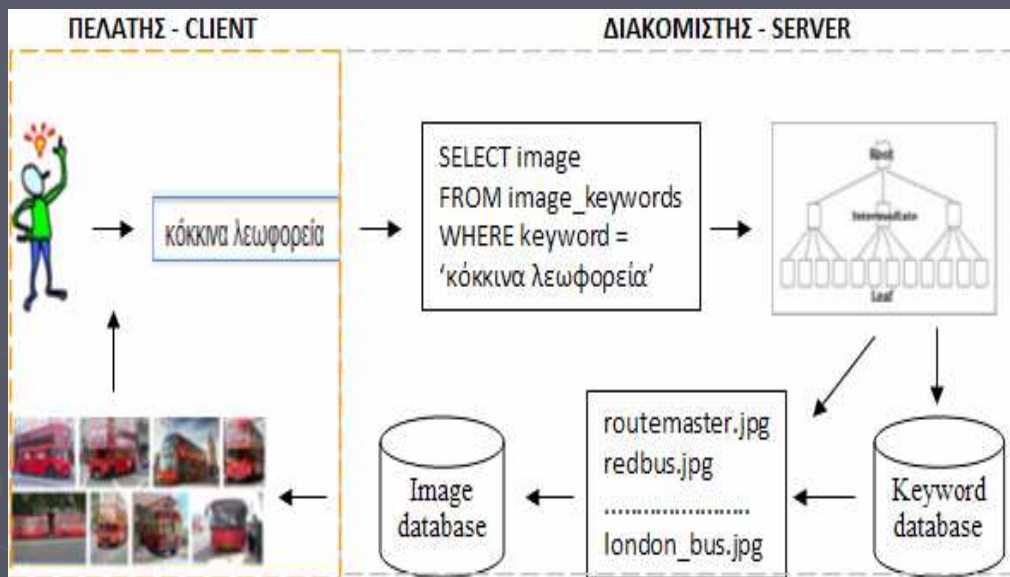


Πανεπιστήμιο Μακεδονίας
Π.Μ.Σ Εφαρμοσμένης Πληροφορικής

Σύγκριση Μεθόδων Ανάκτησης Εικόνας Βασισμένης στο Περιεχομένο με Παράλληλη Υλοποίηση σε Java

Διπλωματική εργασία
Καραφωτιάς Γιώργος 11/20

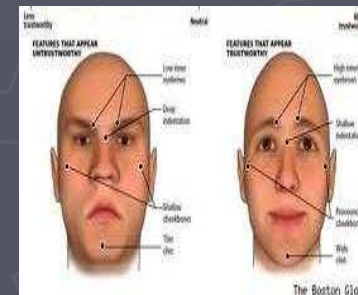
Αναζήτηση με λέξεις κλειδιά



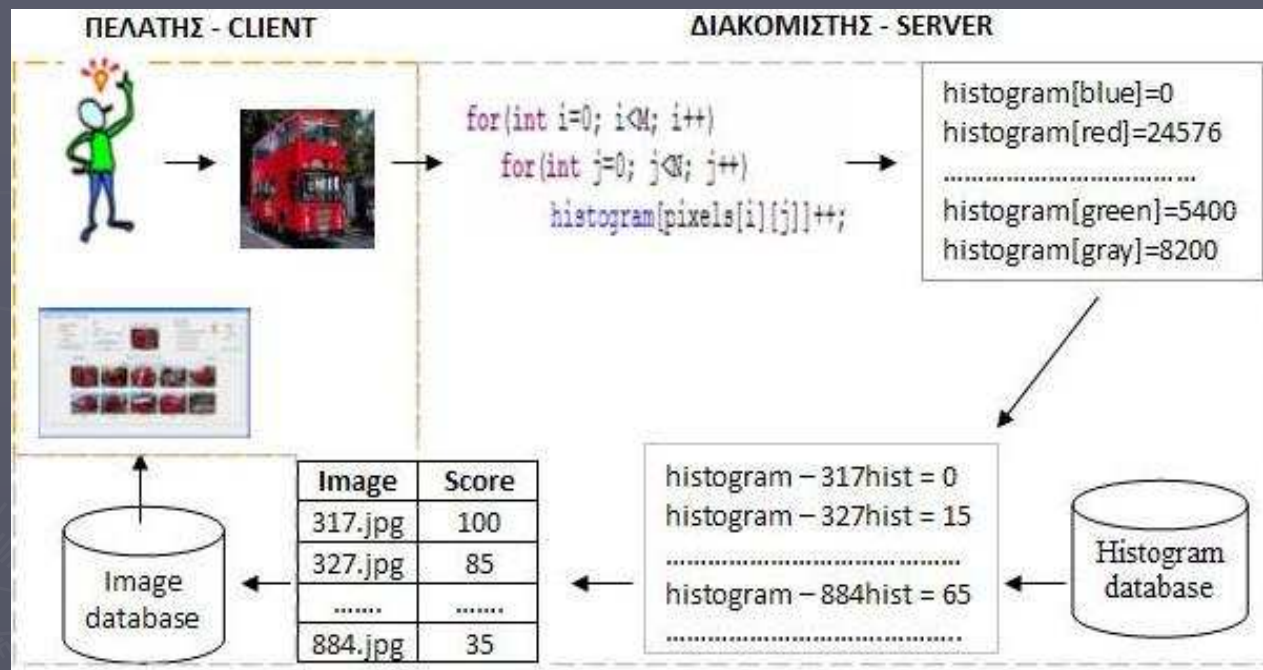
Ευρετήρια

Βάση
Δεδομένων

- ✓ Υφιστάμενες τεχνικές
- ✗ Σχολιασμός εικόνων
- ✗ Ορολογία



Ανάκτηση βάσει περιεχομένου - CBIR



Διατύπωση
ερωτήματος

Εξαγωγή
περιγραφέα

Αποθήκευση
περιγραφέα

Σύγκριση
περιγραφών

Περιγραφείς περιεχομένου

Χαρακτηριστικά

- ▶ Χρώμα
- ▶ Υφή
- ▶ Σχήμα

Ιδιότητες

- ▶ Ακρίβεια
- ▶ Υπολογιστικό κόστος
- ▶ Μέγεθος
- ▶ Αμεταβλητότητα

Κατηγορίες

- ▶ Global
- ▶ Block-based
- ▶ Regional

Παράμετροι επιλογής

- ▶ Όλες οι ιδιότητες
- ▶ Τύπος εικόνων

Μέτρηση ομοιότητας (1)



[15, 28, ..., 10]

[12, 98, ..., 10]

Minkowski

$$d_p(H, K) = \left(\sum_m |h_i - k_i|^p \right)^{1/p}$$

L1 για $p=1$

L2 για $p=2$

Τομή ιστογραμμάτων

$$d_{HI} = 1 - \sum_m \min(h_i, k_i)$$

Εσωτερικό γινόμενο

$$d_{IP} = 1 - H \bullet K = 1 - \sum_m h_i k_i$$

Μέτρηση ομοιότητας (2)

Γεωμετρικός μέσος

$$d_{GM} = 1 - \sum_m \sqrt{h_i k_i}$$

Οικογένεια χ^2

$$d_{sq\chi^2} = \sum_{i=1}^n \frac{(h_i - k_i)^2}{h_i + k_i}$$

Θεωρία πληροφορίας

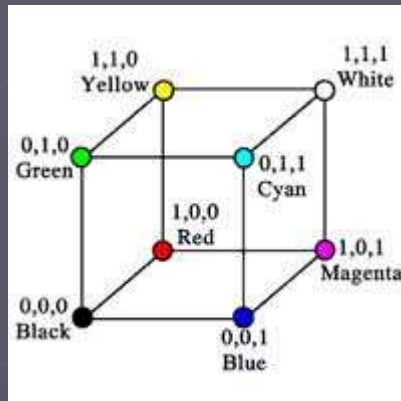
$$d_{KL}(H, K) = \sum_m h_i \log \frac{h_i}{k_i}$$

- Cross – bin συναρτήσεις απόστασης
- Γραφήματα (Graph Matching)
- Γραμμικός προγραμματισμός (Earth Mover's Distance)

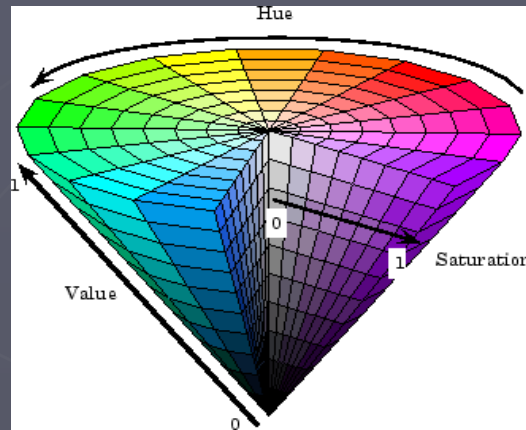
Χρώμα

Κβαντοποίηση

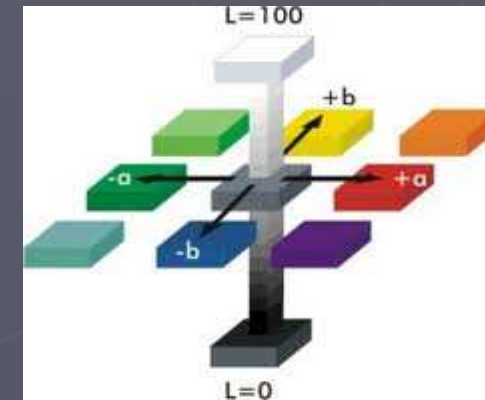
RGB



HSV



Cie Lab

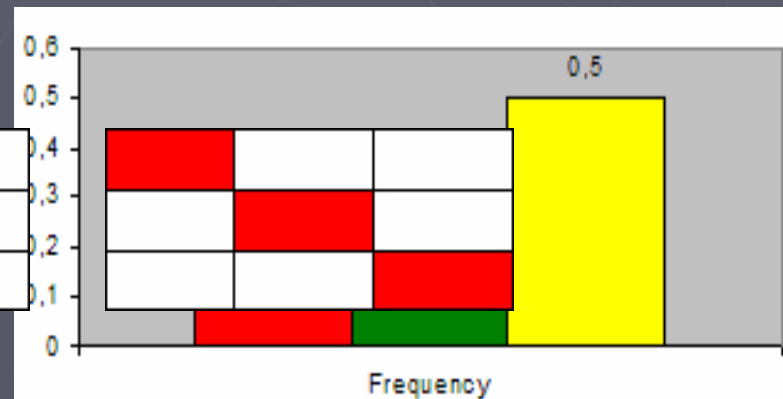
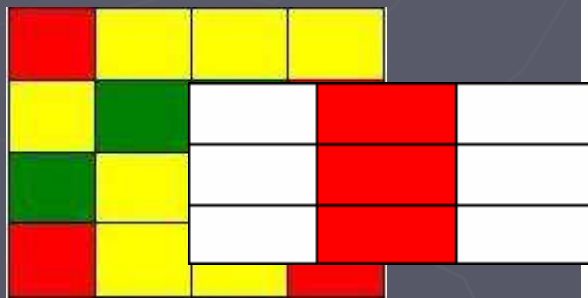


