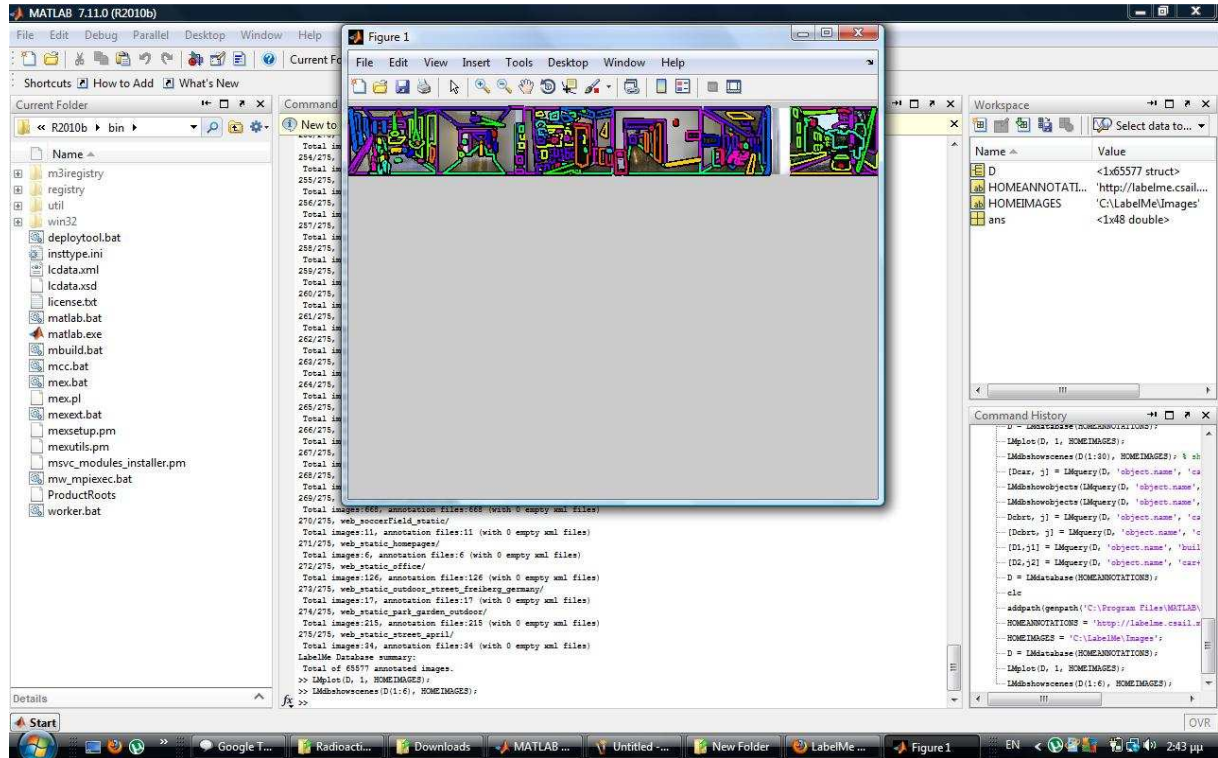
Εικόνα 25

Μπορούμε επίσης να οπτικοποιήσουμε ένα σετ εικόνων ή επιλεγμένων αντικειμένων :

```
Lmdbshowscenes(D(1:6), HOMEIMAGES);
```

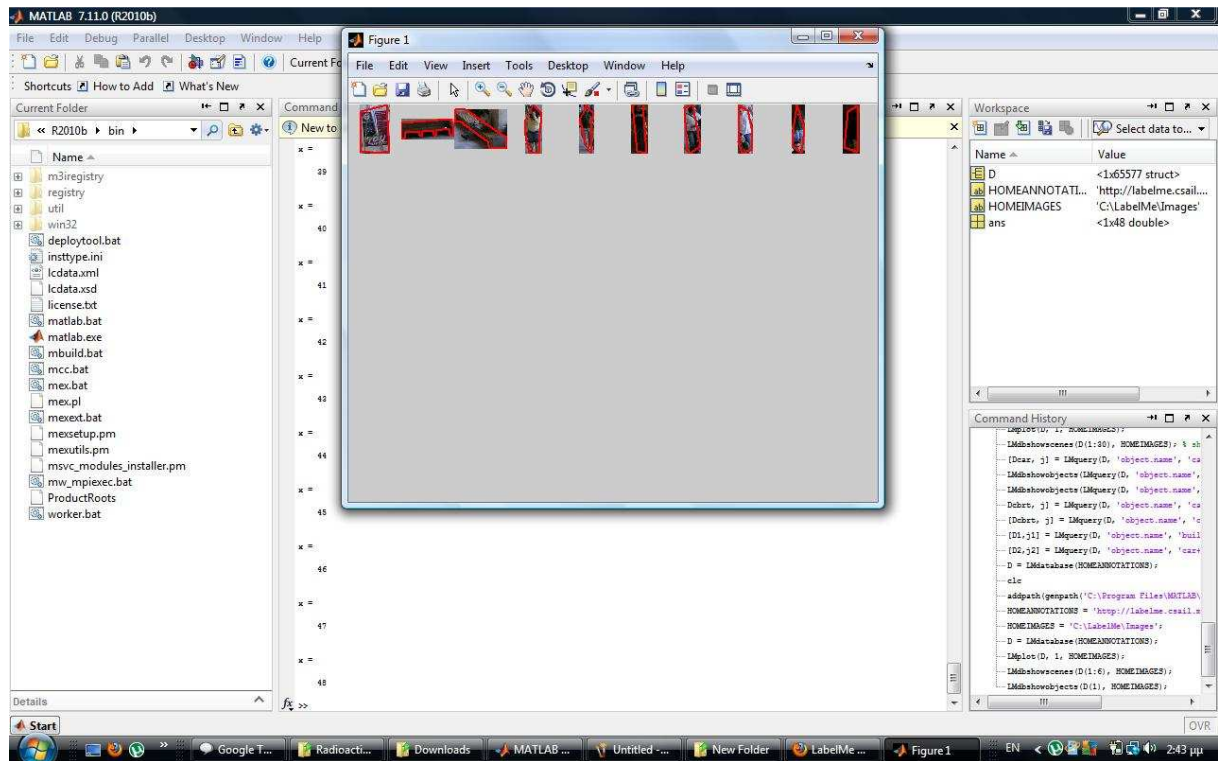
δείχνει τις 6 πρώτες εικόνες





Εικόνα26

LMdbshowobjects(D(1), HOMEIMAGES); δείχνει όλες τις επιμέρους επιλογές όλων των αντικειμένων στη πρώτη εικόνα.



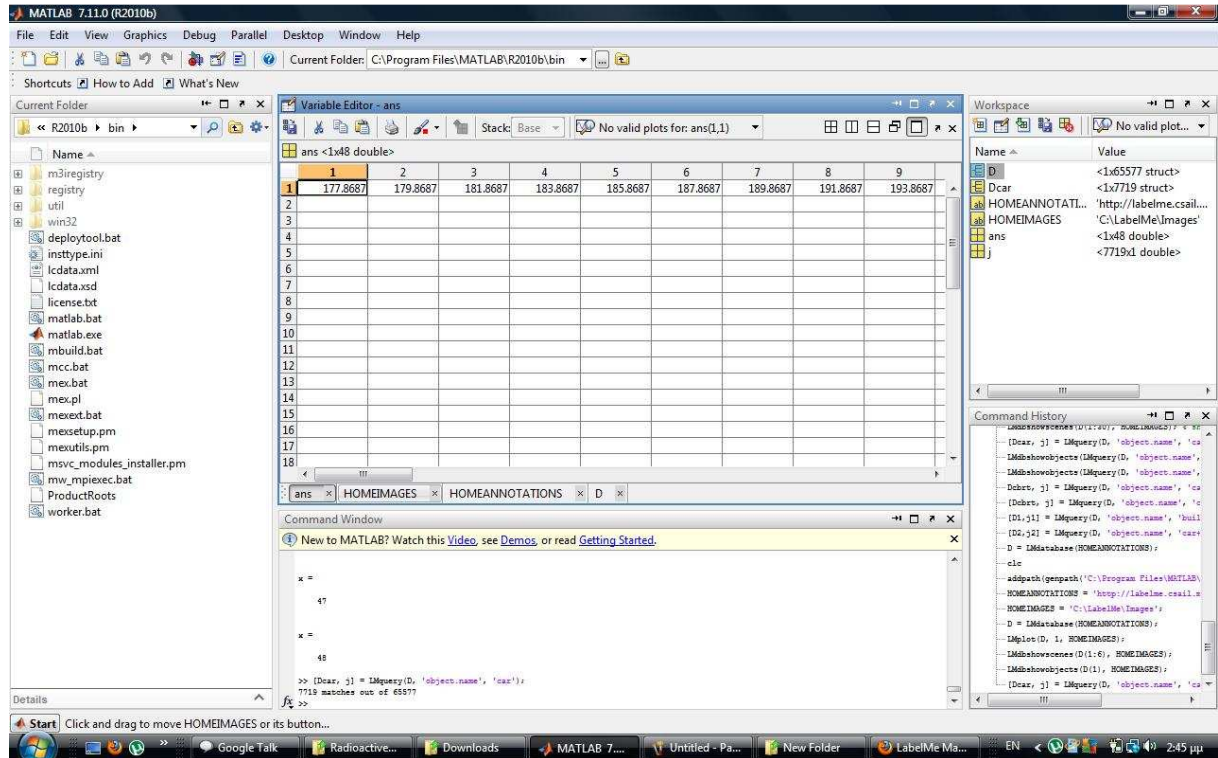
Εικόνα 27



## 4.6 Ερωτήματα

Για την εκτέλεση αναζητήσεων εικόνων, σκηνών αντικειμένων κτλ, μπορούμε να χρησιμοποιήσουμε την συνάρτηση LMquery. Η συνάρτηση αυτή επιτρέπει την αναζήτηση περιεχομένου σε οποιοδήποτε πεδίο.

```
[Dcar, j] = LMquery(D, 'object.name', 'car');
```



