

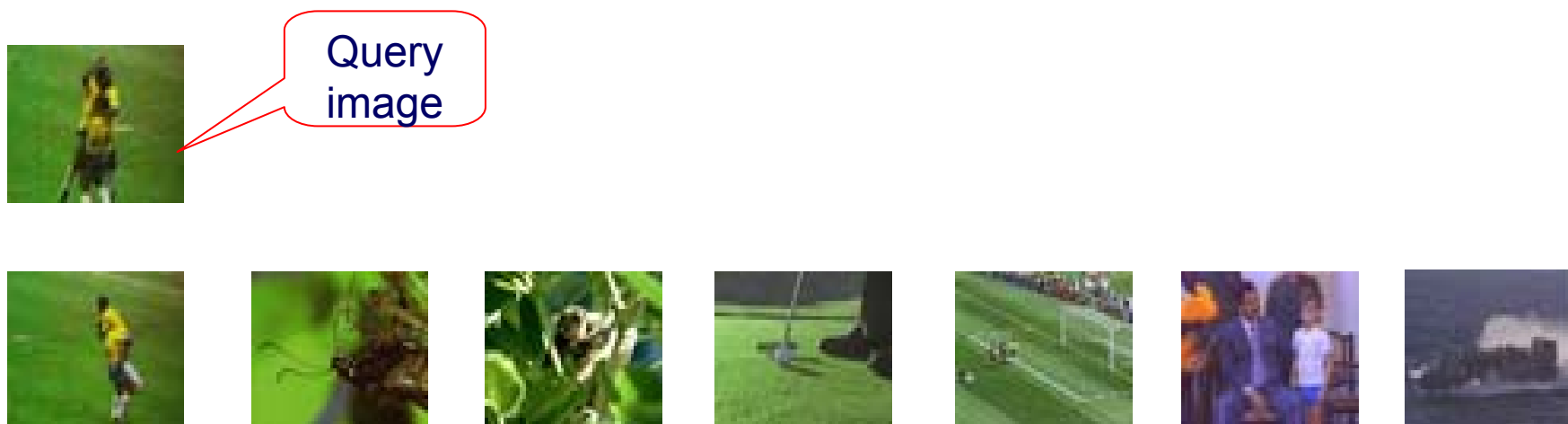
MPEG7

Multimedia Content Description Interface

Αναφέρεται σε «βάσεις» εικόνων, video και ήχου



Πως ένα σύνολο εικόνων θα μπει σε μία διάταξη κατά σειρά ομοιότητας;



Οι 7 εικόνες στη 2^η γραμμή έχουν τοποθετηθεί σύμφωνα με την σειρά ομοιότητας με την εικόνα αναζήτησης (1^η γραμμή)

Content-based image Retrieval

- ανάκτηση βάσει του περιεχομένου

Τρόποι:

- Με σχεδιάγραμμα (query by sketch)
- Με παράδειγμα μία εικόνα (query by example)
- Με σύνολο εικόνων (query by group example)
- Με «σχετική» ανατροφοδότηση (relevance feedback) (θετικά και αρνητικά παραδείγματα σε αλληλεπίδραση με τον χρήστη)
- Με «semantics»



1. Επιλογή χαρακτηριστικών
2. Μέτρα ομοιότητας

Το πρωτόκολλο MPEG7

Εισαγωγικά

Τυποποιεί την περιγραφή του περιεχομένου των πολυμέσων
(video – audio)

Δεν επεξεργάζεται αλλά....

- Συλλέγει χαρακτηριστικά που χρειάζονται για περιγραφή δεδομένων εικόνας-ήχου
- Κάνει δόμηση της περιγραφής για δημιουργία «κοινής γλώσσας» ώστε να..
- πραγματοποιείται indexing – retrieval – browsing κλπ.

Διαδικασίες περιγραφής (Descriptors)

Οπτικές (Visual)

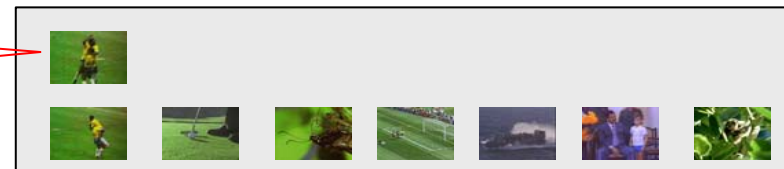
- ✓ Περιγραφή χρώματος
- Περιγραφή Υφής
- Περιγραφή Σχήματος
- Περιγραφή «κίνησης»
- Περιγραφή προσώπου (με Βασικό εργαλείο: PCA)

Ακουστικές (Audio)

- Ταξινόμηση ήχου και ομοιότητα
- Περιεχόμενο προφορικού λόγου

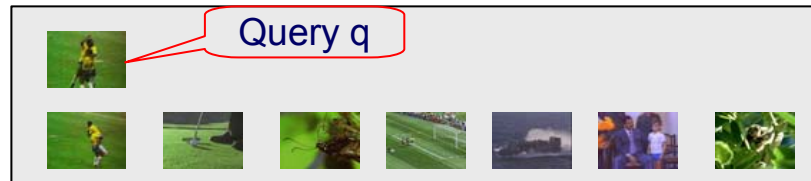
Ποσοτική εκτίμηση της περιγραφής

Query q



- ❑ Βασίζεται στη διαδικασία «retrieval - ανάκτησης»
(Query-by-example QBE)
- Λαμβάνει υπόψη:
- Αν οι όμοιες εικόνες είναι πρώτες στη σειρά ομοιότητας
- Αν παρουσιάζονται στην ανάκτηση «λάθος» εικόνες
- Αν παραλείπονται εικόνες που είναι όμοιες με την εικόνα αναζήτησης

Ποσοτική εκτίμηση της περιγραφής (συνέχεια)



γενικά

- έχουμε δύο εκτιμητές: **precision** και **recall**

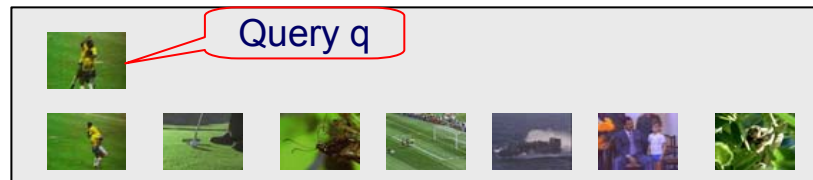
$$\text{precision} = \frac{|Q(q) \cap R(q)|}{|Q(q)|} \quad \text{recall} = \frac{|Q(q) \cap R(q)|}{|R(q)|}$$

Όπου: **q** είναι η εικόνα αναζήτησης (query)

Q(q) ο αριθμός των εικόνων που βρέθηκαν

και **R(q)** ο πραγματικός αριθμός των ομοίων εικόνων

Ποσοτική εκτίμηση της περιγραφής (συνέχεια)



Για το MPEG7

- Κλασσικός εκτιμητής: Retrieval Rank,

$$RR(q) = \frac{NF(a, q)}{NG(q)}$$

Όπου: **q** είναι η εικόνα αναζήτησης (query)

NF(a,q) ο αριθμός των «σωστών» εικόνων που βρέθηκαν σε **aNG(q)** επιλογές ($a > 1$)

και **NG(q)** ο πραγματικός αριθμός των ομοίων εικόνων

Ποσοτική εκτίμηση της περιγραφής (συνέχεια)