Χώρος	Uniform	Perceptual	Παρατηρήσεις
RGB	✗	✗	+ Προκαθορισμένος για τις ψηφιακές εικόνες + Συμβατότητα με μέσα απεικόνισης (οθόνες)
HSV	✗	✓	+ Γραμμικός μετασχηματισμός από RGB
CIE Lab	✓	✓	- Μη γραμμικός μετασχηματισμός από RGB

- 1) Ιστόγραμμα
- 2) Μέσο χρώμα
- 3) Κυρίαρχα χρώματα

Χρωματικό Ιστόγραμμα (GCH)

Σύγκριση της χρωματικής κατανομής 2 εικόνων

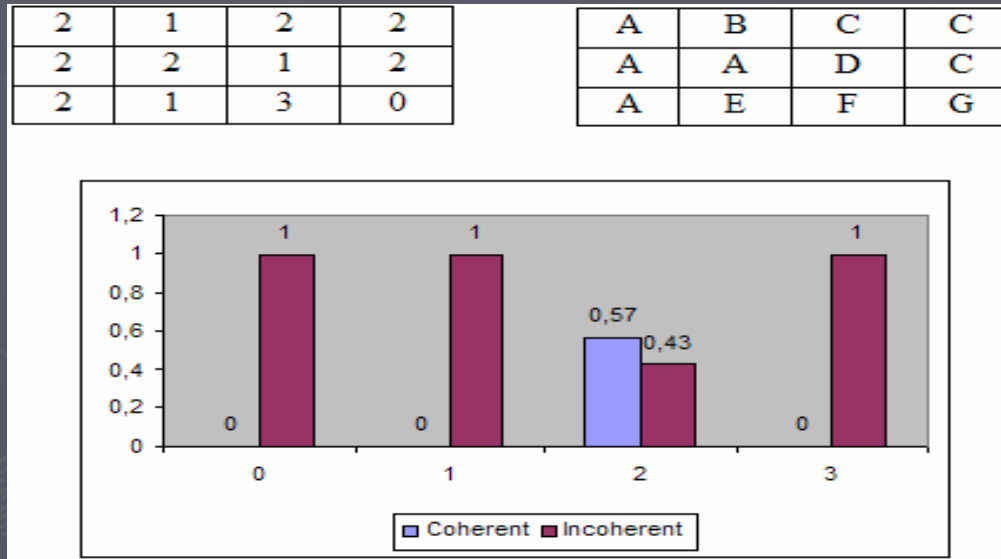


1. Επιλογή χρωματικού χώρου
2. Κβαντοποίηση εικόνας
3. Υπολογισμός ιστογράμματος
4. Σύγκριση ιστογραμμάτων

- ✓ Απλότητα υπολογισμού
- ✓ Ταχύτητα υπολογισμού
- ✓ Αμετάβλητο στην περιστροφή
- ✗ Χωρική κατανομή των pixel

Χρωματικό διάνυσμα συνοχής (CCV)

Βελτιώνει το ιστόγραμμα χωρίζοντας τα pixel σε 2 κατηγορίες

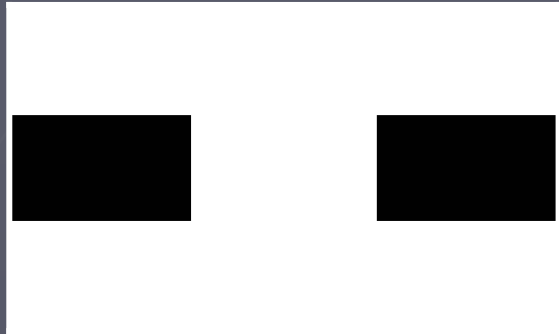


Εικόνα

Συνδεδεμένα
στοιχεία

1. Εύρεση των συνδεδεμένων στοιχείων μιας εικόνας
2. Κατάταξη των pixel κάθε χρώματος σε συνεκτικά και μη-συνεκτικά ανάλογα με τα στοιχεία στα οποία ανήκουν

Χρωματικό συσχετογράφημα (ACC)



$$|p_1 - p_2| = \max(|x_1 - x_2|, |y_1 - y_2|)$$

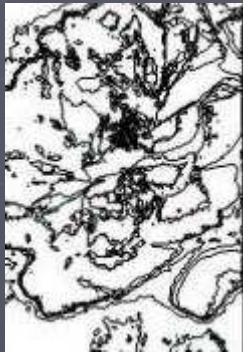
$$\gamma_{c_i, c_j}^{(k)}(\mathcal{I}) \triangleq \Pr_{p_1 \in \mathcal{I}_{c_i}, p_2 \in \mathcal{I}} [p_2 \in \mathcal{I}_{c_j} \mid |p_1 - p_2| = k]$$

Η πιθανότητα ένα pixel χρώματος c_i να βρίσκεται σε απόσταση k από ένα pixel χρώματος c_j

- ▶ Πρόβλημα: Υπολογιστικό κόστος και αποθηκευτικός χώρος
- ▶ Λύση: Χρωματικό αυτοσυσχετογράφημα
- ▶ Ενδιαφέρουσα παραλλαγή: Συσχετογράφημα διαμέσου

BIC

- ▶ Παρόμοιος σε λογική με το CCV
- ▶ Κατατάσσει τα ρίχελ σε συνοριακά και εσωτερικά
- ▶ Απλούστερος υπολογιστικά από το CCV



$$dLog(q, d) = \sum_{i=0}^{i < M} |f(q[i]) - f(d[i])|$$

$$f(x) = \begin{cases} 0, & \text{if } x = 0 \\ 1, & \text{if } 0 < x \leq 1 \\ \lceil \log_2 x \rceil + 1, & \text{otherwise} \end{cases}$$

Βελτιστοποίηση για τη μείωση του αποθηκευτικού χώρου

Τοπικό Ιστόγραμμα (LCH)

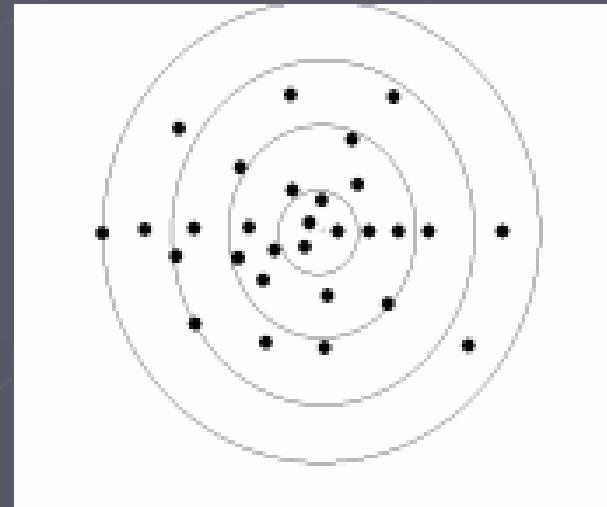
- ▶ Block-based περιγραφείας
- ▶ Η εικόνα χωρίζεται σε μπλοκ και εξάγεται το ιστόγραμμα του καθενός από αυτά
- ✘ Αποθηκευτικό κόστος
- ✘ Είναι ευαίσθητο σε περιστροφές της εικόνας



Ευελιξία προσαρμογών στον τρόπο χωρισμού σε μπλοκ

Γεωμετρικό ιστόγραμμα (ANNHIST)

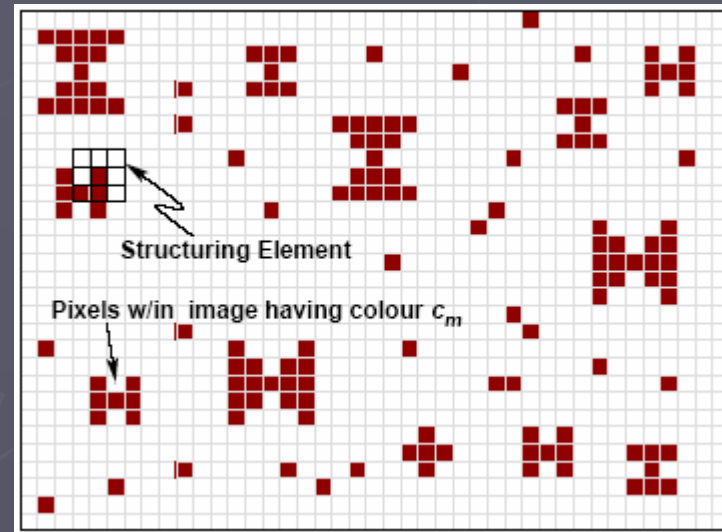
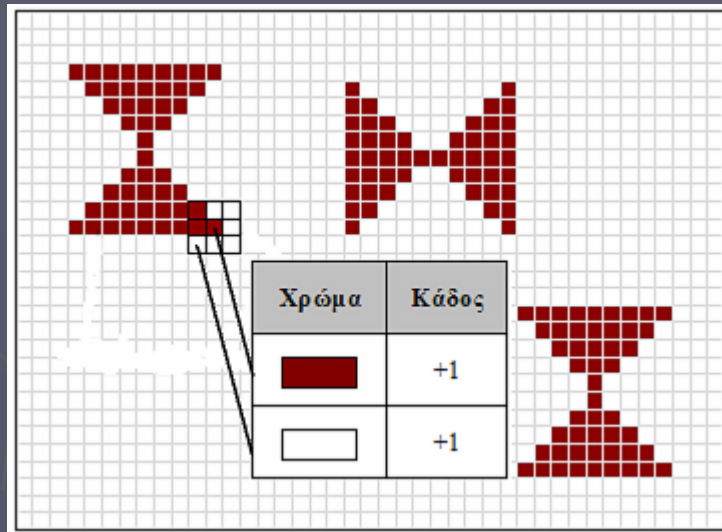
- ▶ Διατηρεί πληροφορία για τη θέση των pixel κάθε χρώματος
- ▶ Τα pixel είναι σημεία στο διδιάστατο επίπεδο