Εικόνα28

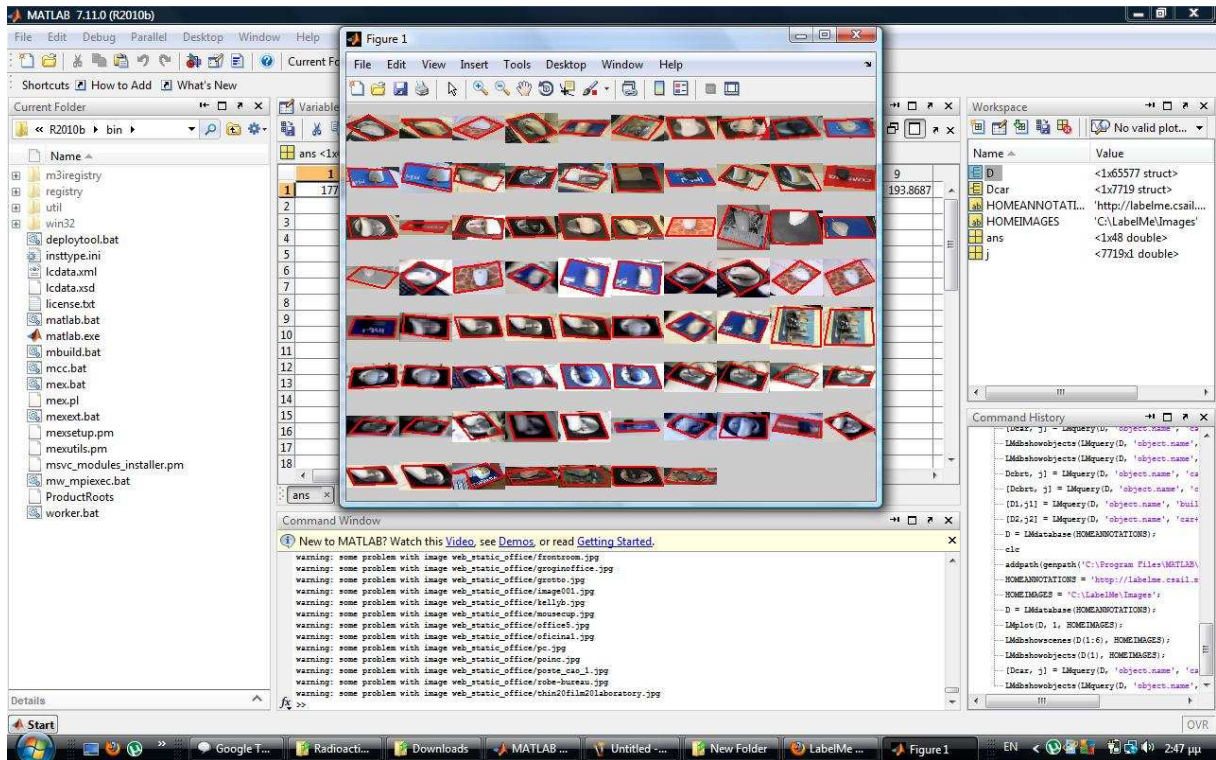
Η νέα δομή Dcar εμπεριέχει όλες τις εικόνες με αυτοκίνητα και όλα τα άλλα αντικείμενα θα έχουν απομακρυνθεί. Το ευρετήριο πίνακα j δείχνει

Στο αυθεντικό ευρετήριο D. Η δομή D(j) εμπεριέχει όλες τις εικόνες με αυτοκίνητα χωρίς να εξαιρεί τα άλλα αντικείμενα.

Η συνάρτηση LMquery δεν μπορεί να λάβει προκαθορισμένες λίστες πεδίων. Μπορούμε να χρησιμοποιήσουμε αυτή τη συνάρτηση για να θέσουμε ερώτημα σχετικά με σε κάθε πεδίο. Επομένως εάν προσθέσουμε νέα πεδία μέσα σε XML αρχεία επισημάνσεων, μπορούμε ακόμα να χρησιμοποιήσουμε LMquery για να ψάξουμε με σεβασμό το περιεχόμενο των νέων πεδίων.

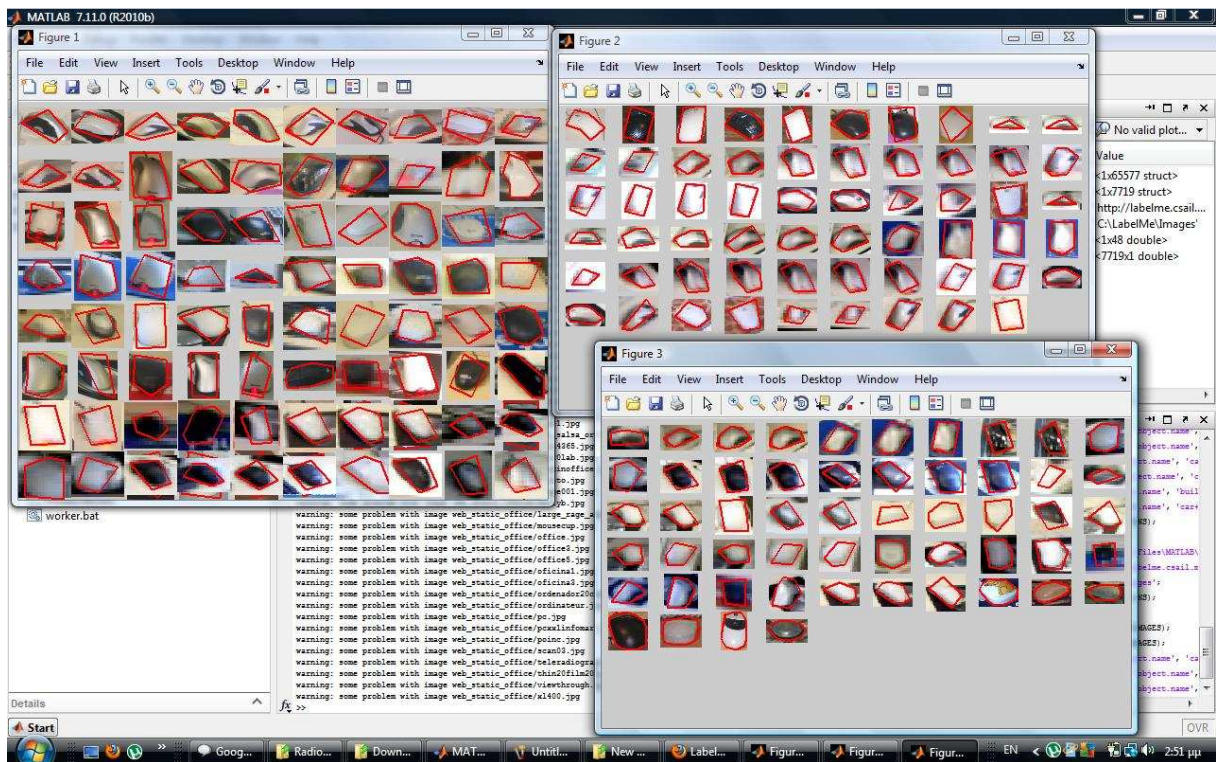
Ο αποκλεισμός μπορεί να χρησιμοποιηθεί για να περιορίσει μία αναζήτηση.

```
LMdbshowobjects(LMquery(D, 'object.name', 'mouse+pad'), HOMEIMAGES);
```



Εικόνα29

`LMdbshowobjects(LMquery(D, 'object.name', 'mouse-pad'), HOMEIMAGES);`



Εικόνα30

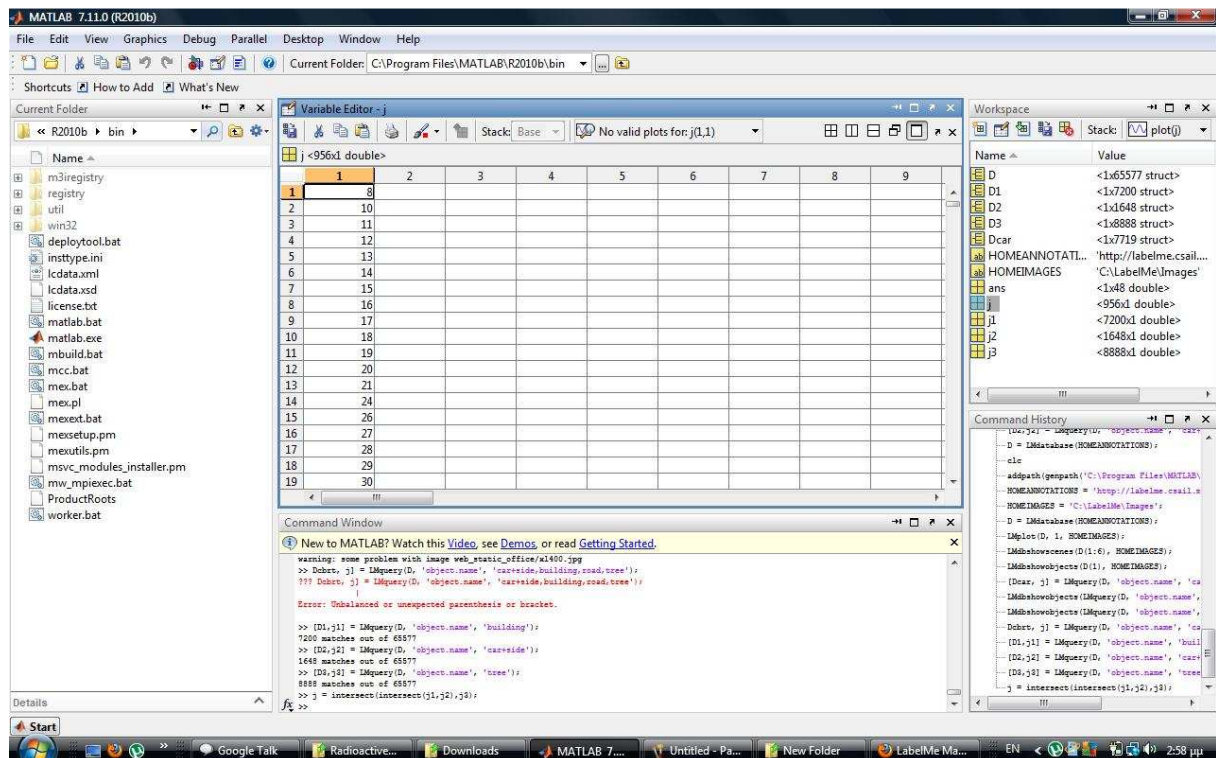


Μπορούμε επίσης να συνδυάσουμε αναζητήσεις. Παρακάτω επιλέγονται αντικείμενα τα οποία ανήκουν σε μία από τις παρακάτω ομάδες: 1) πλάγιες όψεις αυτοκινήτων , 2)κτίρια ,3) δρόμοι ή 4)δέντρα

```
[Dcbtr, j] = LMquery(D, 'object.name', 'car+side,building,road,tree');
```