- Μέση τιμή του RR:

$$ARR = \frac{1}{NQ} \sum_{q=1}^{NQ} RR(q)$$

- Βελτίωση AVR(q):

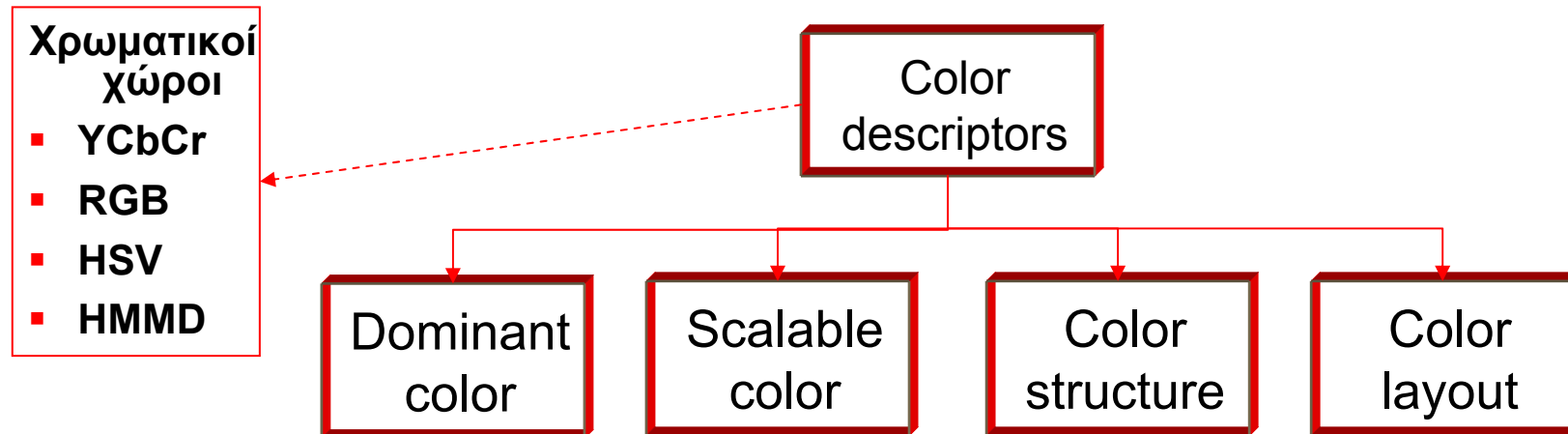
$$AVR(q) = \frac{\sum_{k=1}^{NG(q)} \text{Rank} * (k)}{NG(q)}$$

Όπου rank η σειρά ομοιότητας των εικόνων (1 → πλέον όμοια)

Modified Retrieval Rank $MRR = AVR(q) - 0.5[1 + NG(q)]$

Μικρές τιμές του MRR → καλή συμπεριφορά

Περιγραφή του χρώματος (Color descriptors)



Χρωματικοί χώροι (color spaces)

Επικρατούντα χρώματα (dominant colors)

Χρωματική δομή (color structure)

Χρωματικός σχεδιασμός (color layout)

Κλιμακωτή περιγραφή χρωμάτων (scalable color)



Χρωματικοί χώροι

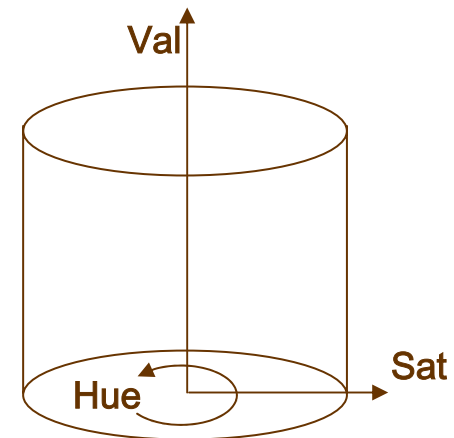
Οι χρωματικοί χώροι που χρησιμοποιούνται στο πρωτόκολλο MPEG7 είναι:

RGB

YCbCr (από τα πρωτόκολλα MPEG 1/2/4)

HSV Σχετίζεται:

- με την περιγραφή ιστογράμματος συνόλου στιγμιότυπων (group of frames histogram descriptor)
- με την κλιμακωτή περιγραφή χρωμάτων (scalable color descriptor)

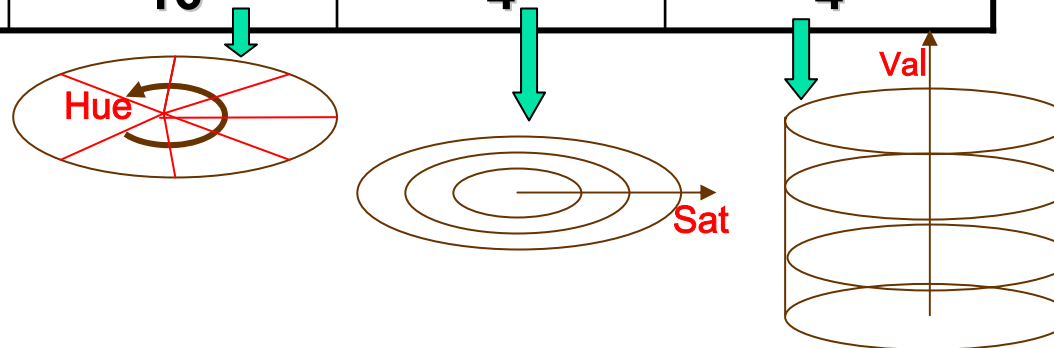




Χρωματικοί χώροι (συνέχεια)

HSV κβάντιση του χώρου

Συνολικός αριθμός bins	αριθμός bins για το H	αριθμός bins για το S	αριθμός bins για το V
16	4	2	2
32	8	2	2
64	8	2	4
128	8	4	4
256	16	4	4





Χρωματικοί χώροι (συνέχεια)

HMMD

1. Είναι *πιο κοντά* στο χώρο ομοιόμορφης αντίληψης (perceptually uniform color space)
2. Η ονομασία του οφείλεται στα μεγέθη:
Hue, Max, Min, Diff (Sum)

που σχετίζονται με τα αντίστοιχα **RGB** ως εξής:

$$\text{Max}=\max(\text{R},\text{G},\text{B})$$

$$\text{Min}=\min(\text{R},\text{G},\text{B})$$

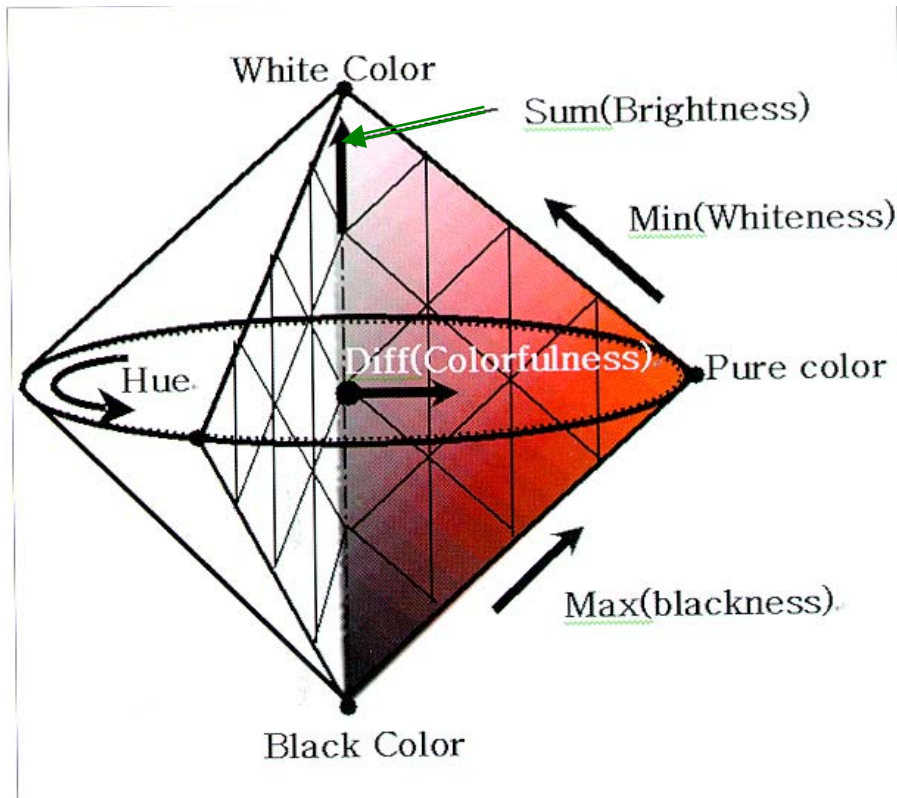
$$\text{Diff}=\text{Max}-\text{Min}$$

$$\text{Sum}=(\text{Max}+\text{Min})/2$$



Χρωματικοί χώροι (συνέχεια)

Ο χώρος HMMD με άξονες Diff, Sum, Hue



Κβάντιση του HMMD χώρου

# bins	256		128		64		32	
Υπο-διάστημα	H	S	H	S	H	S	H	S
0	1	32	1	16	1	8	1	8
1	4	8	4	4	4	4	4	4
2	16	4	8	4	4	4	4	4
3	16	4	8	4	8	2	4	1
4	16	4	8	4	8	1	4	1

Ο άξονας Diff διαμερίζεται στα εξής 5 διαστήματα:

[0,6],[6,20],[20,60],[60,110), και [110,255)



Περιγραφή με τα επικρατούντα χρώματα (dominant color descriptor)

Προσφέρει μία περιγραφή των βασικών χρωμάτων μίας εικόνας με σκοπό την εύρεση (απόκτηση) των ομοίων εικόνων

Ορισμός: $F = \{(c_i, p_i, u_i), s\} \quad i=1,2,\dots,N$

N ο αριθμός των (επικρατούντων) χρωμάτων. Ορίζεται ίδιος για όλη την βάση (συνήθης τιμή $N=8$)

c_i είναι το χαρακτηριστικό διάνυσμα (πχ. RGB) της ομάδας (3bits)

p_i είναι το ποσοστό των pixels που αντιστοιχούν στην ομάδα i

u_i είναι η διακύμανση των pixels της ομάδας i .

s αριθμός που περιγράφει την χωρική ομοιομορφία

RGB είναι ο συνήθης χρωματικός χώρος

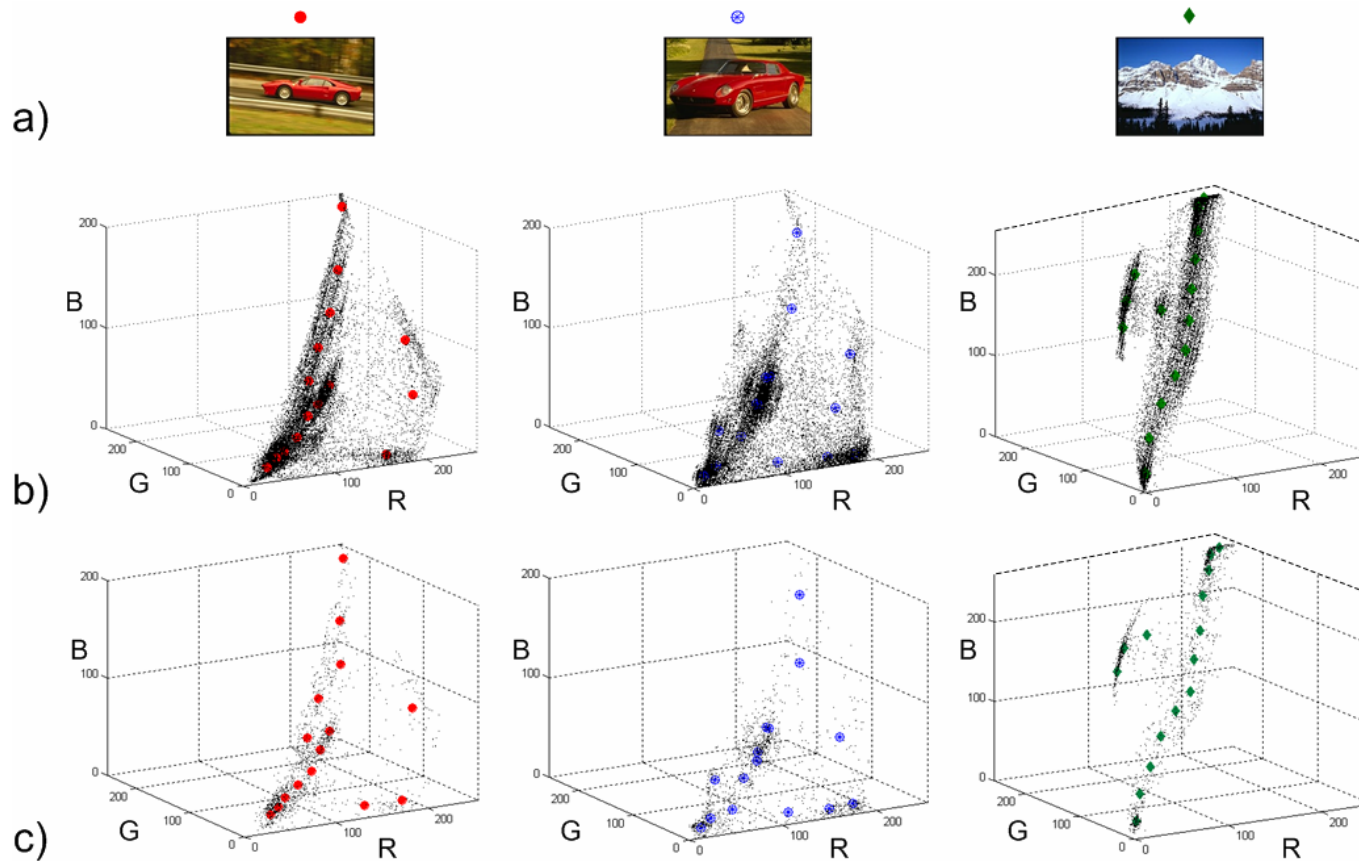


$F=\{(c_i, p_i, u_i), s\}$ Υπολογισμοί

- Ομαδοποίηση των pixels και υπολογισμός των c_i
Προτείνεται μία παραλλαγή τού Lloyd αλγόριθμου. (Μπορεί να χρησιμοποιηθεί ο *fcm*)
- Υπολογισμός των ποσοστών p_i ($\sum p_i=1$)
- Υπολογισμός της διακύμανσης u_i κάθε ομάδας
- Υπολογισμός της συνεκτικότητας για εύρεση ομάδων pixels με ίδιο χρώμα στο επίπεδο της εικόνας. Η τιμή s είναι η μέση τιμή όλων των ομάδων