$$r^q = \max_{(x,y) \in S_q} \sqrt{(x - x^q)^2 + (y - y^q)^2}$$

Μετράμε τα pixel που βρίσκονται σε κάθε μία περιοχή

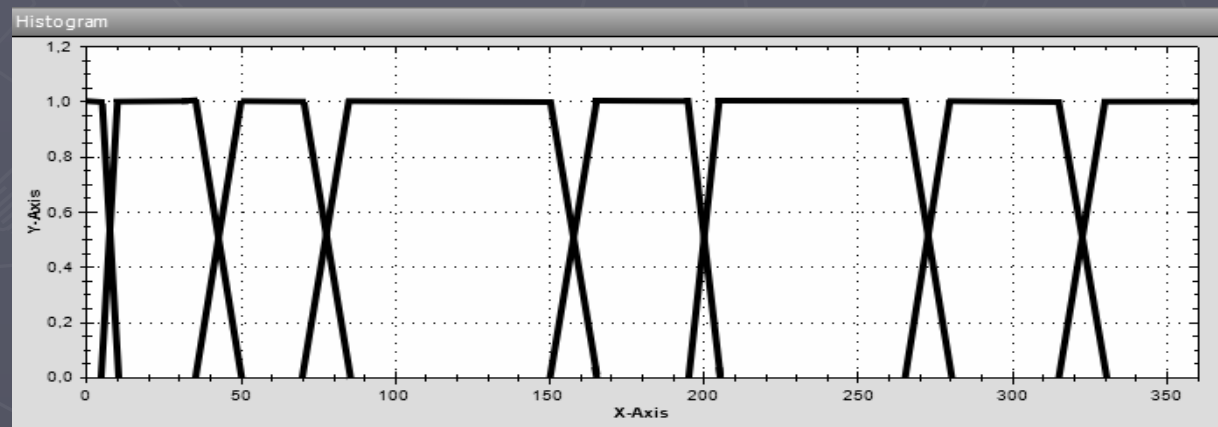
Περιγραφέας δομής χρώματος (CSD)



- ▶ Περιγραφέας του προτύπου MPEG-7
- ▶ Ο υπολογισμός του γίνεται στο χρωματικό χώρο HMMD
- ▶ Διαφορετική ερμηνεία τιμών του ιστογράμματος

Ασαφές Ιστογράμμο (FUZZYHIST)

- ▶ Με τη χρήση ασαφούς λογικής κάθε pixel της εικόνας κβαντοποιείται σε 1 από 10 πιθανά χρώματα
- ▶ Δημιουργείται ένα ιστογράμμο με 10 μόνο τιμές



Ασαφής κανόνας: Αν η συνιστώσα Hue είναι μπλε και η φωτεινότητα 70% τότε το χρώμα είναι γαλάζιο

Χρωματικές ροπές (CM)

- ▶ Τιμές χρώματος των pixel
= στατιστική κατανομή
- ▶ Μια κατανομή μπορεί να
χαρακτηριστεί από μια
σειρά τιμών (μέσος,
διακύμανση)
- ✓ Συμπαγής περιγραφέας
- ✗ Διακριτική ικανότητα

$$E_i = \frac{1}{N} \sum_{j=1}^N p_{ij}$$

$$\sigma_i = \sqrt{\left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^2 \right)}$$

$$s_i = \sqrt[3]{\left(\frac{1}{N} \sum_{j=1}^N (p_{ij} - E_i)^3 \right)}$$

Υφή (Texture)

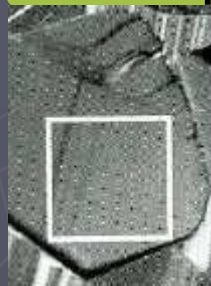
- ▶ Έλλειψη ξεκάθαρα ορισμού
- ▶ Επαναλαμβανόμενα μοτίβα
- ▶ Δε σχετίζεται με το χρώμα



Προσεγγίσεις

- ▶ Στατιστικοί
- ▶ Πεδίο συχνοτήτων
- ▶ Συντακτικοί

Παράδειγμα



11,79



74,63

Τυπική απόκλιση

Πίνακες σύμπτωσης (GLCM)

- ▶ Η ίδια ιδέα με το συσχετογράφημα για εικόνες του γκρι

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Αρχική εικόνα

4/7	2/7	1/7	0
2/7	4/7	0	0
1/8	0	6/8	1/8
0	0	1/3	2/3

Πίνακας σύμπτωσης 0 μοίρες

4/5	1/5	0	0
1/5	2/5	2/5	0
0	2/7	4/7	1/7
0	0	1	0

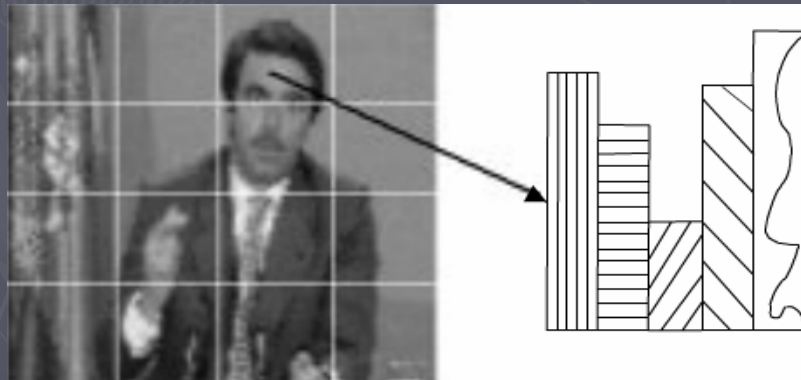
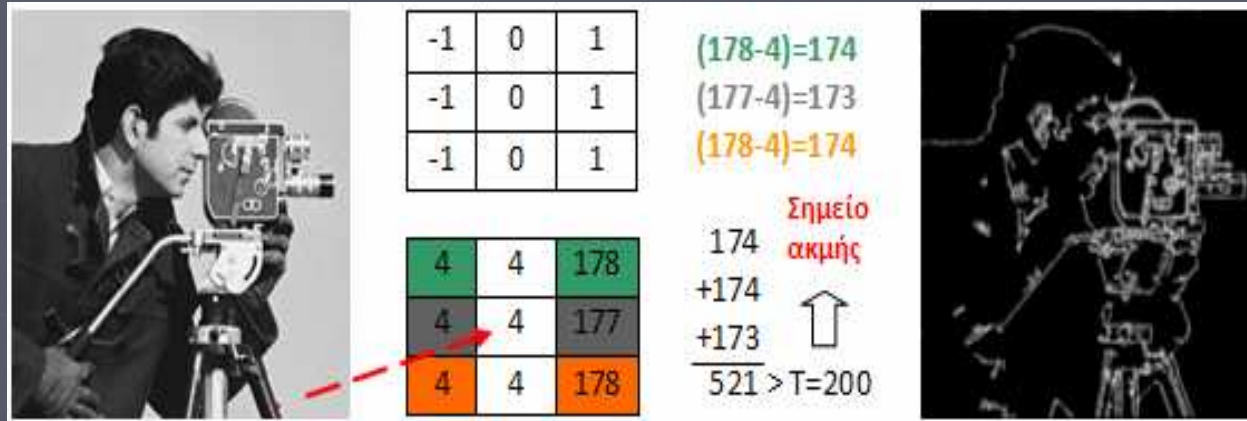
Πίνακας σύμπτωσης 45 μοίρες

- 1) Υπολογισμός πινάκων σύμπτωσης
- 2) Υπολογισμός στατιστικών μέτρων από αυτούς

$$\sum_i \sum_j p(i, j)^2$$

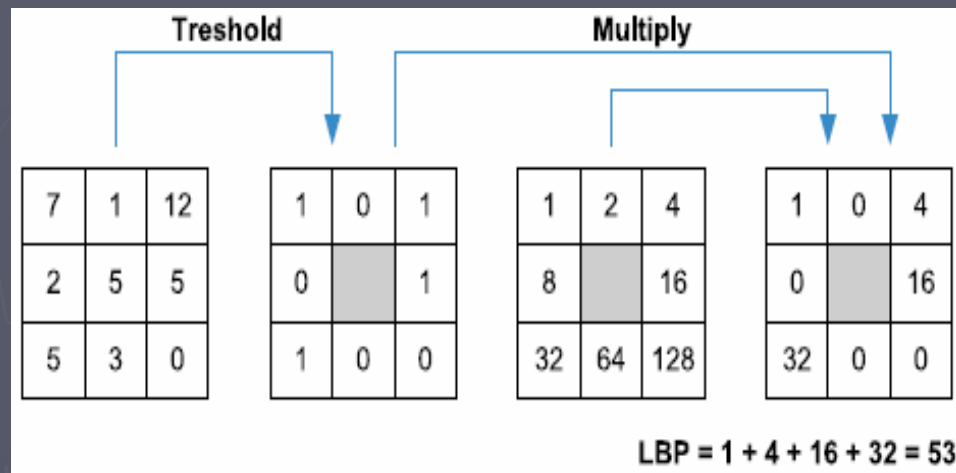
$$\sum_{n=0}^{N-1} n^2 \sum_{i=1}^N \sum_{j=1}^N p(i, j), \quad |i - j| = n$$

Ιστογράμμα ακμών (EHD)

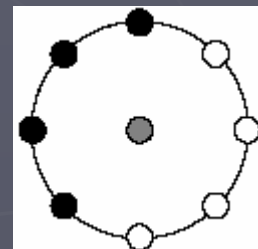


- ▶ Η εικόνα χωρίζεται σε 16 υποεικόνες
- ▶ Για κάθε υποεικόνα μετράμε το πλήθος καθενός από 5 είδη ακμών → Ιστογράμμα 80 τιμών

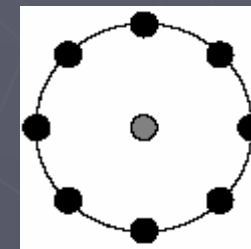
Τοπικά δυαδικά πρότυπα (LBP)



Ακμή



Σημείο



Μικροπρότυπα υφής
σε μια εικόνα

- ▶ Κυκλικές γειτονιές διαφορετικού μεγέθους
- ▶ Ομοιόμορφα πρότυπα
- ▶ Προσθήκη αμεταβλητότητας με μετασχηματισμούς
- ▶ Πρόβλημα: Αλλαγή μεγέθους εικόνας

Σύγκριση περιγραφέων - Πλαίσιο

Πυλώνες

- ▶ Ακρίβεια
- ▶ Κόστος

Ακρίβεια

Μέτρα αξιολόγησης

Benchmark συλλογές

Κόστος

Χώρος

Χρόνος

Εργαλείο

- ▶ Σύστημα java

Κλιμάκωση

- ▶ Παραλληλοποίηση – Open MP

Σύστημα CBIR - Java

JPSim

Search Indexing Performance

Similarity Mode

- Single
- Hierarchical
- Parallel

Restore Default

Filter


Score >

Best

Open an image ...