Μπορούμε επίσης συνδυασμούς AND χρησιμοποιώντας αρκετά ερωτήματα. Επί παραδείγματι για να πάρουμε μία λίστα εικόνων οι οποίες εμπεριέχουν κτίρια, πλάγιες όψεις αυτοκινήτων και δέντρα κάνουμε το εξής:

```
[D1,j1] = LMquery(D, 'object.name', 'building');
[D2,j2] = LMquery(D, 'object.name', 'car+side');
[D3,j3] = LMquery(D, 'object.name', 'tree');
j = intersect(intersect(j1,j2),j3);
```



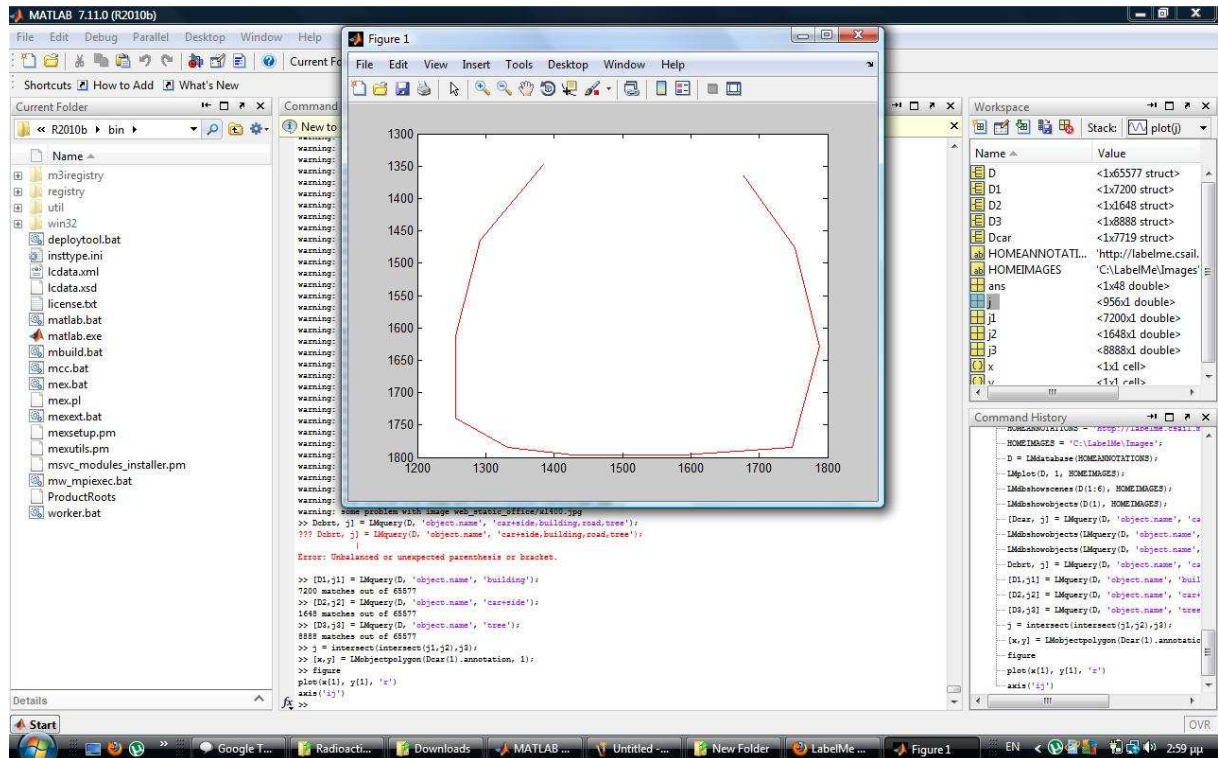
Εικόνα 31

Το ευρετήριο πίνακα j δείχνει σε όλες τις εικόνες οι οποίες εμπεριέχουν και τα τρία αντικείμενα. Αξίζει να σημειωθεί ότι ο D(j) μπορεί να περιέχει και άλλα αντικείμενα ,είναι όμως σίγουρο ότι θα περιέχει τα προηγούμενα τρία.

## 4.7Εξαγωγή πολυγώνων και τμημάτων

Για να εξάγουμε τις συντεταγμένες ενός πολυγωνικού αντικειμένου χρησιμοποιούμε την εξής συνάρτηση:

```
[x,y] = LMobjectpolygon(Dcar(1).annotation, 1);
figure
plot(x{1}, y{1}, 'r')
axis('ij')
```



Εικόνα 32

Σε αυτή τη περίπτωση η συνάρτηση επιστρέφει το πρώτο πολύγωνο στην πρώτη εικόνα του ευρετηρίου. Η συνάρτηση `Lmobjectpolygon` επιστρέφει ένα πίνακα κελιών. Χρειάζεται μία καταχώρηση για κάθε πολύγωνο.

Για την εξαγωγή μασκών κατάτμησης χρησιμοποιούμε την συνάρτηση `Lmobjectmask`:

```
[mask, class] = Lmobjectmask(D(1).annotation, HOMEIMAGES);
imshow(colorSegments(mask))
```

Με τη χρήση αυτής της συνάρτησης μπορούμε να εξάγουμε μάσκες κατάτμησης για όλα τα αντικείμενα τα οποία ανήκουν σε μία κατηγορία ή για ανεξάρτητα πολύγωνα.

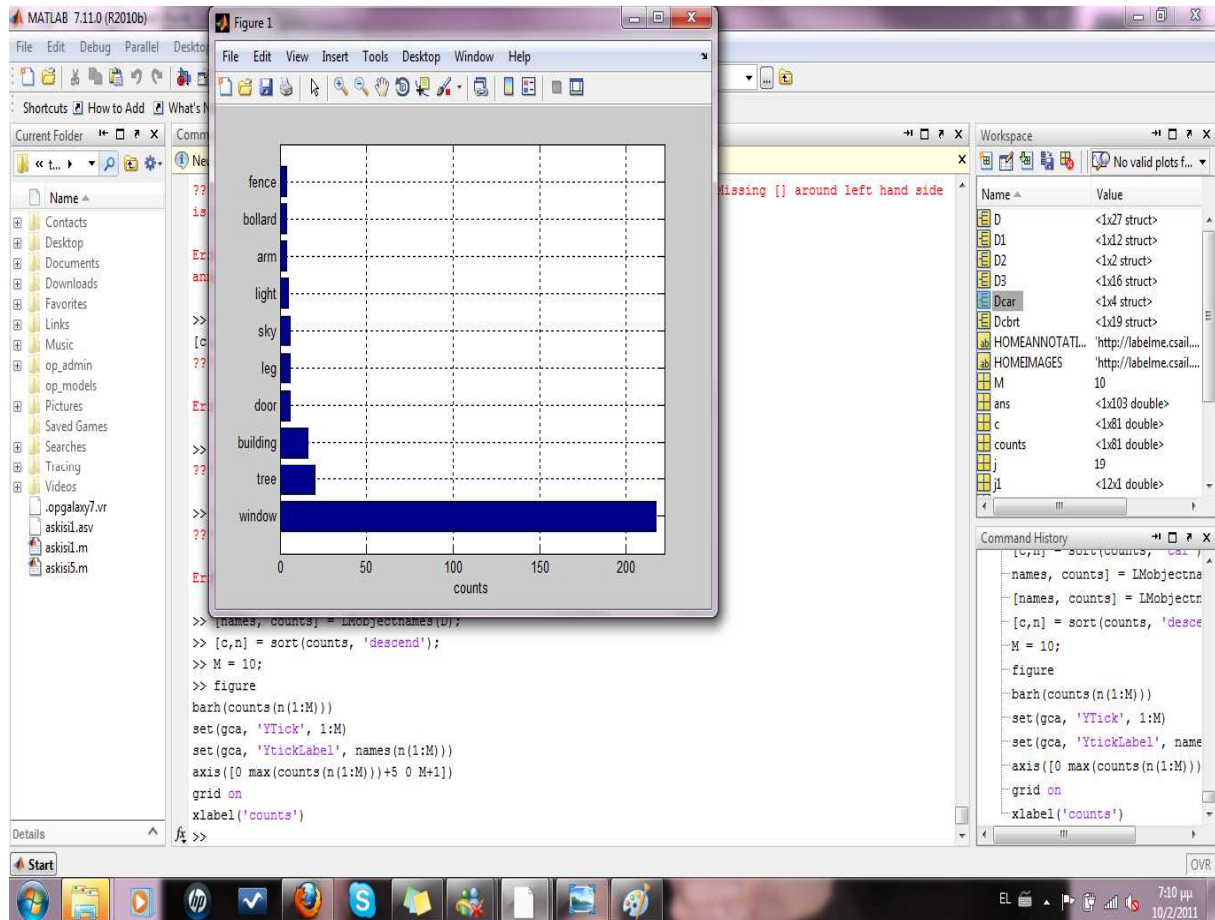
## 4.8 Συλλογή στατιστικών σχολιασμών

Η βάση δεδομένων περιέχει πολλά διαφορετικά ονόματα αντικειμένων. Προκειμένου να δούμε λίστα με τα ονόματα των αντικειμένων και τον αριθμό εμφάνισης κάθε αντικειμένου χρησιμοποιούμε τη συνάρτηση `Lmobjectnames`.

```
[names, counts] = Lmobjectnames(D);
[c,n] = sort(counts, 'descend');
```

```
M = 10; % number of objects to show
figure
barh(counts(n(1:M)))
set(gca, 'YTick', 1:M)
set(gca, 'YtickLabel', names(n(1:M)))
axis([0 max(counts(n(1:M)))+5 0 M+1])
```

grid on  
xlabel('counts')