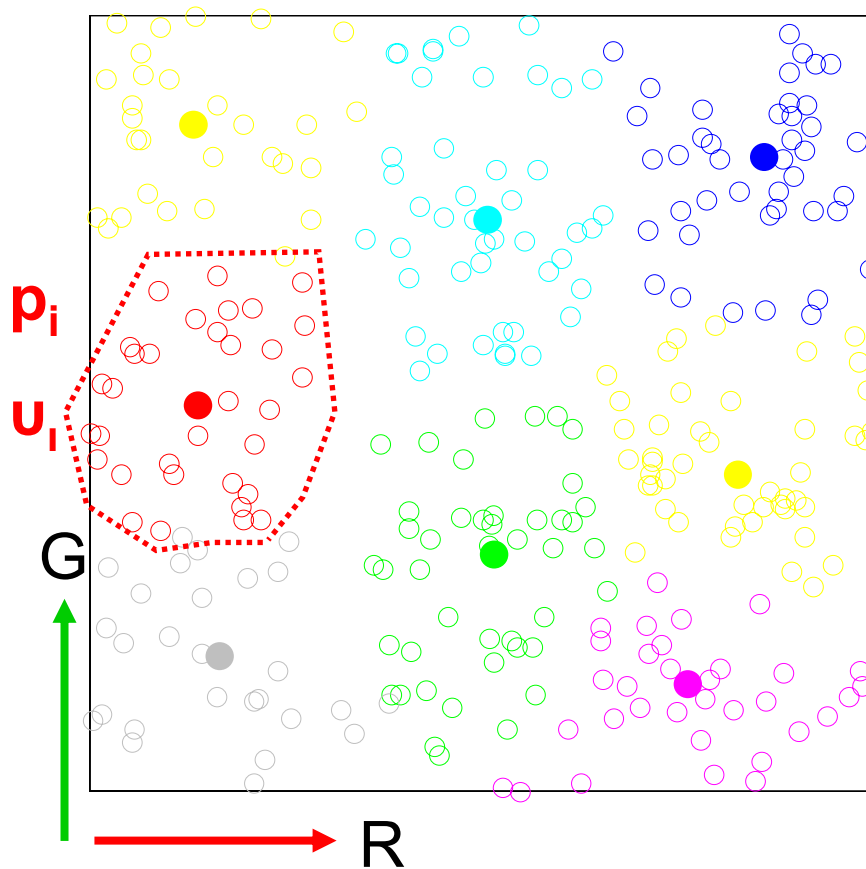
Ο χώρος των χαρακτηριστικών - Ομαδοποίηση



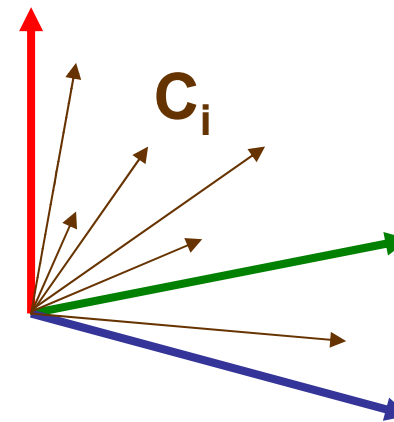


$$F = \{(c_i, p_i, u_i), s\}$$

Υπολογισμός των c_i, p_i, u_i



Τα pixels της εικόνας
ομαδοποιούνται στο **χώρο**
των χαρακτηριστικών (RGB)
σε **(N=8) ομάδες**

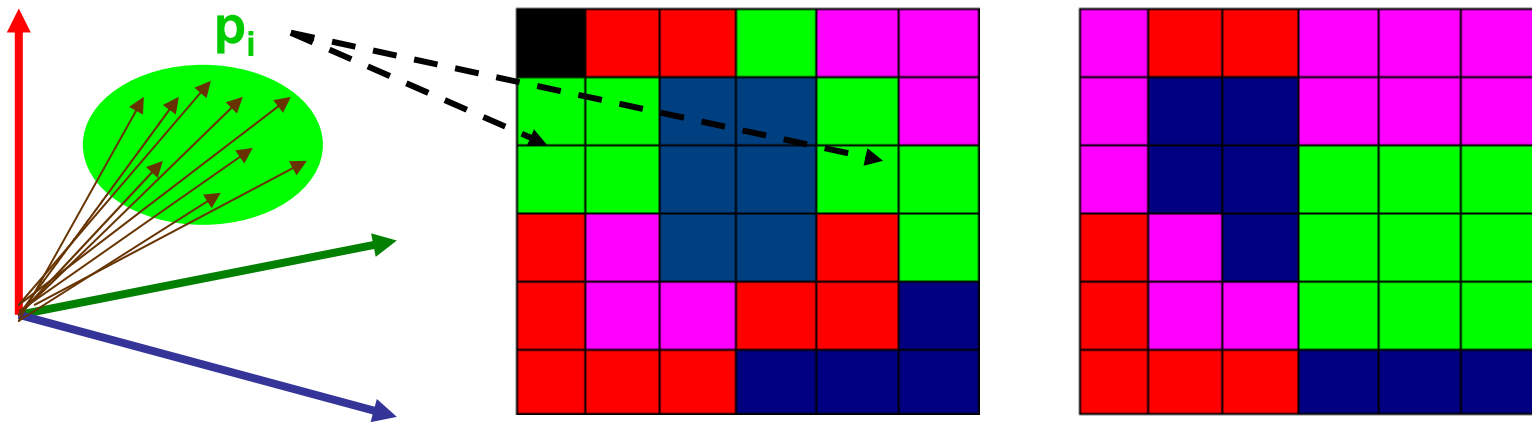




$F = \{(c_i, p_i, u_i), s\}$
Υπολογισμός του s

Συνεκτικότητα:

Δείχνει την σχέση μεταξύ χώρου χαρακτηριστικών και επιπέδου εικόνας



Η μέση τιμή των τμημάτων που «χωρίζεται» το p_i στο επίπεδο της εικόνας εκφράζει την τιμή s

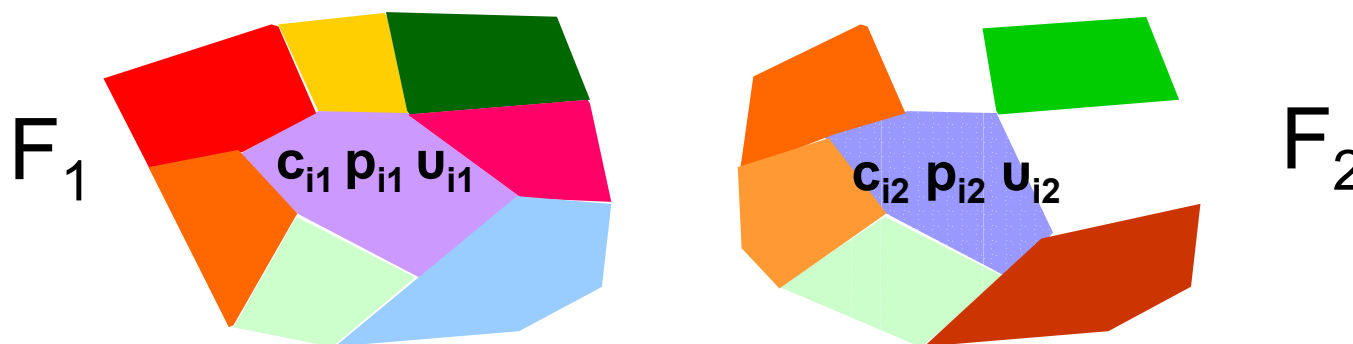


Μέτρα ομοιότητας

$$D^2(F_1, F_2) = \sum_{i=1}^{N_1} p_{1i}^2 + \sum_{j=1}^{N_2} p_{2j}^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{1i,2j} p_{1i} p_{2j}$$

$$a_{kl} = \begin{cases} 1 - d_{kl} / d_{\max} & d_{kl} \leq T_d \\ 0 & d_{kl} > T_d \end{cases}$$

$$d_{\max} = aT_d$$





Μέτρα ομοιότητας

Παραλλαγές

1. Για χωρική συνάφεια
$$D_s = w_1 \text{abs}(s_1 - s_2) D + w_2$$
2. Για να συμπεριληφθεί και η u_s