Run

Dataset Wang



Descriptors











- Histogram
- Color Coherence Vector
- AutoCorrelogram
- Local Histogram
- Edge Histogram

Category

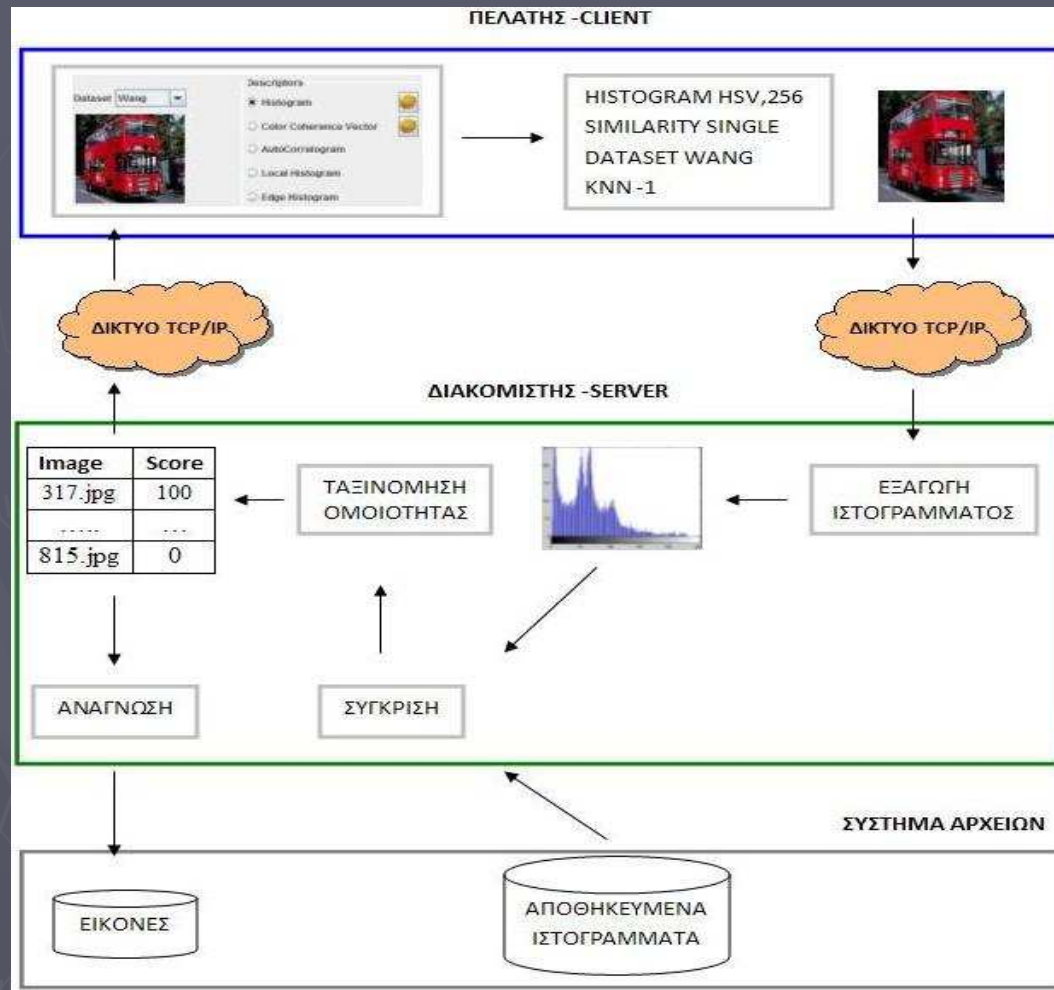
- Color
- Texture
- All

Previous Elapsed Time = 0.094 seconds Next

SINGLE: 1. Histogram L1,HSV,256

				
100	61,34	60,91	60,43	60,38
				
59,5	59,23	59,2	58,97	57,32

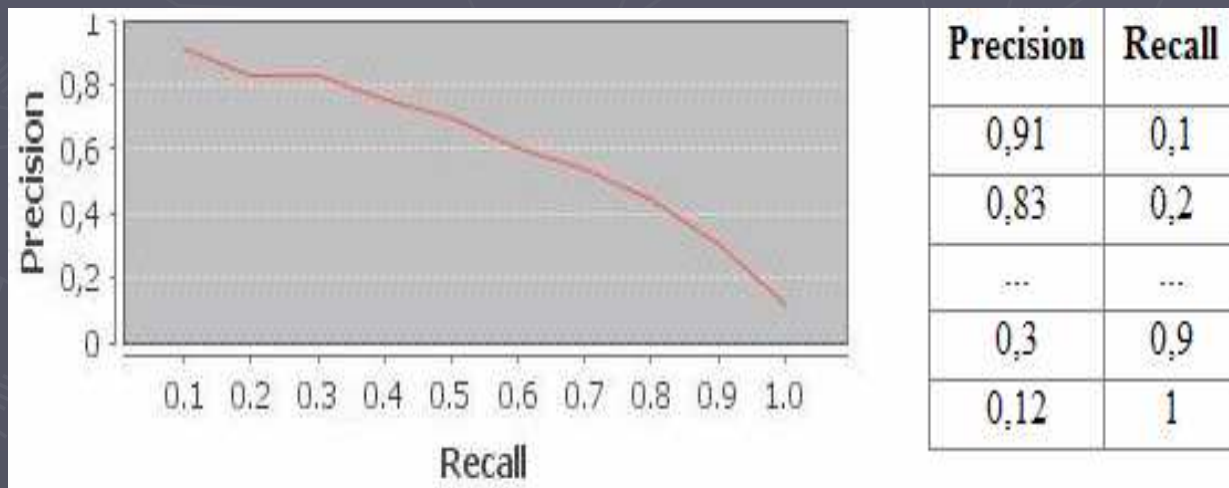
Σκαρίφημα συστήματος



- ▶ Client – server ή Desktop Based
- ▶ Πρωτόκολλο Sockets
- ▶ Αποθήκευση των περιγραφών σε δυαδικά αρχεία
- ▶ Αυτοματοποίηση πειραμάτων

Μέτρα ακρίβειας

- ▶ Γραφήματα ακρίβειας – ανάκλησης
- ▶ Mean Average Precision (MAP)
- ▶ Προϋπόθεση: groundtruth file
- ▶ Συλλογές: WANG, UCID, Brodatz

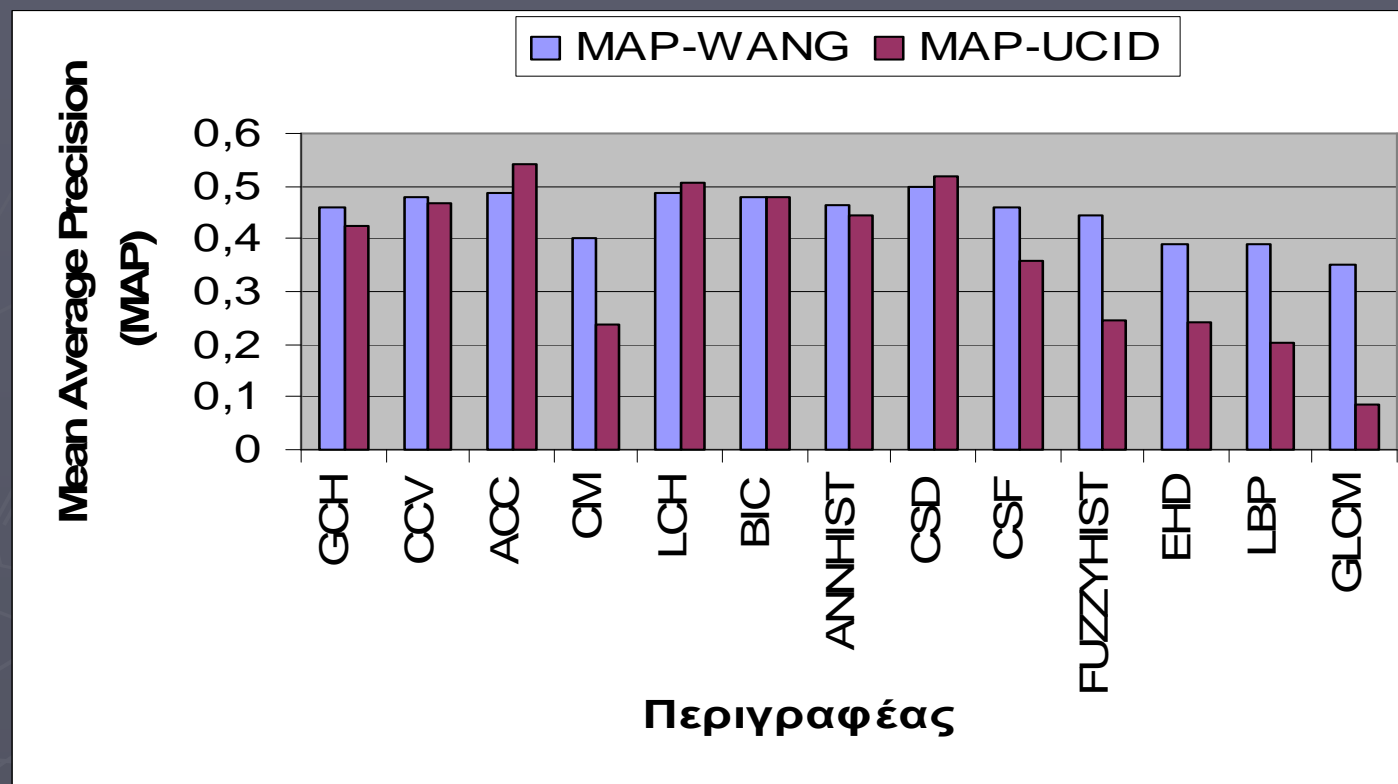


Mean Average Precision (MAP)

ucid00001.tif: ucid00015.tif ucid00038.tif ucid00240.tif	ucid00001.tif: ucid00015.tif ucid00017.tif ucid00121.tif ucid00038.tif ucid00240.tif	$N_R = 3, q = \text{ucid00001.tif}$ <table border="1"><thead><tr><th>n</th><th>R_n</th><th>$P_q(R_n)$</th></tr></thead><tbody><tr><td>1</td><td>20</td><td>$1/1 = 1$</td></tr><tr><td>2</td><td>80</td><td>$2/4 = 0,5$</td></tr><tr><td>3</td><td>100</td><td>$3/5 = 0,6$</td></tr></tbody></table> $AP = (1+0,5+0,6)/3 = 0,7$	n	R_n	$P_q(R_n)$	1	20	$1/1 = 1$	2	80	$2/4 = 0,5$	3	100	$3/5 = 0,6$
n	R_n	$P_q(R_n)$												
1	20	$1/1 = 1$												
2	80	$2/4 = 0,5$												
3	100	$3/5 = 0,6$												
Σωστή απάντηση	Πραγματική απάντηση													

$$AP(q) = \frac{1}{N_R} \sum_{n=1}^{N_R} P_q(R_n) \quad MAP = \frac{1}{|Q|} \sum_{q \in Q} AP(q)$$