Εικόνα33

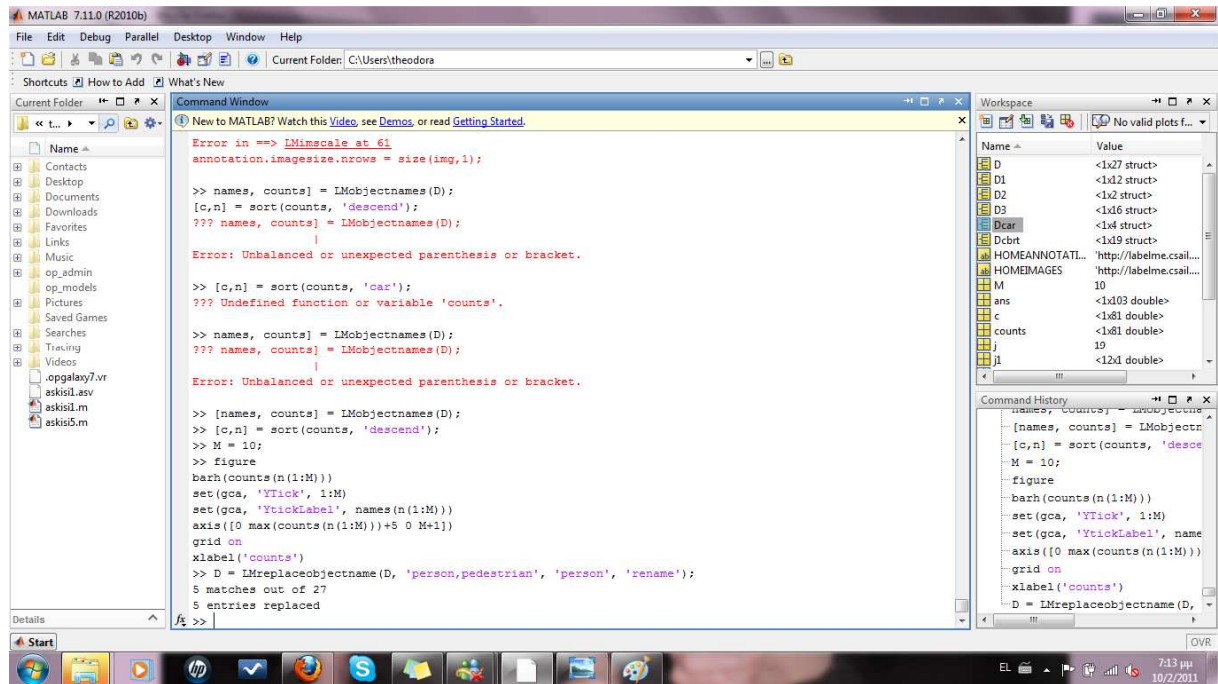
## 4.9 Αντιμετώπιση συνωνύμων και επισήμανση θορύβου

Δεδομένου ότι δεν υπάρχουν συγκεκριμένες οδηγίες για το πως πρέπει να εισαχθούν οι ετικέτες όταν χρησιμοποιούμε το εργαλείο σχολιασμού σε απευθείας σύνδεση οδηγούμεστε στο φαινόμενο διαφορετικές περιγραφές κειμένου να χρησιμοποιούνται για την ίδια κατηγορία αντικειμένων. Έτσι ένα άτομο μπορεί να περιγραφεί ως «άτομο», «πεζός», «κινούμενο άτομο», «παιδί» κτλ. Επομένως είναι σημαντικό να ενοποιηθούν τα σχόλια. Ο τρόπος με τον οποίο τα σχόλια θα ενοποιηθούν εξαρτάται από το τι θέλουμε να κάνουμε.

### 4.10 Lmreplaceobjectname

Η συνάρτηση αυτή είναι χρήσιμη όταν θέλουμε να αντικαταστήσουμε μερικά ονόματα αντικειμένων. Προκειμένου να αντικαταστήσουμε το όνομα ενός αντικειμένου χρησιμοποιούμε τη συνάρτηση Lmreplaceobjectname. Έτσι για να αντικαταστήσουμε το όνομα όλων των αντικειμένων που έχουν τη λέξη person ή pedestrian με τη λέξη person εκτελούμε την εξής εντολή:

```
D = Lmreplaceobjectname(D, 'person,pedestrian', 'person', 'rename');
```



Εικόνα34

## 4.11 Σχολιάστε τις δικές σας εικόνες

Η συνάρτηση **LMphotoalbum** δημιουργεί μία ιστοσελίδα με εικονίδια συνδεδεμένα με το εργαλείο σχολιασμού σε απευθείας σύνδεση. Με τη συνάρτηση αυτή μπορούμε να δημιουργήσουμε μία σελίδα στην οποία θα εμφανίζονται εικόνες προς σχολιασμό. Κάτι τέτοιο είναι χρήσιμο εάν θέλουμε άλλοι χρήστες να μας βοηθήσουν καθώς μπορούμε να δημιουργήσουμε μία σελίδα για κάθε χρήστη με διαφορετικό διαχωρισμό στα σετ εικόνων που θέλουμε να επισημανθούν.

**LMphotoalbum(folderlist, filelist, webpage, HOMEIMAGES);**

Για παράδειγμα εάν θέλουμε να δημιουργήσουμε μία ιστοσελίδα με αντικείμενα κουζίνας κάνουμε το εξής:

```
D = LMquery(D, 'folder', 'kitchen');
LMphotoalbum(D, 'myphotoalbum.html');
```

Εάν θέλουμε να σχολιάσουμε τις εικόνες μας πρέπει να τις ανεβάσουμε πρώτα στο LabelMe. Οι φωτογραφίες που ανεβάζουμε μαζί με τους σχολιασμούς που παρέχουμε είναι προς διάθεση για την έρευνα όρασης υπολογιστών ως μέρος της βάσης δεδομένων LabelMe.

## 4.12 ΑΛΛΕΣ ΧΡΗΣΙΜΕΣ ΕΝΤΟΛΕΣ

### 4.12.1 Για να κατεβάσουμε ένα μόνο φάκελο από την εργαλειοθήκη του MATLAB

```
HOMEIMAGES='/desired/path/to/Images';  
HOMEANNOTATIONS='/desired/path/to/Annotations';  
folderlist={'05june05_static_street_porter'};  
LMinstall(folderlist, HOMEIMAGES, HOMEANNOTATIONS);
```

### 4.12.2 Για να πάρουμε όλα τα ονόματα των αντικειμένων σε ένα φάκελο

```
HOMEANNOTATIONS = 'http://labelme.csail.mit.edu/Annotations';  
HOMEIMAGES = 'http://labelme.csail.mit.edu/Images';  
D = LMdatabase(HOMEANNOTATIONS, {'static_nature_web_outdoor_animal'}); [names, counts] =  
LMobjectnames(D);
```

### 4.12.3 Πως μπορούμε να βάλουμε μια ετικέτα (label) σε ένα υποσύνολο εικόνων

```
HOMEANNOTATIONS = 'http://labelme.csail.mit.edu/Annotations';  
HOMEIMAGES = 'http://labelme.csail.mit.edu/Images';  
D=LMdatabase(HOMEANNOTATIONS, {'spatial_envelope_256x256_static_8outdoorcategories'});  
for i = 1:length(D);  
folderlist{i} = D(i).annotation.folder;  
filelist{i} = D(i).annotation.filename;  
end  
LMthumbnailsbar(folderlist(1:500), filelist(1:500), 'page1.html', HOMEIMAGES);  
LMthumbnailsbar(folderlist(501:1000), filelist(501:1000), 'page2.html', HOMEIMAGES);
```

Ο κώδικας αυτός θα δημιουργήσει δύο σελίδες: page1.html και page2.html

### 4.12.4 Να αφαιρέσουμε όλα τα αντικείμενα που είναι πολύ μικρά

Χρησιμοποιώντας την εντολή `addsmallobjectlabel`

```
D = addsmallobjectlabel(D, height, width);
```

```
D = LMquery(D, 'object.name', '-smallobject'); % για να απαλλαγούμε από τα μικρά αντικείμενα
```

### 4.12.5 Να βρούμε μια λίστα με τις εικόνες που είναι πλήρως συμπληρωμένες



Πολλές φορές οι εικόνες οι οποίες έχουμε δεν έχουν παρά μόνο ελάχιστα σχολιασμένα αντικείμενα. Προκειμένου να φτιάξουμε μία λίστα από εικόνες με πολλά σχολιασμένα αντικείμενα εκτελούμε:

```
relativearea = LMlabeledarea(D, objectname);
```

```
j = find(relativearea>.9);
```

#### 4.12.6 Όταν ψάχνουμε σχολιασμένη εικόνα από τις κατηγορίες σκηνών(scene categories);

```
Dq = LMQuery(D, 'scenedescription', 'kitchen');
```

### 4.13 Κάνοντας τις εικόνες μας «Label Me»

Μέχρι τώρα σε αυτό το κεφάλαιο έχουμε δει σχολιασμένες εικόνες από διάφορα άλμπουμ που μας παρέχει το LABEL ME κάνοντας διάφορες αναζητήσεις των αντικειμένων τους. Σε αυτή την ενότητα βλέπουμε πως μπορούμε να ανεβάσουμε δικές μας εικόνες και να τις σχολιάσουμε. Παρακάτω δείχνουμε τα βήματα πώς να το κάνουμε. Στο τέλος δείχνουμε κάποιες δικές μας εικόνες που έχουν σχολιαστεί. Οι εικόνες αυτές είναι από διάφορα τοπία από διαφορετικές χώρες και δίκες μου προσωπικές.

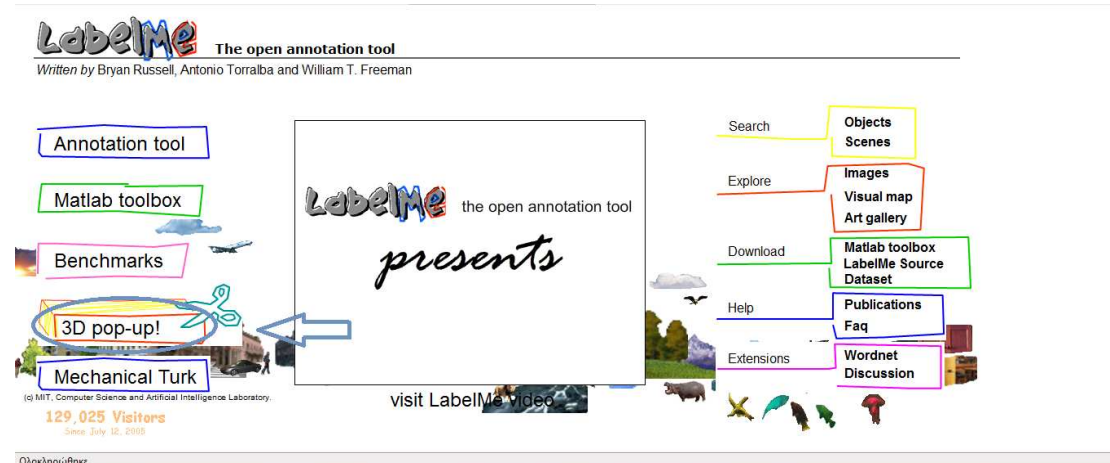
Βήμα 1. Από την αρχική σελίδα του label me κάνουμε κλικ στο 3D pop-up ( εικόνα 35)