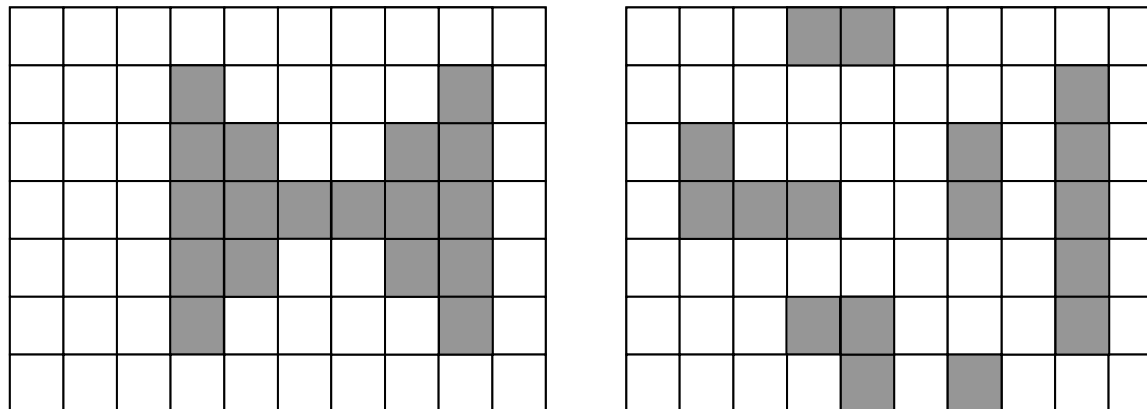


Περιγραφή Χρωματικής δομής

(Color Structure Descriptor)

Είναι ένα έγχρωμο ιστόγραμμα με πληροφορία και για την δομή του χρώματος.

Για τον υπολογισμό χρησιμοποιείται ένα **δομικό στοιχείο**



Τα ιστογράμματα των δύο τμημάτων είναι ίδια αλλά η χρωματική δομή διαφορετική



Color Structure Descriptor (συνέχεια)

$$\bar{h}_s(m) \quad m \in \{1, \dots, M\}$$

- s** = το μέγεθος του δομικού στοιχείου (πχ $s=3^2$)
- M** = το μέγεθος του κβαντισμένου χρωματικού χώρου που είναι ο **HMMD** ($M=256, 128, 64, 32$)
- Η περιγραφή της χρωματικής δομής μίας εικόνας αναφέρεται στη κατανομή του χρώματος αφενός και στη χωρική δομή αφετέρου
- Μοιάζει με έγχρωμο ιστόγραμμα



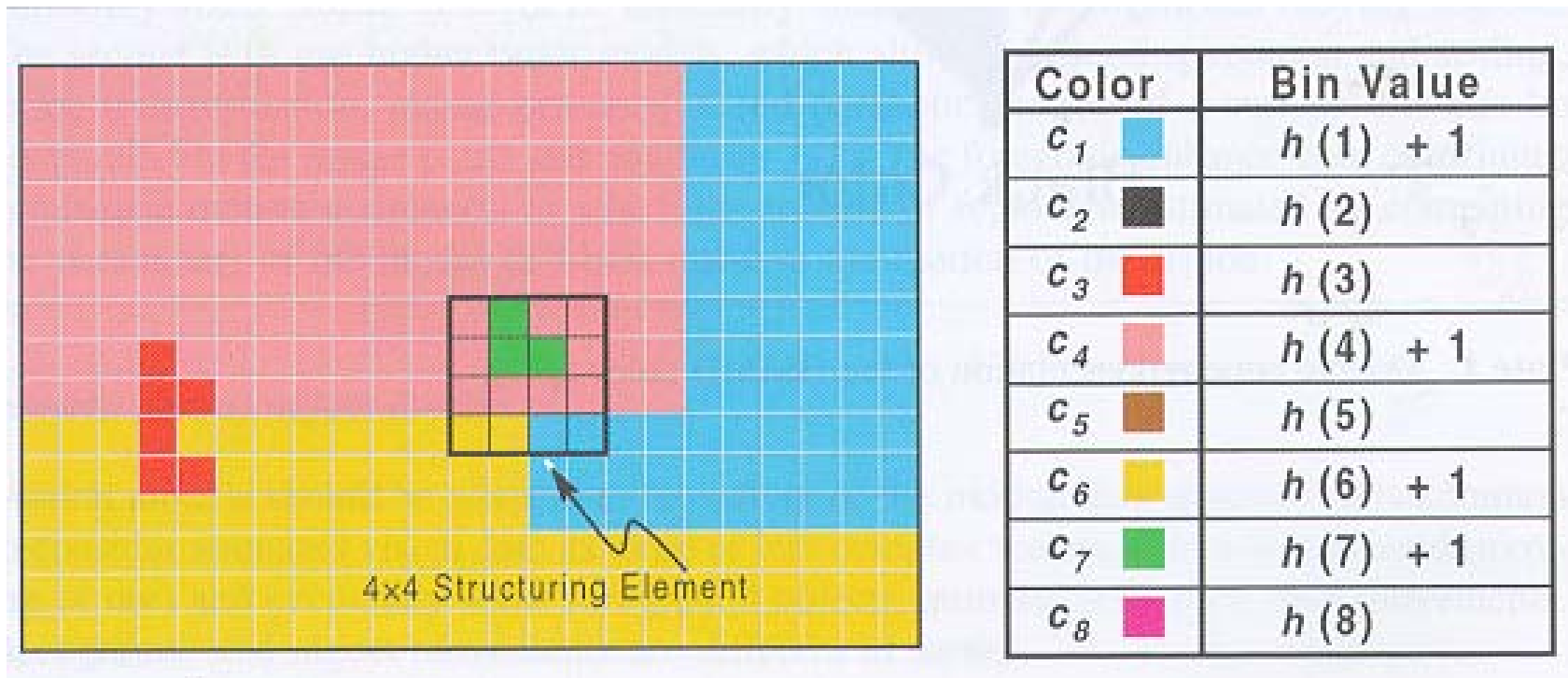
Color Structure Descriptor Υπολογισμός του $\overline{h_s}$ (m)

1. Ορίζεται ο αριθμός των bins πχ 256
2. Γίνεται κβάντιση του χρωματικού χώρου HMMD μέσα από διαδικασίες ομαδοποίησης
3. Ορίζεται ένα δομικό στοιχείο (structuring element) δηλ ένα παράθυρο (συνήθως 8×8) που διατρέχει την εικόνα
4. Σε κάθε σημείο του παραθύρου εντοπίζονται τα υπάρχοντα χρώματα και για κάθε ένα (χρώμα όχι pixel) προστίθεται μία μονάδα στο αντίστοιχο bin



Color Structure Descriptor

Υπολογισμός του $\bar{h}_s(m)$ -παράδειγμα





Περιγραφή Χρωματικού σχεδιασμού (Color Layout Descriptor-CLD)