MAP – Σταθερή κβαντοποίηση

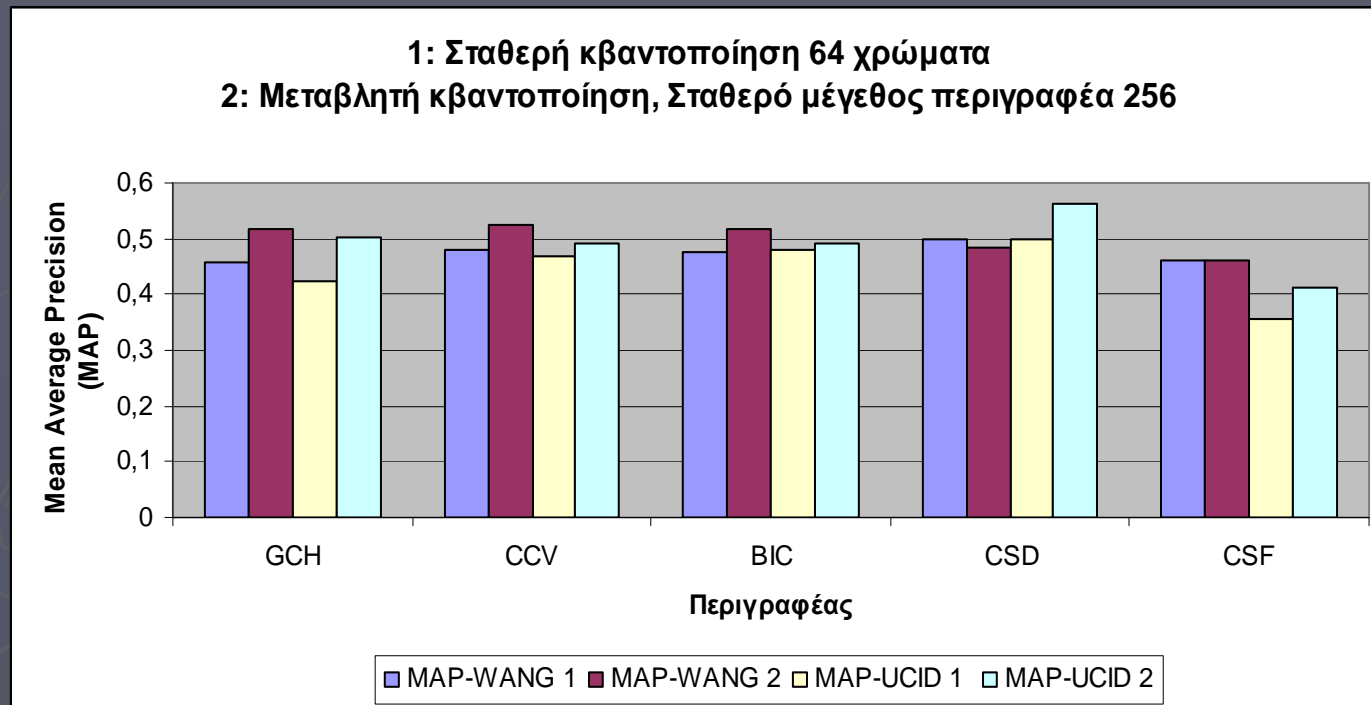


1. CSD

2. ACC

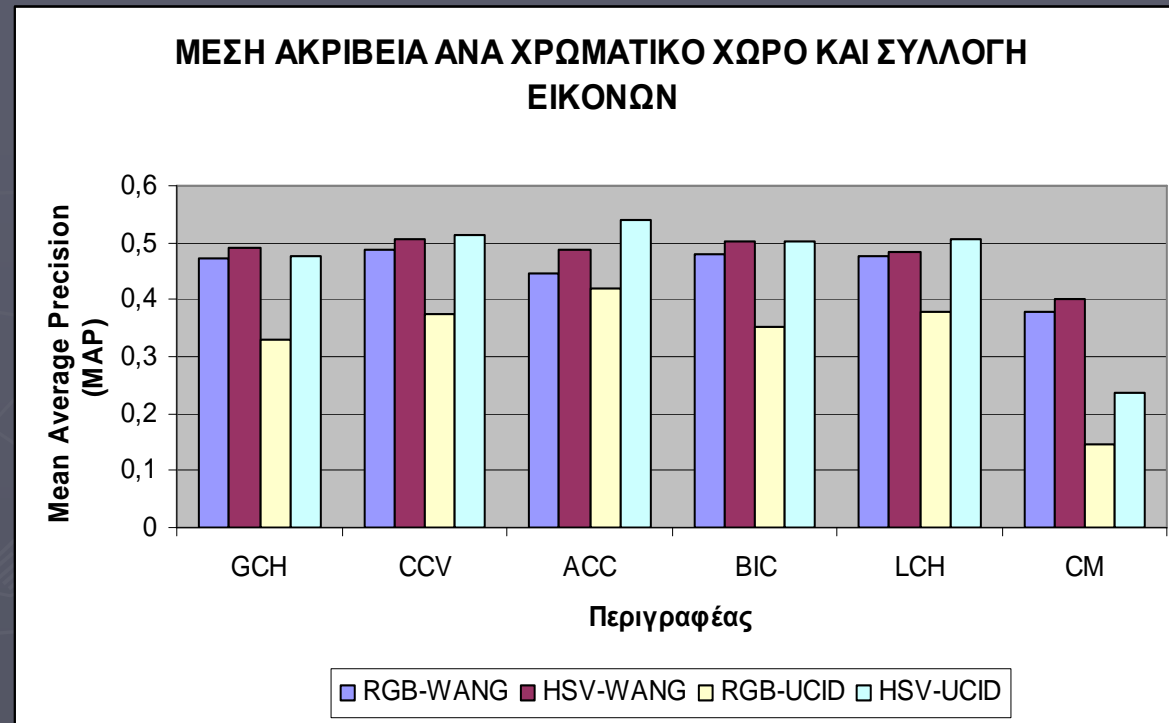
3. LCH

MAP – Σταθερό μέγεθος περιγραφέα



Μεγαλύτερα περιθώρια βελτίωσης στο ιστόγραμμα

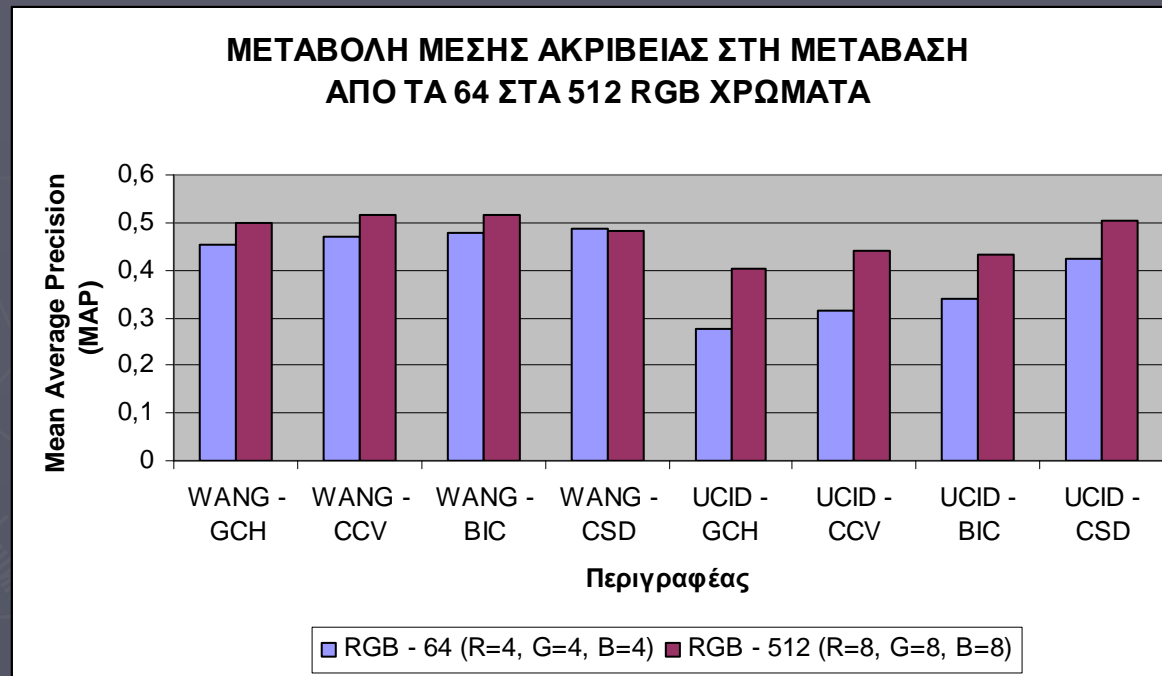
MAP – Αλλάζοντας χρωματικό χώρο



Ο HSV καλύτερος του RGB

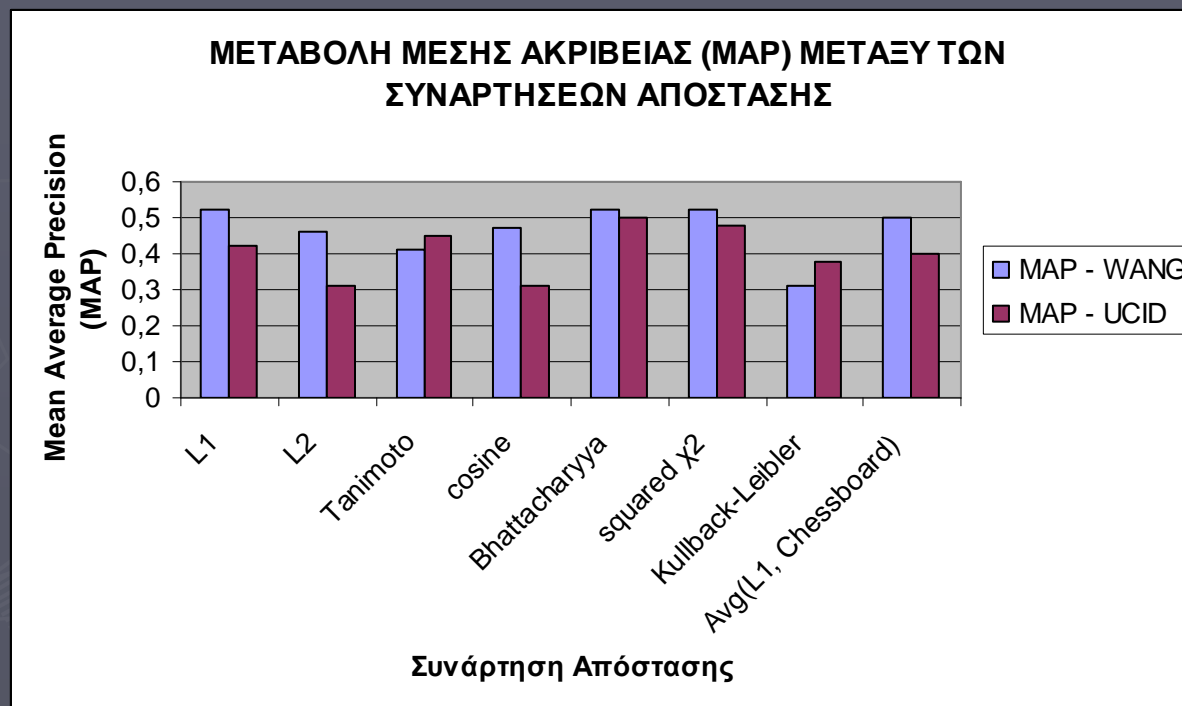
Ο CieLab αργός

Αλλάζοντας σχήμα κβαντοποίησης



Αύξηση της ακρίβειας με ταυτόχρονη αύξηση του κόστους