Βήμα 2 .Στην συνέχεια πηγαίνουμε στο upload image (εικόνα 36)

Βήμα 3. Από το κουμπί αναζήτηση βρίσκουμε την εικόνα που θέλουμε. Η εικόνα πρέπει να είναι σε \*.jpg και μέγεθος μέχρι 1mb και μετά την ανεβάζουμε πατώντας το upload image (εικόνα 37)

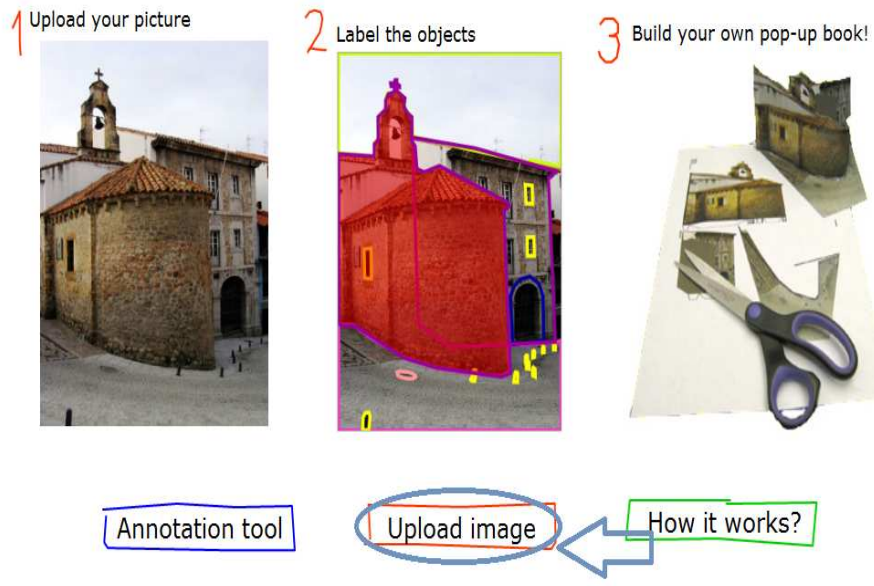
Βήμα 4. Δημιουργούμε διάφορα πολύγωνα στην εικόνα μας και τα ονομάζουμε

Έτσι έχουμε την εικόνα μας έτοιμη σχολιασμένη.



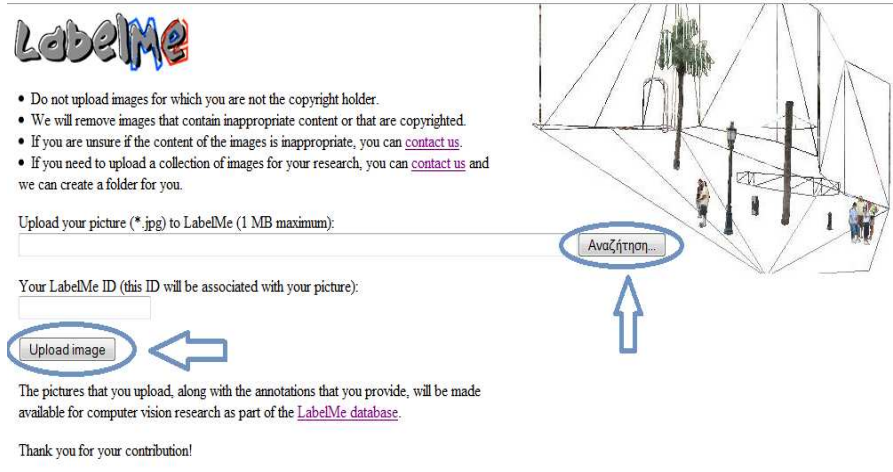
Εικόνα 35





<http://people.csail.mit.edu/brussell/research/AIM-2005-025-new.pdf>

Εικόνα 36



Εικόνα 37

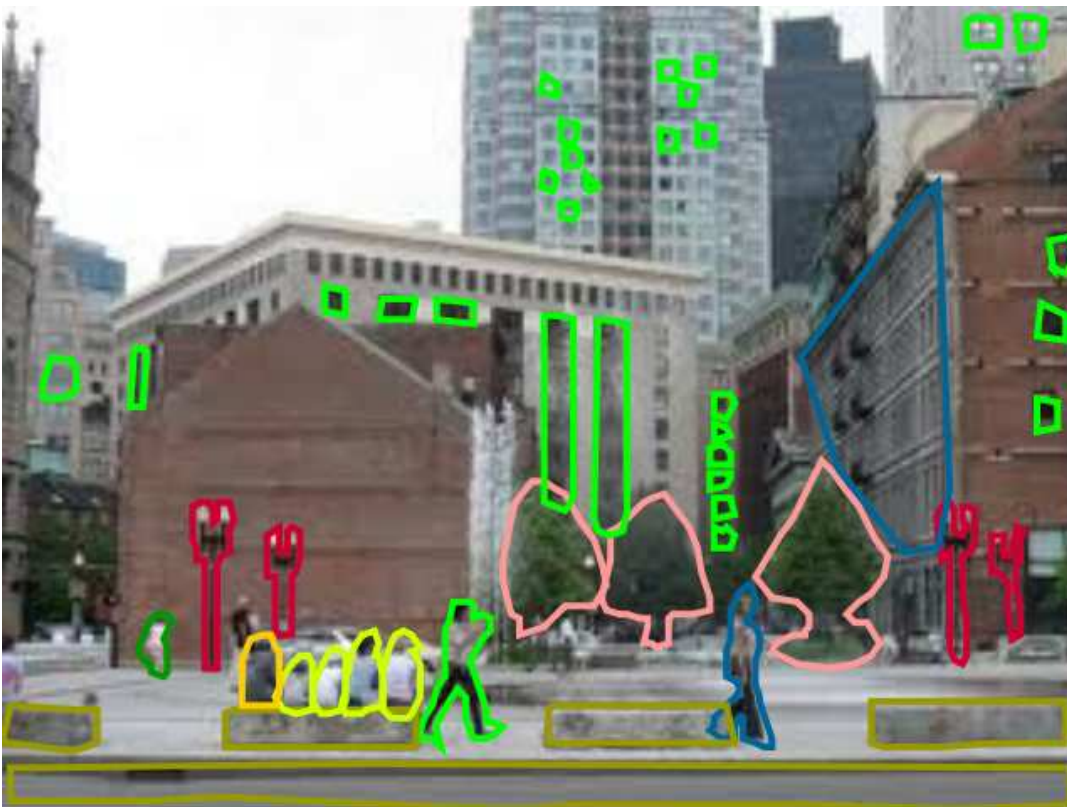
Ακολουθούν οι δικές μας σχολιασμένες εικόνες:



- [krewataki paralias](#)
- [krewataki paralias](#)
- [krewataki paralias](#)
- [omprella](#)
- [omprella](#)
- [parathuro](#)
- [parathuro](#)
- [hotel](#)
- [foinikas](#)
- [foinikas](#)
- [dwmatio](#)
- [dwmatio](#)
- [foinikas](#)
- [sunnefo](#)
- [krewataki paralias](#)
- [mpalkoni hotel](#)
- [foinikas](#)
- [tsoulithra](#)
- [ammos](#)

Εικόνα 38

Εικόνα 39



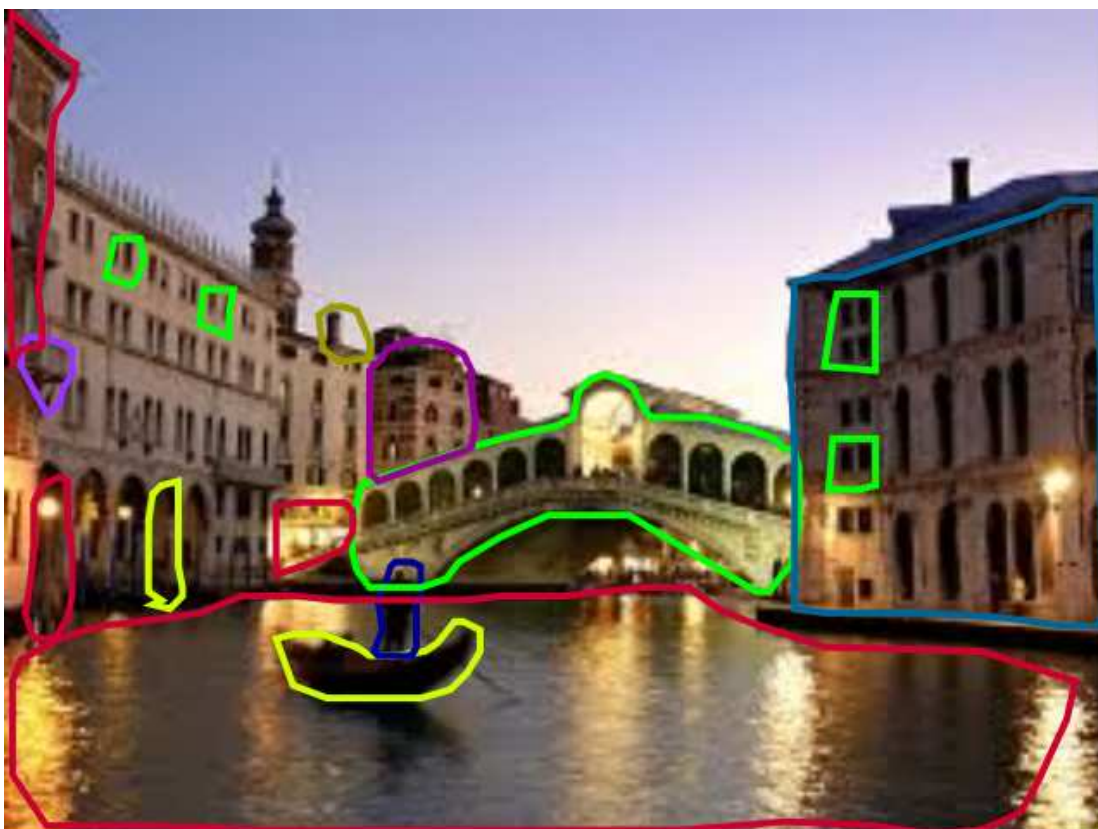
- [antras se kinisi](#)
- [gunaka se kinisi](#)
- [pezouli](#)
- [pezouli](#)
- [pezouli](#)
- [dentro](#)
- [dentro](#)
- [dentro](#)
- [parathuro](#)
- [parathuro](#)
- [parathuro](#)
- [parathuro](#)
- [parathuro](#)
- [parathuro](#)
- [parathuro](#)
- [gunaiika se kinisi](#)
- [lampa](#)
- [lampa](#)
- [lampa](#)
- [lampa](#)
- [ilikiomenos](#)
- [antras](#)
- [paidaki](#)
- [paidaki](#)
- [gria](#)
- [parathuro](#)
- [pezouli](#)
- [ktirio](#)
- [dromos](#)