- Ενδείκνυται για γρήγορη ανάκτηση
- Είναι ανεξάρτητος της ανάλυσης
- Αναδεικνύει την χωρική κατανομή του χρώματος
- Χρησιμοποιείται ο YCbCr χρωματικός χώρος

Υλοποιείται σύμφωνα με το διάγραμμα:

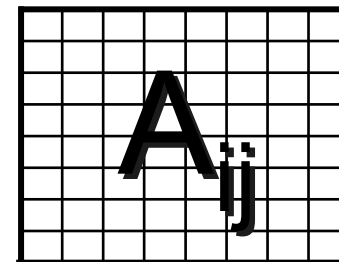




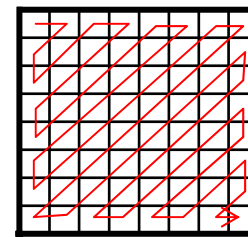
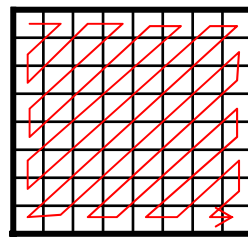
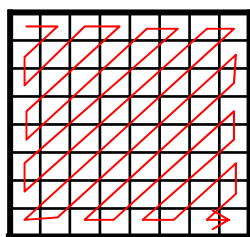
Αρχική εικόνα



64 blocks



Εξαγωγή 64
διανυσμάτων



64 συντελεστές DCT για κάθε συνιστώσα, Y, Cb, Cr

Zig-zag scanning και επιλογή 12 συντελεστών - κβάντιση



Μέτρα ομοιότητας

Σύγκριση δύο εικόνων με CLDs $\{DY, DCb, DCr\}$ και $\{DY', DCb', DCr'\}$ γίνεται με την εξής σχέση:

$$D = \sqrt{\sum_i w_y (DY_i - DY'_i)^2} + \sqrt{\sum_i w_b (DCb_i - DCb'_i)^2} \sqrt{\sum_i w_r (DCr_i - DCr'_i)^2}$$

- Με τα βάρη w δίνεται έμφαση στις χαμηλές συχνότητες
- Η σειρά των συντελεστών καθορίζεται από το διάνυσμα του zig-zag

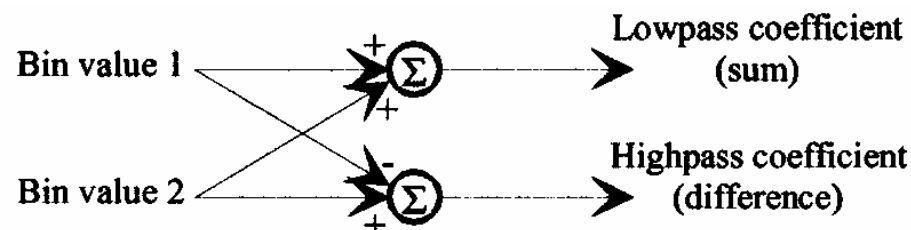


Κλιμακωτή περιγραφή χρωμάτων (scalable color descriptor)

- Ο περιγραφέας SCD ορίζεται στον χρωματικό χώρο HSV
- Χρησιμοποιεί το σχήμα κωδικοποίησης του μετασχηματισμού Haar στο έγχρωμο ιστόγραμμα.
- Ο μετασχηματισμός Haar δημιουργεί την κλιμακωτή περιγραφή που αποτελεί και το χαρακτηριστικό για διαδικασίες σύγκρισης εικόνων.
- Ο περιγραφέας αυτός επεκτείνεται σε ομάδες εικόνων ή ομάδες στιγμιότυπων σε δεδομένα video (*group of frames-GoF /group of pictures - GoP*)



scalable color descriptor (συνέχεια)

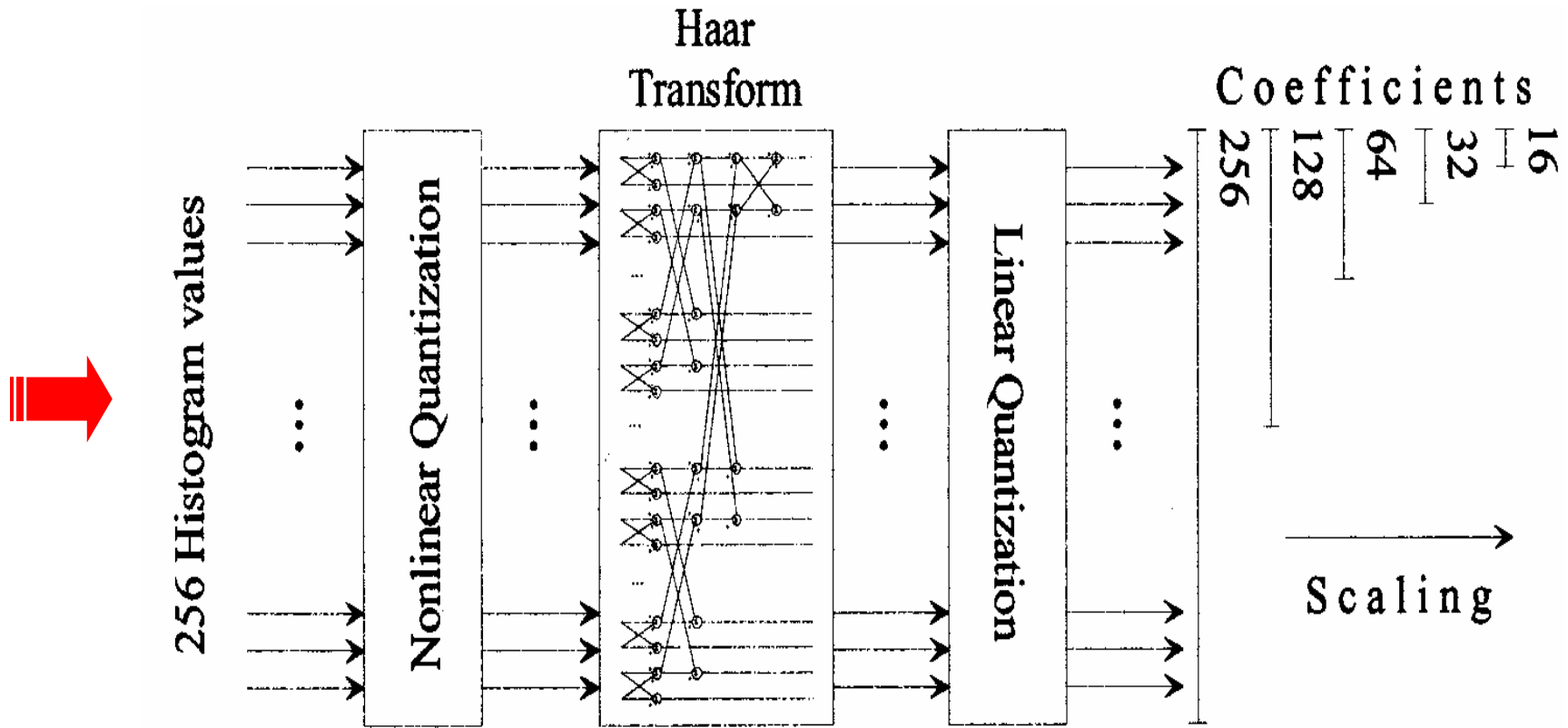


- Η βασική μονάδα του SCD αποτελείται από μία πράξη άθροισης και μία διαφοράς
- Εφαρμόζεται σε διαδοχικές τιμές του (ακεραίας τιμής 4-bit) ιστογράμματος.
- Η κλιμάκωση γίνεται με την άθροιση όπου σε διαδοχικά βήματα λαμβάνουμε 256, 128, 64 ή 32 bin
- Η διαφορά δεν έχει σημαντική πληροφορία και γιαυτό χρησιμοποιεί μικρό αριθμό bits