Αλλάζοντας τη συνάρτηση απόστασης

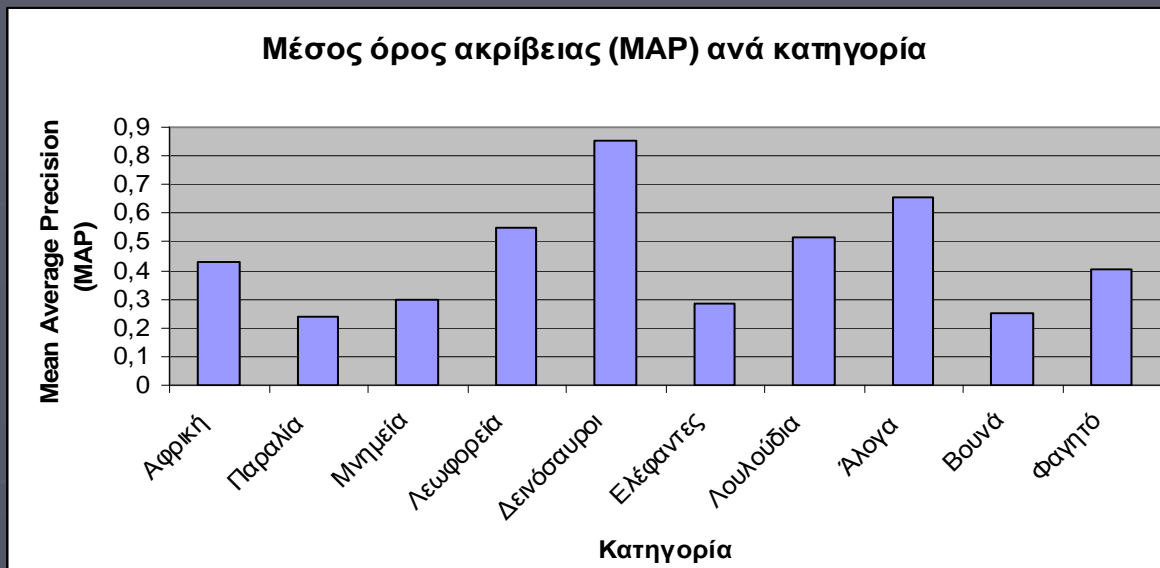


1. Bhattacharyya

2. Squared χ^2

3. L1

Ακρίβεια ανά κατηγορία εικόνων



LBP, EHD



FUZZYHIST, CSF



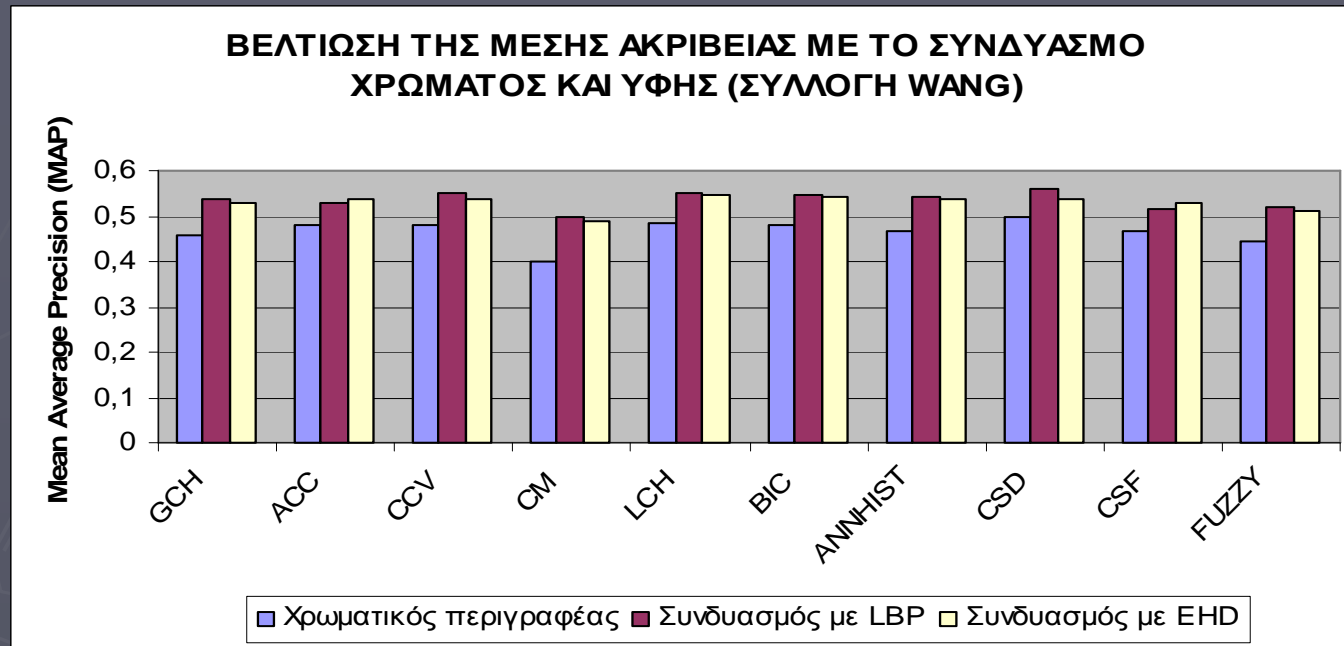
LBP, ACC



CM, FUZZYHIST (!)



Συνδυασμός χρώματος-υφής



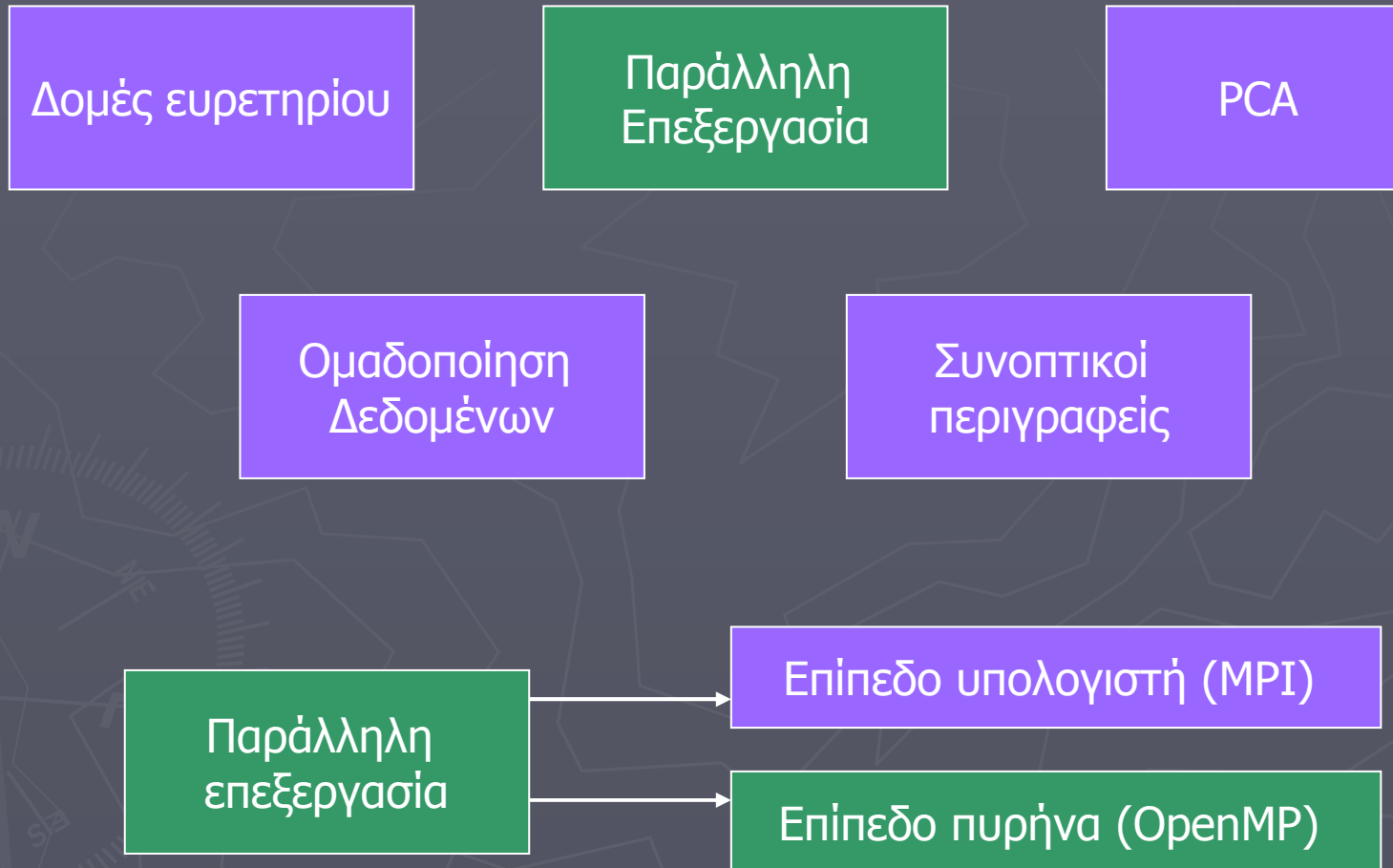
- ▶ Βελτίωση της ακρίβειας
- ▶ Αύξηση του κόστους (Δυνατότητα παραλληλισμού)
- ▶ **Πρόβλημα:** Κανονικοποίηση