- [trapezi](#)
- [bouno](#)
- [fvtaki](#)
- [fagito](#)
- [potiri me kراسi](#)
- [mpoukali kراسi](#)
- [maxairi](#)
- [pswmi](#)
- [piato](#)
- [tree](#)
- [kagkela](#)
- [sea](#)
- [karekla](#)
- [iliovasilema](#)
- [kaliva](#)
- [pirouni](#)
- [ladi](#)
- [potiri me nero](#)
- [vouno](#)

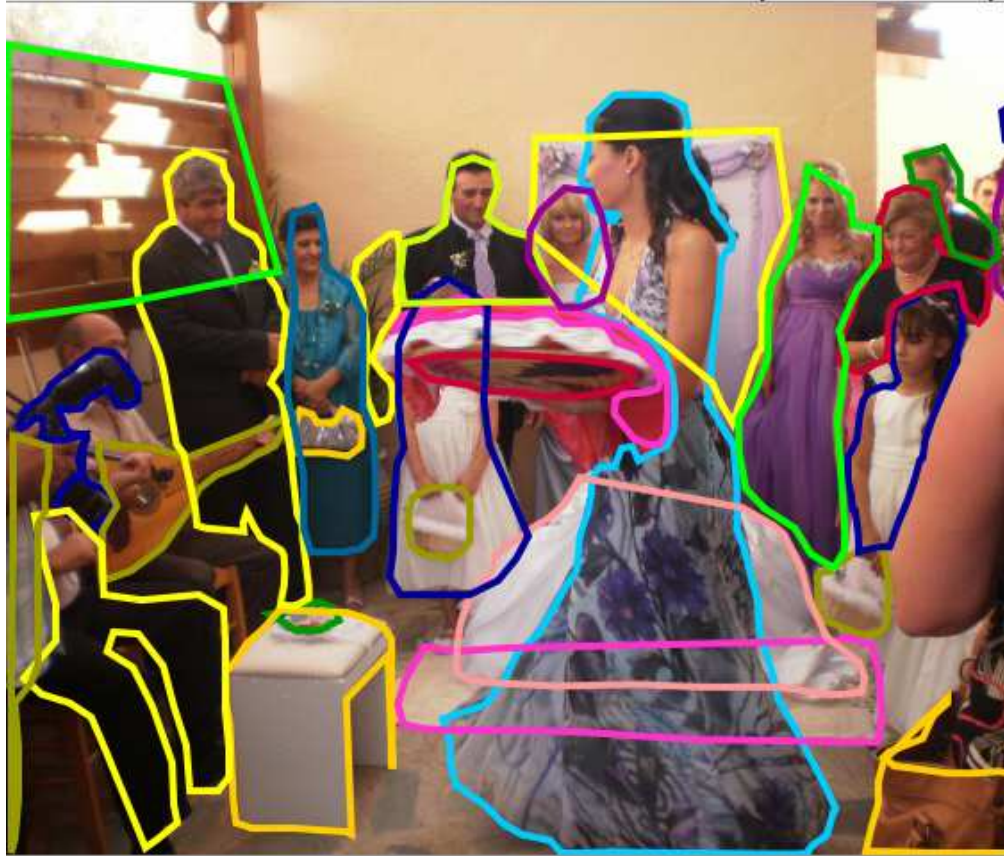
Εικόνα40



- [varka](#)
- [antras se ora douleias](#)
- [limni](#)
- [gefura](#)
- [ktirio](#)
- [lampa](#)
- [parathuro](#)
- [kamara](#)
- [parathuro](#)
- [kaminada](#)
- [hotel](#)
- [mpalkoni](#)
- [parathuro](#)
- [katastima](#)
- [parathuro](#)
- [diamerismata](#)

Εικόνα 41





[egw otan imoun koumpara](#)  
[psifiaki fotografiki](#)  
[tsanta](#)  
[kalathaki](#)  
[paneri](#)  
[kokkino mantili](#)  
[euro](#)  
[trapezaki](#)  
[paradosiako xali](#)  
[nufiko](#)  
[laouto](#)  
[futo](#)  
[paranumfaki](#)  
[giagia nufis](#)  
[louloudia](#)  
[diakosmitiko fonto](#)  
[antras me koustoumi\\_petheros nufis](#)  
[kalathaki](#)  
[tsanta](#)  
[paranumfaki](#)  
[ksilo](#)  
[father nufis](#)  
[mother nufis](#)  
[kopela se poza](#)  
[adelfi nufis](#)  
[pethera nufis](#)  
[xeri kopelas](#)  
[antras pou paizei violi](#)  
[adelfos nufis](#)  
[qualia iliou](#)  
[agori pou parakoloutha](#)  
[fotografos](#)

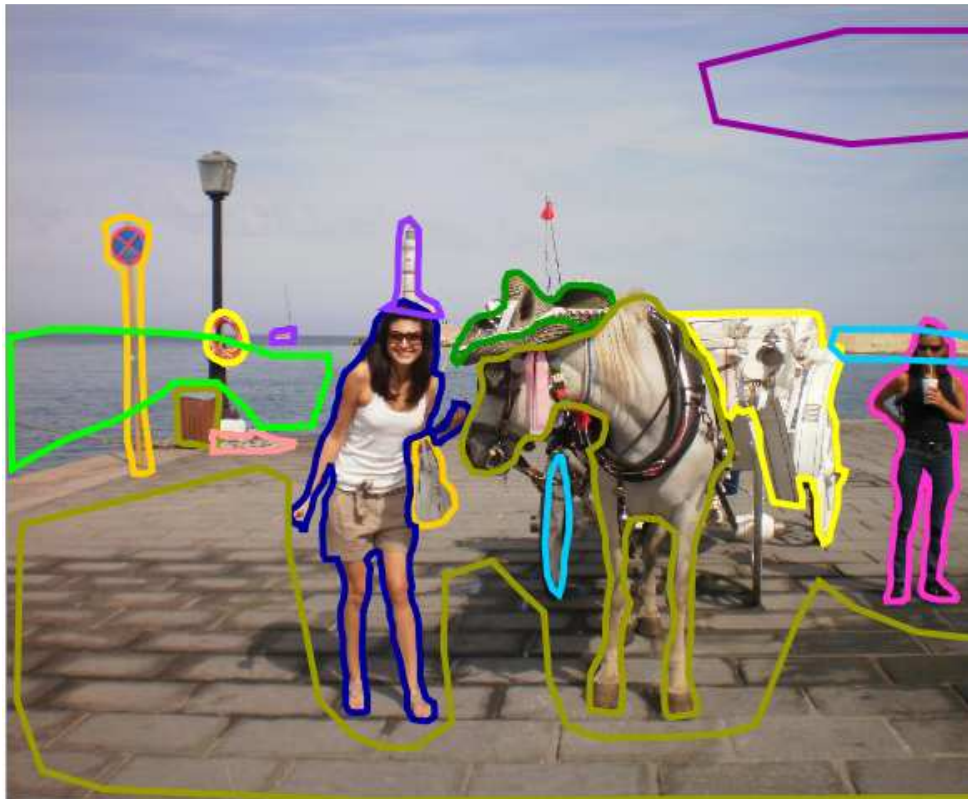
Εικόνα 42



[rosida touristria](#)  
[eleni xoreontas xasapiko](#)  
[epixirimatias](#)  
[egw xoreontas me tin eleni](#)  
[ksanthia pou koitazei ena agori](#)  
[papoutsi](#)  
[lampa](#)  
[lampa](#)  
[diakosmitika stolidia](#)  
[mallia antra](#)  
[kopella pou xoreuei](#)  
[fwtaki](#)  
[fwtaki](#)  
[fwtaki](#)  
[fwtaki](#)  
[kourtina](#)  
[faraklos](#)  
[lampa](#)  
[lampa](#)  
[papoutsi antra](#)

Εικόνα 43





[sosivio](#)  
[kados axristwn](#)  
[sxoini](#)  
[tsanta](#)  
[kapelo alogou](#)  
[kopela en wra egrasias](#)  
[pagkaki](#)  
[varkoula](#)  
[egw stin palia poli xaniwn](#)  
[troxos](#)  
[amaksa](#)  
[alogo](#)  
[faros](#)  
[sima troxaias](#)  
[limani](#)  
[pezodromio](#)  
[sunefo](#)  
[toixi](#)

Εικόνα 44



[egw sta matata](#)  
[kafeteria](#)  
[apartment](#)  
[kafeteria](#)  
[apartment](#)  
[vouno](#)  
[dentro](#)  
[gualia iliou](#)  
[vrxkos apo tin spilia](#)  
[pelta](#)  
[vrxkos apo tin spilia](#)  
[omprella](#)  
[omprella](#)  
[omprella](#)  
[touristas apo gallia](#)  
[touristas apo gallia](#)  
[touristas apo gallia](#)  
[zeugari pou kanoun iliotherapia katw stin ammo](#)  
[skini](#)  
[omprella](#)  
[omprella](#)  
[estiatorio](#)  
[touristas pou kolimpaie](#)  
[zeugari pou kolimpane](#)  
[italos pou kanei iliotherapia](#)  
[germanida pou perpataei](#)  
[paidia pou paizoun stin ammo](#)  
[ammos](#)  
[spiti](#)  
[parathuro](#)  
[apartment](#)  
[apartment](#)  
[mpiraria](#)  
[italios touristas](#)  
[italos ilikiomenos pou kolimpaie](#)  
[italida pou koitazei tin germanida](#)  
[paralia sta matata](#)  
[skini pou tha koimithi to erwteumeno](#)  
[zeugari](#)  
[pateras ellinas pou prosexei to paidi tou kuma](#)  
[mpalkoni](#)  
[eksoxiko](#)  
[touristes- parea apo ellines](#)  
[omprella](#)  
[omprella](#)  
[omprella](#)  
[omprella](#)  
[omprella](#)  
[omprella](#)  
[omprella](#)  
[kopella ksaplwmeni sto krewataki](#)  
[spiti](#)  
[apartment](#)  
[apartment](#)  
[omprella](#)  
[omprella](#)  
[tsenta](#)