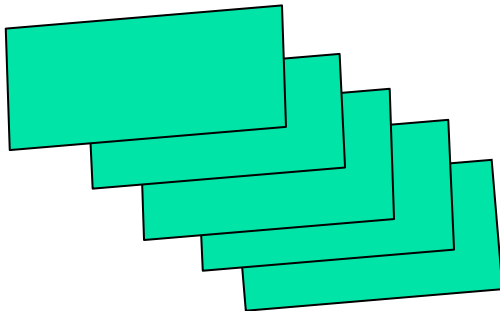
(συνέχεια)



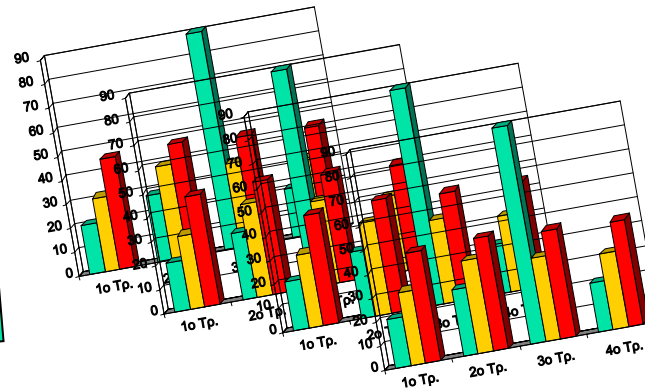


group of frames-GoF group of pictures - GoP

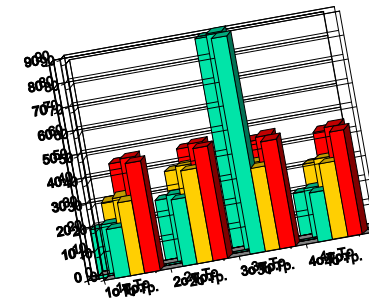
Αρχικά frames



εξαγωγή ιστογραμμάτων



«άθροιση»



Η «άθροιση» γίνεται:

- Με μέση τιμή των ιστογραμμάτων (σε κάθε bin)
- Με διάμεση τιμή (σε κάθε bin)
- Με επιλογή του ελάχιστου (σε κάθε bin) -intersection



Περιγραφή υφής

- Περιγραφέας ομοιογένειας (homogeneous texture descriptor)
- Περιγραφέας (texture browsing descriptor)
- Περιγραφέας ιστογράμματος ακμών (edge histogram descriptor)

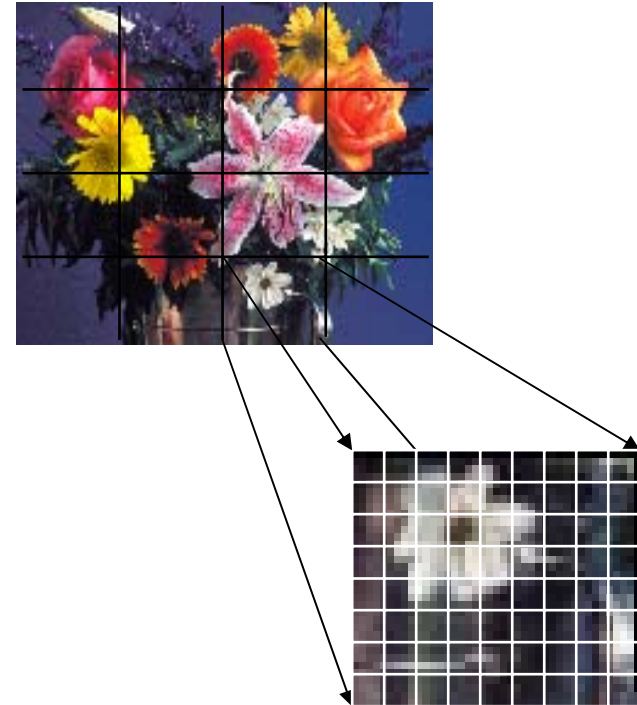


Περιγραφή ιστογράμματος ακμών

(edge histogram descriptor -EHD)

Τα βήματα:

- Χωρίζεται η εικόνα σε $4 \times 4 = 16$ υποεικόνες
- Κάθε υποεικόνα χωρίζεται σε τετραγωνικά blocks (~1100)
- Και κάθε block χαρακτηρίζεται με ένα τύπο ακμής: οριζόντια
κάθετη
διαγώνια 45
διαγώνια 135
χωρίς κατεύθυνση





- Το ιστόγραμμα των ακμών υπολογίζεται ως εξής:

$h(0)$ =σχετικός αριθμός καθέτων ακμών στην υποεικόνα (0,0)

$h(1)$ =σχετικός αριθμός οριζοντίων ακμών στην υποεικόνα (0,0)

$h(2)$ =σχετικός αριθμός 45° ακμών στην υποεικόνα (0,0)

$h(3)$ =σχετικός αριθμός 135° ακμών στην υποεικόνα (0,0)

$h(4)$ =σχετικός αριθμός ακμών χωρίς κατεύθυνση στην υποεικόνα (0,0)

.....
.....

$h(79)$ =σχετικός αριθμός ακμών χωρίς κατεύθυνση στην υποεικόνα (3,3)



Τελεστές για αναγνώριση των ακμών.

1	-1
1	-1

1	1
-1	-1

$\sqrt{2}$	0
0	$-\sqrt{2}$

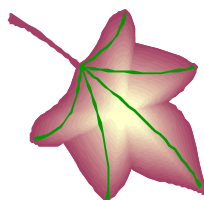
0	$\sqrt{2}$
$-\sqrt{2}$	0

2	-2
-2	2

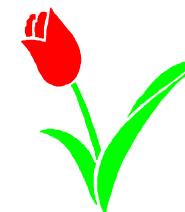
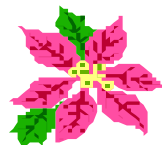
Η μέγιστη τιμή εφόσον είναι μεγαλύτερη από ένα κατώφλιο χαρακτηρίζει το block

Εφαρμογές – Διαχείριση δεδομένων φυτών

Φύλλα



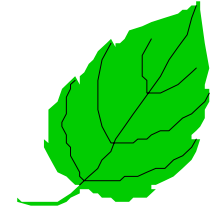
Άνθη



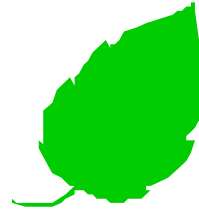
Βλαστοί



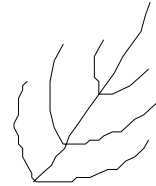
Περιγραφή φύλλων



Σχήμα



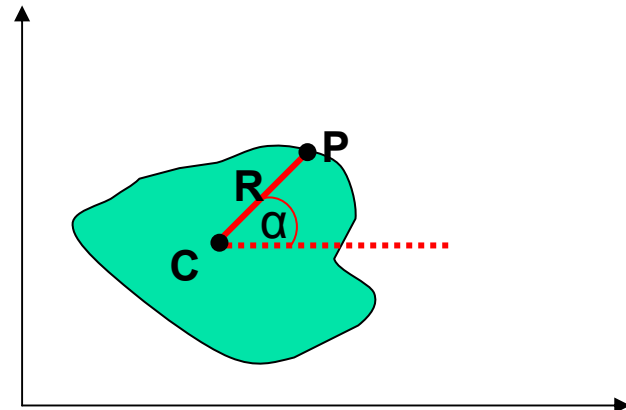
Νεύρωση



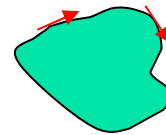
Περιγραφή φύλλων (συνέχεια)