Ταχύτητα υπολογισμού περιγραφών

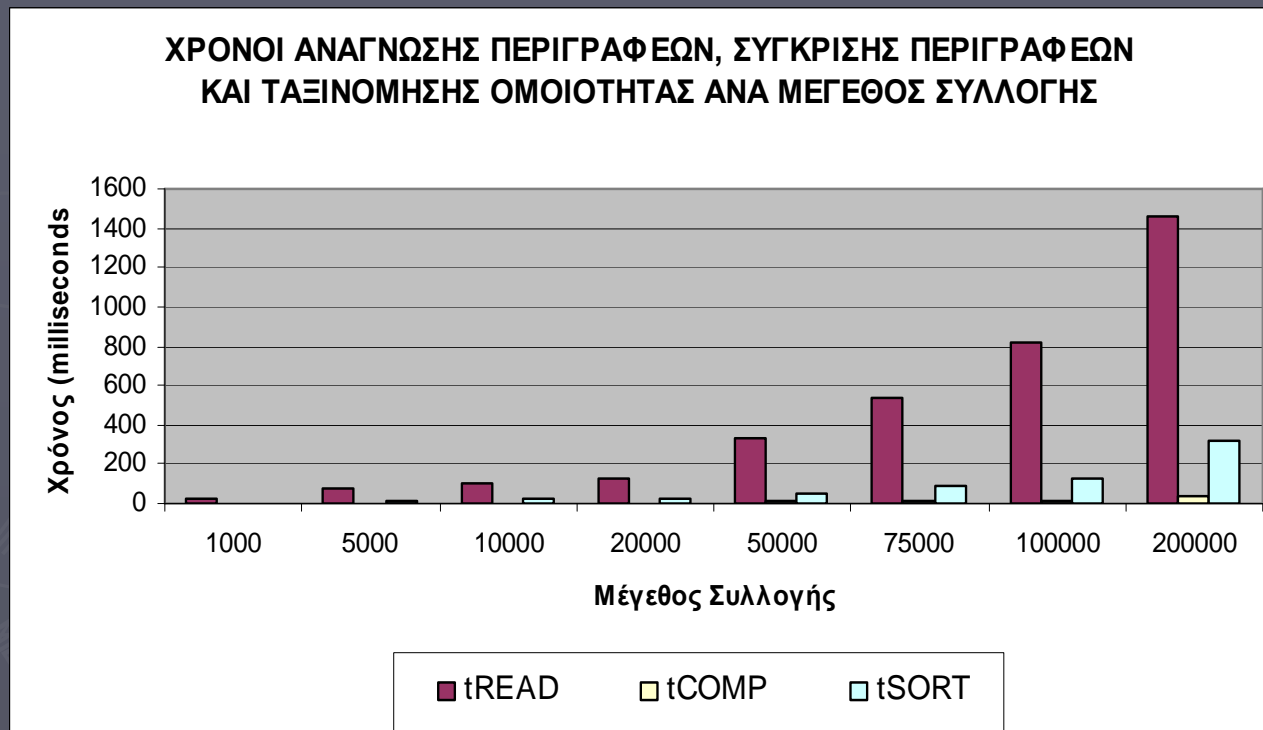


- ▶ 3 γκρουπ περιγραφών
- ▶ Ταύτιση με τις θεωρητικές πολυπλοκότητες

Κλιμάκωση



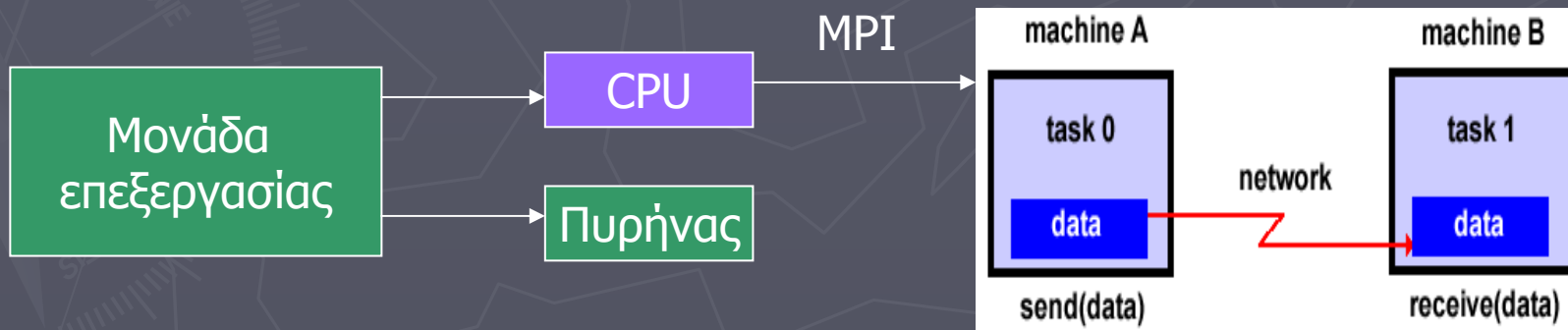
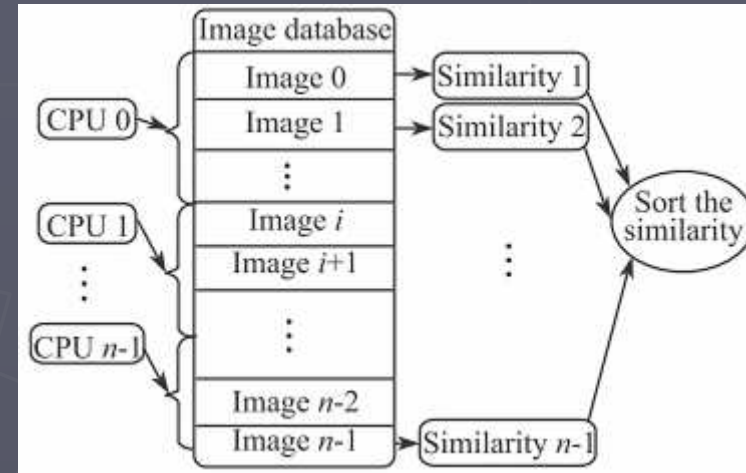
Χρόνος απόκρισης



- ▶ Η επιβάρυνση του χρόνου ανάγνωσης μπορεί να αποφευχθεί
- ▶ Κόστος μεταφοράς δεδομένων στο δίκτυο

Η λογική του παραλληλισμού

- Χωρίζουμε τις εικόνες σε n τμήματα
- Κάθε **επεξεργαστική μονάδα** αναλαμβάνει ένα τμήμα

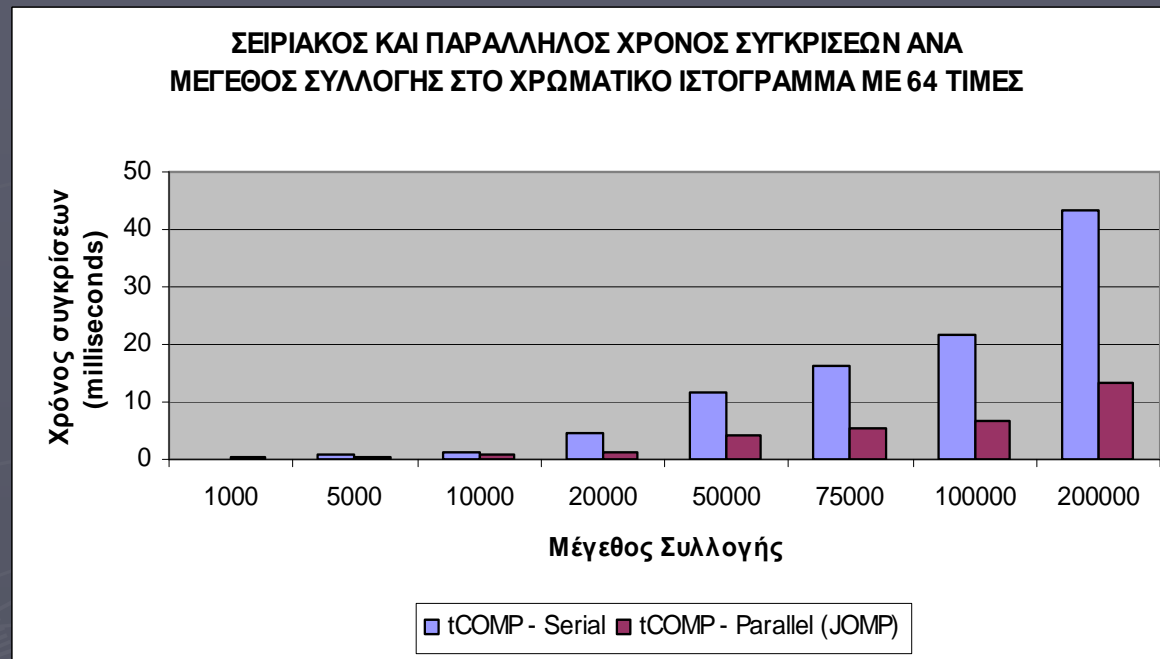


OpenMP – Παράλληλες συγκρίσεις

```
//omp parallel shared(similarities, features,  
feature, datasetSize) private(omp_threads)  
{  
    //omp for schedule(static, datasetSize/omp_threads)  
    for(int i = 0; i < datasetSize; i++)  
        similarities[i] = feature.compare(features[i]);  
}
```