Εικόνα 45

## ΠΑΡΑΡΤΗΜΑ

### 1. Χειρισμός εικόνων

Για να πάρουμε ένα τμήμα και να αλλάξουμε το μέγεθος εικόνων και σχολιασμών χρησιμοποιούμε τις συναρτήσεις :

LMimscale – διαβαθμίζει μία εικόνα και τον αντίστοιχο σχολιασμό

LMimcrop – παίρνει ένα τμήμα μίας εικόνας και τον αντίστοιχο σχολιασμό

LMimpad – επικαλύπτει μία εικόνα με PADVAL και τροποποιεί τον σχολιασμό

### 2. Αναγνώριση Σκηνής

Σύνοψη περιγραφέα

Για να υπολογίσουμε τη σύνοψη περιγραφέα σε μία εικόνα χρησιμοποιούμε τη συνάρτηση LMgist. Στο παράδειγμα διαβάζουμε μία εικόνα και υπολογίζουμε τον περιγραφέα.

```
% Load image
img = imread('demo1.jpg');

% Parameters:
param.imageSize = 128;
param.orientationsPerScale = [8 8 8 8];
param.numberBlocks = 4;
param.fc_prefilt = 4;

% Computing gist:
[gist, param] = LMgist(img, "", param);

% Visualization
figure
subplot(121)
imshow(img)
title('Input image')
subplot(122)
title('Descriptor')
showGist(gist, param)
```

Μπορούμε επίσης να υπολογίσουμε την κεντρική ιδέα μίας συλλογής εικόνων :

```
gist = LMgist(D, HOMEIMAGES, param);
```

Το αποτέλεσμα είναι ένας πίνακας μεγέθους [Nscenes Nfeatures], όπου Nscenes = length(D).



## **ΚΕΦΑΛΑΙΟ 5**

### **5.1 Η εικόνα ως πολυμεσικό αντικείμενο**

Από την αρχή αυτής της πτυχιακής εργασίας έγινε σαφές πως αντιμετωπίσαμε την εικόνα ως πολυμεσικό αντικείμενο. Με μία πρώτη ματιά κάτι τέτοιο φαίνεται παράδοξο καθώς έχουμε συνηθίσει την έννοια πολυμέσα να την συσχετίζουμε άμεσα με κινούμενη εικόνα και ήχο. Αναλύοντας όμως τόσο την ετοιμολογία της λέξης αλλά και τον ορισμό των πολυμέσων εύκολα βλέπουμε πως η εικόνα ως αντικείμενο σχολιασμού και αλληλεπίδρασης μεταξύ των χρηστών, από ένα στατικό αντικείμενο μπορεί να μετατραπεί σε ένα πολυμεσικό διαδραστικό αντικείμενο. Φυσικά κάθε εικόνα δεν είναι πολυμεσικό αντικείμενο μπορεί όμως να γίνει με την χρήση Εργαλείων Σχολιασμού Εικόνας όπως αυτά που αναφέρθηκαν στο τρίτο κεφάλαιο. Με τον όρο σχόλιο δεν εννοούμε κατά ανάγκη μόνο τον σχολιασμό κειμένου αλλά ακόμα και τη προσθήκη ήχου σε μία εικόνα.

Είναι γεγονός ότι η ανάπτυξη των Εργαλείων Σχολιασμού Εικόνας ξεκίνησε με την ανάπτυξη της ψηφιακής εικόνας, η οποία είχε με τη σειρά της ως αποτέλεσμα την αύξηση του όγκου των δεδομένων που είχε ένας απλός χρήστης στη διάθεση του. Αρκεί να αναλογιστούμε την πληθώρα ψηφιακών εικόνων που έχει ο κάθε ένας μας στον υπολογιστή του και θα δούμε πως η αναζήτηση αλλά και η ταξινόμηση των εικόνων είναι αρκετά χρονοβόρα και δύσκολη υπόθεση πρόβλημα αυτό ήρθαν να αντιμετωπίσουν τα Εργαλεία Σχολιασμού Εικόνας τα οποία ανάλογα με τις δυνατότητές τους μπορούν να δημιουργήσουν ερωτήματα αναζήτησης αλλά και ταξινόμησης των εικόνων μας. Πέραν όμως της χρησιμότητάς τους στον απλό χρήστη τα Εργαλεία Σχολιασμού Εικόνας Βρίσκουν εφαρμογή σε πιο εξειδικευμένες εφαρμογές. Οι ηλεκτρονικές βιβλιοθήκες είναι που θα αναλύσουμε σε αυτό το κεφάλαιο.

### **5.2 Ηλεκτρονικές Βιβλιοθήκες**

Οι ψηφιακές βιβλιοθήκες αποτελούν συστήματα που υλοποιούν υπηρεσίες πληροφόρησης και υποστηρίζουν καθημερινές πρακτικές και ζωτικές λειτουργίες κοινοτήτων όπως η ακαδημαϊκή και η ερευνητική, ενώ η συμβολή των ακαδημαϊκών βιβλιοθηκών στην ανάπτυξη τους, ειδικότερα στην χώρα μας, κρίνεται ιδιαίτερα σημαντική. Η τεχνολογική εξέλιξη και η σύζευξή της με τις απαιτήσεις των χρηστών για την εκτέλεση κρίσιμων εργασιών έχει αυξήσει τους ρυθμούς εξέλιξης τους και έχει καταστήσει συχνότερες τις αναπροσαρμογές στο ύψος και τη μορφολογία τους. Ειδικότερη μέριμνα απαιτεί ο τομέας ανάπτυξης υπηρεσιών ψηφιακών βιβλιοθηκών, ο οποίος σε μεγάλο μέρος κρίνει και την φυσιογνωμία τους συγκριτικά με τις φυσικές βιβλιοθήκες, αλλά και επεμβαίνει σε κρίσιμες φάσεις της πληροφοριακής ροής.

Μεγάλος λόγος γίνεται για τις ειδοποιούς διαφορές μεταξύ ψηφιακών και φυσικών βιβλιοθηκών σε μια προσπάθεια οριοθέτησης όχι μόνο του αντικειμένου, αλλά και των συνεπαγόμενων ρόλων και διαδικασιών. Η προσπάθεια προσδιορισμού των ψηφιακών βιβλιοθηκών περνά μέσα από ποικίλους διαύλους, οι οποίοι προωθούν είτε μια ιστορική αντιμετώπιση, είτε μια προσέγγιση με τεχνικό-επιστημονικές μεθόδους. Μια εξ αυτών είναι η ανίχνευση και ο εντοπισμός των υπηρεσιών αυτών που διαφέρουν και οι οποίες συνεπάγονται είτε πρότερης προεργασίας και υποδομής, είτε ύστερων αποτελεσμάτων και οφελών. Σε αυτό το στάδιο εισέρχεται η έννοια της προστιθέμενης αξίας, η οποία είναι άμεσα συνδεδεμένη με τις αυξητικές επιδράσεις του ηλεκτρονικού περιβάλλοντος. Η ανίχνευση όμως αυτή πολύ δε περισσότερο, η προτυποποίηση των υπηρεσιών αυτών, αποτελεί μια δύσκολη διαδικασία, η οποία έχει επιχειρηθεί αρκετά στο παρελθόν. Χαρακτηριστικό παράδειγμα αποτελεί η εργασία του

Goncalves (2004), ο οποίος συγκέντρωσε ένα σύνολο υπηρεσιών, τις οποίες κατηγοριοποίησε σε υπηρεσίες ικανοποίησης και υπηρεσίες υποδομής. Σύμφωνα με τον Goncalves οι υπηρεσίες προστιθέμενης αξίας αποτελούν υποκατηγορία των υπηρεσιών υποδομής μαζί με τις υπηρεσίες σχολιασμού, οπτικοποίησης κ.ά. Από την άλλη οι Candela et al. (2006) διαίρεσαν τις υπηρεσίες σε τέσσερις μεγάλες κατηγορίες, τις υπηρεσίες πρόσβασης, υποβολής, διαχείρισης και εξατομίκευσης. Σε όλες αυτές τις φάσεις, η δυνατότητα συνεργατικής αναζήτησης και ανάκτησης της πληροφορίας, αποτελεί πλέον μια διεθνή πρακτική, ειδικά στο χώρο της επιστημονικής πληροφόρησης, για την οποία δεν έχει επαρκώς ερευνηθεί και κατανοηθεί η συμπεριφορά των χρηστών, έτσι ώστε να δοθεί η δυνατότητα ανάπτυξης υπηρεσιών προστιθέμενης αξίας από συστήματα ψηφιακών βιβλιοθηκών.

Οι παράγοντες που επηρεάζουν τη χρήση μίας υπηρεσίας χωρίζονται σε τρεις κατηγορίες: (α) τους παράγοντες του συστήματος, (β) τους παράγοντες της υπηρεσίας και (γ) τους προσωπικούς παράγοντες. Οι παράγοντες του συστήματος σχετίζονται με την χρηστικότητα (σχεδιασμός, δόμηση και αισθητική της διεπαφής) και την απόδοση του συστήματος (αξιοπιστία, ταχύτητα απόκρισης) της ψηφιακής βιβλιοθήκης. Σε ένα δεύτερο επίπεδο οι παράγοντες υπηρεσίας σχετίζονται με το σύνολο και τη φύση των υπηρεσιών που προσφέρει μία ψηφιακή βιβλιοθήκη. Οποιαδήποτε λοιπόν επικάλυψη ή ομοιότητα μεταξύ δύο ή περισσότερων υπηρεσιών λειτουργεί αποτρεπτικά σε βάρος αυτής που δεν έχει συνηθίσει να χρησιμοποιεί ο χρήστης. Σε ενδεχόμενο δίλημμα μεταξύ δύο υπηρεσιών οι χρήστες φαίνεται πως στρέφονται προς εκείνες που είναι πιο κοντά στις συνήθειές τους. Παραδείγματος χάριν, οι υπηρεσίες του Δικτύου Συγγραφέων και του Θεματικού Χάρτη είχαν μικρή απήχηση, διότι οι χρήστες θεώρησαν πως καλύπτονται από τα πεδία «Συγγραφέας» και «Θέμα» στην αναζήτηση.