A. Εξαγωγή Χαρακτηριστικών σχήματος

- κώδικες αλυσίδας
- μετασχ. «medial axis»
- Fourier descriptors
- Καμπύλη αποστάσεων από το κέντρο βάρους (centroid contour distance)



- Κυκλικότητα $\alpha = \frac{4\pi (\text{εμβαδό})}{(\text{περίμετρος})^2}$



- Ιστόγραμμα κώδικα γωνίας

- Εκκεντρότητα (eccentricity) $\epsilon = \frac{\text{μικρότερη ιδιοτιμή}}{\text{μεγαλύτερη ιδιοτιμή}}$

Περιγραφή φύλλων (συνέχεια)

B. Μέτρα ομοιότητας σχήματος

B1. Καμπύλης αποστάσεων (από το κέντρο βάρους) -CCD

$$D = \sqrt{\frac{\sum_{i=1}^n |f_1(i) - f_2(i)|}{n}} \quad D_c = \min\{D_1, D_2, \dots, D_j, \dots, D_m\}$$

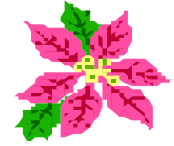
B2. Ιστογράμματος κώδικα γωνίας -ACH

$$D_h(I, J) = \sum_{i=1}^m |H_i(I) - H_i(J)|$$

B3. Εκκεντρότητα - Eccentricity

$$D_e(I, J) = |e_I - e_J|$$

Περιγραφή ανθέων



- Βασικό χαρακτηριστικό: Χρώμα – περιοχή χρώματος
- Άλλα χαρακτηριστικά: Σχήμα - Υφή



Περιγραφή ανθέων (συνέχεια)

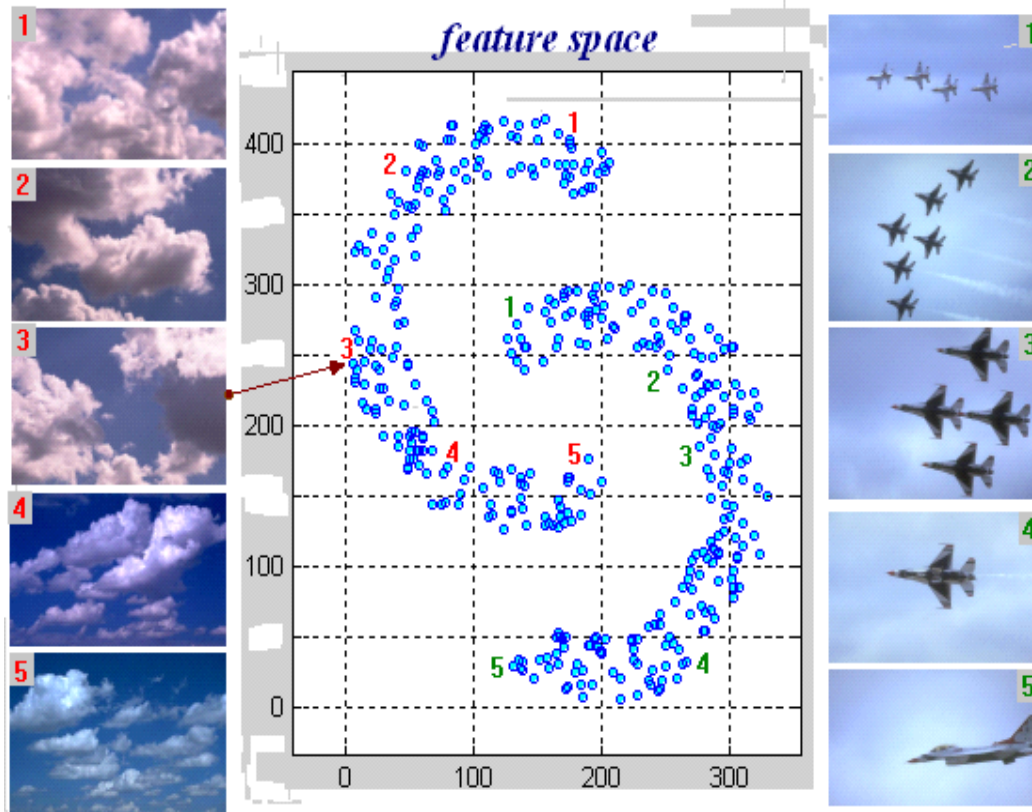
A. Εξαγωγή Χαρακτηριστικών

- A1. Εγχρωμο ιστόγραμμα
- A2. Ομαδες χρωμάτων (clustering)
- A3. Χωρική κατανομή χρωμάτων

B. Μέτρα ομοιότητας

- B1. Απόσταση εγχρώμων ιστογραμμάτων
- B2. Απόσταση “χωρικών” ιστογραμμάτων
- B3. Απόσταση βασικών χρωματων (DCD)

Χώρος των Χαρακτηριστικών και ελάττωση διάστασης



Μια μορφή αναπαράστασης εικόνων στο διδιάστατο χώρο. Κάθε μία εικόνα παριστάνεται με ένα διάνυσμα χαρακτηριστικών. Και όλη η βάση με ένα πίνακα αποστάσεων. Ελαττώνοντας τη διάσταση του χώρου σε 2 «βλέπουμε» την βάση στο επίπεδο

Βιβλιογραφία

1. Introduction to MPEG7, B.Smanjunath, P. Salembier and T. Sikora Eds, J.Wiley&Sons, Ltd, ENGLAND, 2002, Chapter13
2. B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan and Akio Yamada *Color and Texture Descriptors* IEEE Trans. On Circuits and Systems for Video Technology, VOL.11, NO. 6, JUNE 2001
3. <http://www.lgcit.com/MIGR/cmip/hmmd/hmmd.html>
4. http://www.lis.ei.tum.de/research/bv/topics/mmdb/e_mpeg7.html
5. Text of ISO/IEC 15 938-3 Multimedia Content Description Interface- Part 3: Visual. Final Committee Draft, ISO/ IEC/ JTC1/ SC29/WG11, Doc. N4062, Mar. 2001.
6. MPEG-7 Visual Experimentation Model (XM), Version 10. ISO/IEC/JTC1/SC29/WG11, Doc. N4063, Mar. 2001

MPEG7 Ασκήσεις - εργασίες

1. Υλοποίηση face descriptor
2. Υλοποίηση texture descriptor
3. Shape descriptor
4. Διαχείριση δεδομένων σε φυτά
 - α) κλασσικά εργαλεία διαχείρισης(Lucid, Uconn, CALFORA)
 - β) υπολογιστικά (computer aided) εργαλεία διαχείρισης
5. Επεξεργασία εικόνας και ανάκτηση α) «φύλλα», β) άνθη
εφαρμογή στη βάση του Corel
6. Ανάκτηση (βάσει περιεχομένου) ιατρικών δεδομένων
1ας, 2, 3 και 4 διαστάσεων