- ✓ Ευκολότερο από τα νήματα Java
- ✓ Φανερά τα σημεία παραλληλισμού
- ▶ Μοντέλο νημάτων
- ▶ Μοιραζόμενη μνήμη
- ▶ JOMP – Υλοποίηση java
- ✘ Ελλιπής υποστήριξη

Σειριακός και παράλληλος χρόνος

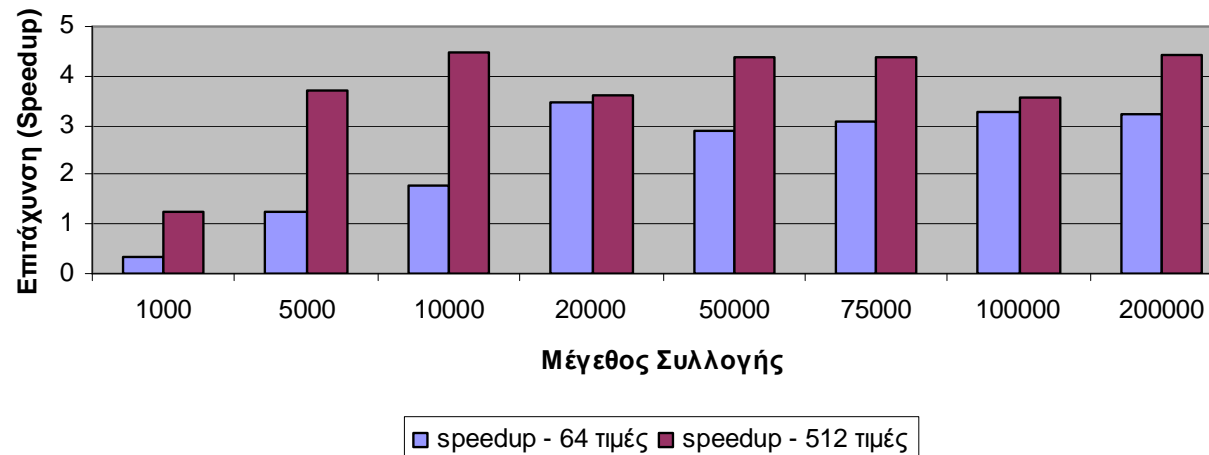


- ▶ Μη-πραγματικές συλλογές
- ▶ 500 δείγματα
- ▶ 4 πυρήνες
- ✓ Επιτάχυνση
- ✓ Μικρός ρυθμός μεταβολής

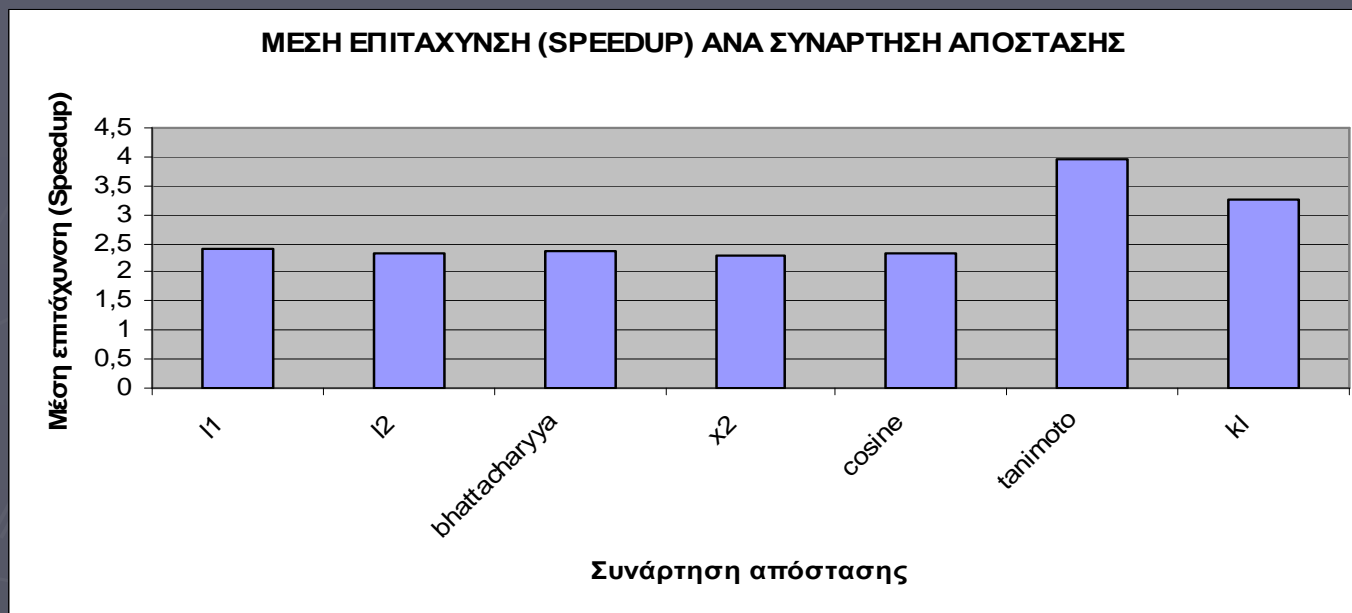
Αυξάνοντας το μέγεθος

Μέγεθος συλλογής	Χώρος αποθήκευσης (Ιστόγραμμα με 64 τιμές)	Χώρος αποθήκευσης (Ιστόγραμμα με 512 τιμές)
100000	25,2 MB	204,8 MB
200000	50,4 MB	409,6 MB

ΕΠΙΤΑΧΥΝΣΗ (SPEEDUP) ΑΝΑ ΜΕΓΕΘΟΣ ΣΥΛΛΟΓΗΣ ΣΤΟ ΧΡΩΜΑΤΙΚΟ ΙΣΤΟΓΡΑΜΜΑ ΜΕ 64 ΚΑΙ 512 ΤΙΜΕΣ ΑΝΤΙΣΤΟΙΧΑ



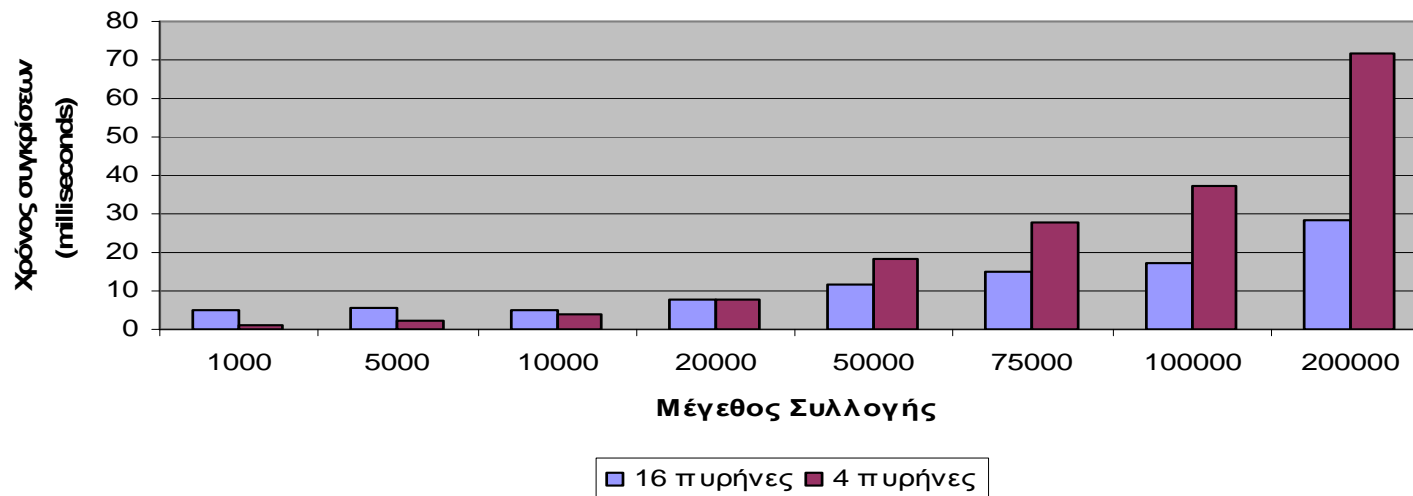
Αλλάζοντας τη συνάρτηση απόστασης



- ▶ Οι διαφορές φαίνονται όταν ξεφύγουμε από τη γραμμική πολυπλοκότητα
- ▶ Απόσταση τετραγωνικού τύπου 50 φορές πιο αργή από την L1

Αυξάνοντας τους πυρήνες

ΧΡΟΝΟΣ ΣΥΓΚΡΙΣΕΩΝ ΑΝΑ ΜΕΓΕΘΟΣ ΣΥΛΛΟΓΗΣ
ΚΑΙ ΠΛΗΘΟΣ ΠΥΡΗΝΩΝ



Μετάβαση	Ποσοστιαία Μεταβολή (4 πυρήνες)	Ποσοστιαία Μεταβολή (16 πυρήνες)
750000-100000	34%	13%
100000-200000	93%	64%

Συμπεράσματα – Ανάκτηση εικόνας

- ▶ Ευρύ πεδίο

Περιγραφείς

Αξιολόγηση
απόδοσης

Κάλυψη
σημασιολογικού
κενού

- ▶ Άνω όριο
περιγραφέων

Ακρίβεια 50%

- ✓ Χρώμα και υφή μαζί
- ✓ Σύνθετοι περιγραφείς
- ✓ Μηχανική μάθηση

Εφαρμογές ?

Κλιμάκωση

- ▶ Απαραίτητος και χρήσιμος ο παραλληλισμός

Τέλος Παρουσίασης