Τέλος, οι προσωπικοί παράγοντες αφορούν στις ιδιότητες και τις εμπειρίες που φέρουν οι χρήστες. Οι χρήστες επηρεάζονται από τις εικόνες που φέρουν από την καθημερινή τους επαφή με τις ψηφιακές βιβλιοθήκες. Η συνήθεια και η ευκολία εκμάθησης μίας υπηρεσίας παίζει καθοριστικό ρόλο στην αποδοχή μίας υπηρεσίας. Μολονότι οι ψηφιακές βιβλιοθήκες παρέχουν ανεπτυγμένα εργαλεία για την καλύτερη κάλυψη των αναγκών των χρηστών, εντούτοις οι τελευταίοι παραμένουν πιστοί σε εργαλεία συνήθους χρήσης. Η καθημερινή επαφή με διαδικτυακές εφαρμογές δημιουργεί εμπειρίες στους χρήστες, οι οποίοι τις μεταφέρουν και σε άλλα περιβάλλοντα. Οποιαδήποτε απόκλιση από τις εικόνες που έχουν διαμορφώσει, λειτουργεί αποτρεπτικά.

Από την ερμηνεία των αποτελεσμάτων προέκυψε, πως οι παραπάνω παράγοντες αλληλεπιδρούν μεταξύ τους όταν πρόκειται για την αποδοχή των υπηρεσιών. Για παράδειγμα το εύρος των υπηρεσιών που προσφέρει μία βιβλιοθήκη ενθαρρύνει τόσο την γενικότερη χρήση της, όσο και την χρήση των ειδικότερων υπηρεσιών. Συνεπώς, οι χρήστες μπορούν να χρησιμοποιήσουν υπηρεσίες επηρεαζόμενοι από την συνολική εικόνα του συστήματος. Το σύστημα, οι υπηρεσίες και ο προσωπικός παράγοντας διαμορφώνουν το πλαίσιο της αποδοχής μίας υπηρεσίας. Για να χρησιμοποιηθεί μία υπηρεσία πρέπει να συνδυάζει χρηστική και συναισθηματική αξία σύμφωνα με τον Boztere (2007). Ο προσωπικός παράγοντας συνδέεται άρρηκτα με την εμπειρία, τις ανάγκες, τις αντιλήψεις και σε τελική ανάλυση την ικανοποίηση, η οποία σημαίνει ταυτόχρονα και αποδοχή.

## 52.1. Υπηρεσίες συνεργατικής ανάκτησης πληροφορίας

Η συνεργατική ανάκτηση πληροφορίας, σύμφωνα με τους Hansen και Jarvelin (2005) αποτελεί μια δραστηριότητα που σχετίζεται με την επίλυση ενός συγκεκριμένου προβλήματος και η οποία περιλαμβάνει την αλληλεπίδραση μεταξύ ανθρώπων, άμεσα ή και έμμεσα (π.χ. μέσω κειμένων, σχημάτων και εικόνων). Επιπρόσθετα η συνεργατική ανάκτηση πληροφορίας λαμβάνει χώρα είτε σε ένα ορισμένο πεδίο εργασίας, είτε σε μια ανοιχτή κοινότητα. Εντός αυτού του πλαισίου αναπτύσσεται η επίγνωση μιας δραστηριότητας (activity awareness) (Carroll et al., 2005), η οποία αποτελείται από τέσσερα χαρακτηριστικά: (α) το κοινό υπόβαθρο (common ground), που αποτελεί το πρωτόκολλο επικοινωνίας μεταξύ των συνεργατών, (β) οι κοινότητες εργασίας (communities of practice), που αναφέρονται στην συνεργασία, στο σχεδιασμό και στην ανταλλαγή συμβουλών, ιδεών και απόψεων, (γ) το κοινωνικό κεφάλαιο (social capital), που είναι τα δίκτυα αλληλεπίδρασης που δημιουργούνται για την επίτευξη κοινών στόχων και (δ) η ανθρώπινη ανάπτυξη (human development), που αποτελεί

την ατομική πρόοδο και την ανάπτυξη των προσωπικών ικανοτήτων των μελών μιας ομάδας μέσα από τη συνεργασία. Έως τώρα η συνεργασία περιοριζόταν σε συναντήσεις που λάμβαναν χώρα σε πραγματικό χρόνο, αλλά πλέον καθίσταται προφανές ότι τα άτομα τείνουν να συνεργάζονται απομακρυσμένα χάρις στην ανάπτυξη του διαδικτύου, μεταλλάσσοντας καθ' αυτό τον τρόπο την έννοια του όρου «ομάδα». Η απομακρυσμένη συνεργατική δραστηριότητα δημιουργεί κρίσιμα ζητήματα, όπως είναι αυτό της εμπιστοσύνης για την επιτυχή ολοκλήρωση των καθηκόντων και την οικοδόμηση συμπαγών δομών εργασίας (Nandhakumar & Baskerville, 2006). Το κατεξοχήν σύγχρονο συνεργατικό εργαλείο είναι η υπηρεσία αποστολής μηνυμάτων, η χρήση της οποίας, είτε μέσα στο διαδίκτυο είτε μέσα σε τοπικά δίκτυα, είναι ευρέως διαδεδομένη. Αρχικά, οι υπηρεσίες σύγχρονης ανταλλαγής μηνυμάτων (chatting) χρησιμοποιούνταν για διασκέδαση, ενώ πλέον δίνουν τη δυνατότητα ανταλλαγής μηνυμάτων και αρχείων εικόνας, κειμένου και βίντεο, δημιουργίας κοινοτήτων, ακόμα και συνομιλίας μέσω κάμερας, δημιουργώντας έναν τύπο ηλεκτρονικής συνεδρίασης. Οι κοινοί πληροφοριακοί χώροι είναι περιοχές που δημιουργούνται από τα μέλη μιας ομάδας, στους οποίους αποθηκεύουν τεκμήρια και ψηφιακά αντικείμενα σχετικά με τα ενδιαφέροντά τους και ακολούθως τους δίνεται η δυνατότητα εξερεύνησης και διαχείρισης του (Davis et al., 2001). Οι κοινοί αυτοί χώροι αποτελούν δεξαμενές πληροφοριών για τις διάφορες ομάδες από τις οποίες μπορούν να ανακτούν επεξεργασμένη (διηθημένη) πληροφορία, μέσα από οργανωτικές (π.χ. ταξινόμηση) και διαχειριστικές (π.χ. σχολιασμός) δραστηριότητες. Ένα από τα πιο σημαντικά συνεργατικά εργαλεία είναι η υπηρεσία σχολιασμού (annotations). Ως σχολιασμός νοείται οποιοδήποτε σχόλιο ή ετικέτα (tag) προστίθεται σε μια υπερσύνδεση ή κείμενο χαρακτηρίζοντάς το. Η ανάρτηση ή επισύναψη σχολιασμών έχει μια τριπλή υπόσταση. Αρχικά, αποτελεί τη δημιουργία μεταδεδομένων, καθώς αποτελεί προστιθέμενη πληροφορία για ένα τεκμήριο, η οποία και το χαρακτηρίζει (Nagao, 2003). Επιπλέον, οι σχολιασμοί αυτοί αποτελούν αυτοτελές περιεχόμενο, γεγονός που σημαίνει ότι μετασχηματίζονται οι σχέσεις μεταξύ των περιεχομένων ή σχηματίζονται νέες, δημιουργώντας ένα είδος υπερκειμένου (Agosti et al., 2004). Τέλος, οι σχολιασμοί αποτελούν επικοινωνιακές πράξεις (Searle, 1979), ενισχύοντας τους δεσμούς συνεργασίας σε μια ομάδα. Το πρόβλημα που προκύπτει από την ύπαρξη σχολιασμών και την επικόλληση ετικετών από διαφορετικά άτομα, είναι κυρίως σημασιολογικό, καθώς δημιουργούνται προβλήματα πολυσημίας και συνωνυμίας σε βασικό επίπεδο. Παρόλα αυτά, εκτιμάται ότι το πρόβλημα αυτό λύνεται λόγω του κοινού υπόβαθρου που υπάρχει ανάμεσα στα μέλη μιας ομάδας, σημειώνοντας καθ' αυτόν τον τρόπο τη σημασία του περιβάλλοντος μέσα στο οποίο λαμβάνουν χώρα κάποιες ενέργειες (στην περίπτωση των ψηφιακών βιβλιοθηκών, οι φάσεις διαχείρισης της πληροφορίας).

## ΒΙΒΛΙΟΓΡΑΦΙΑ- ΠΗΓΕΣ

- [1] <http://en.wikipedia.org/wiki/Multimedia>
- [2] <http://people.csail.mit.edu/brussell/research/AIM-2005-025-new.pdf>
- [3] <http://dewey.at.northwestern.edu/ppad2/documents/help/image.html>
- [4] <http://www.w3.org/2001/sw/BestPractices/MM/resources/Tools.html>
- [5] <http://dret.net/biblio/reference/wie05>
- [6] <http://www.mindswap.org/2003/PhotoStuff/UserGuide.pdf>
- [7] <http://www.download3k.com/Press-FotoTagger-a-new-personal-tool-for-photo.html>
- [8] <http://www.cogitum.com/>
- [9] <http://www.download3k.com/Press-The-FotoTagger-Image-Annotation-Tool-Now-Supports.html>
- [10] Siegfried Handschuh and Steffen Staab. Cream - creating metadata for the semantic web. *Computer Networks*, 42:579–598, AUG 2003. Elsevier.
- [11] <http://www.acemedia.org/aceMedia/files/document/kes06-m-ontomat-paper.pdf>
- [12] <http://mklab.iti.gr/m-onto2>
- [13] <http://annotation.semanticweb.org/ontomat/index.html>
- [14] <http://www.w3.org/2001/sw/BestPractices/MM/resources/Tools.html>
- [15] <http://annotation.semanticweb.org/Members/lago/AnnotationTool.2003-08-25.1258>
- [16] <http://www.w3.org/2001/Annotea/>
- [17] <http://mklab.iti.gr/via/>
- [18] <http://www.springerlink.com/content/76x9j562653k0378/fulltext.pdf>
- [19] <http://labelme.csail.mit.edu/LabelMe3D/labelme3D.html>
- [20] <http://hcil.cs.umd.edu/trs/2005-04/2005-04.html>
- [21] <http://gregelin.net/bio/resume.html>
- [22] <http://www.mediawiki.org/wiki/Extension:Fotonotes>
- [23] <http://drupal.org/project/fotonotes>
- [24] <http://code.google.com/p/jfotonotes/>
- [25] <http://ccnmtl.columbia.edu/projects/iat/>