

Ανάκτηση πολυμεσικού περιεχομένου

Ανίχνευση / αναγνώριση προσώπων

Ανίχνευση / ανάγνωση κειμένου

Ανίχνευση αντικειμένων

Οπτικές λέξεις

Δεικτοδότηση

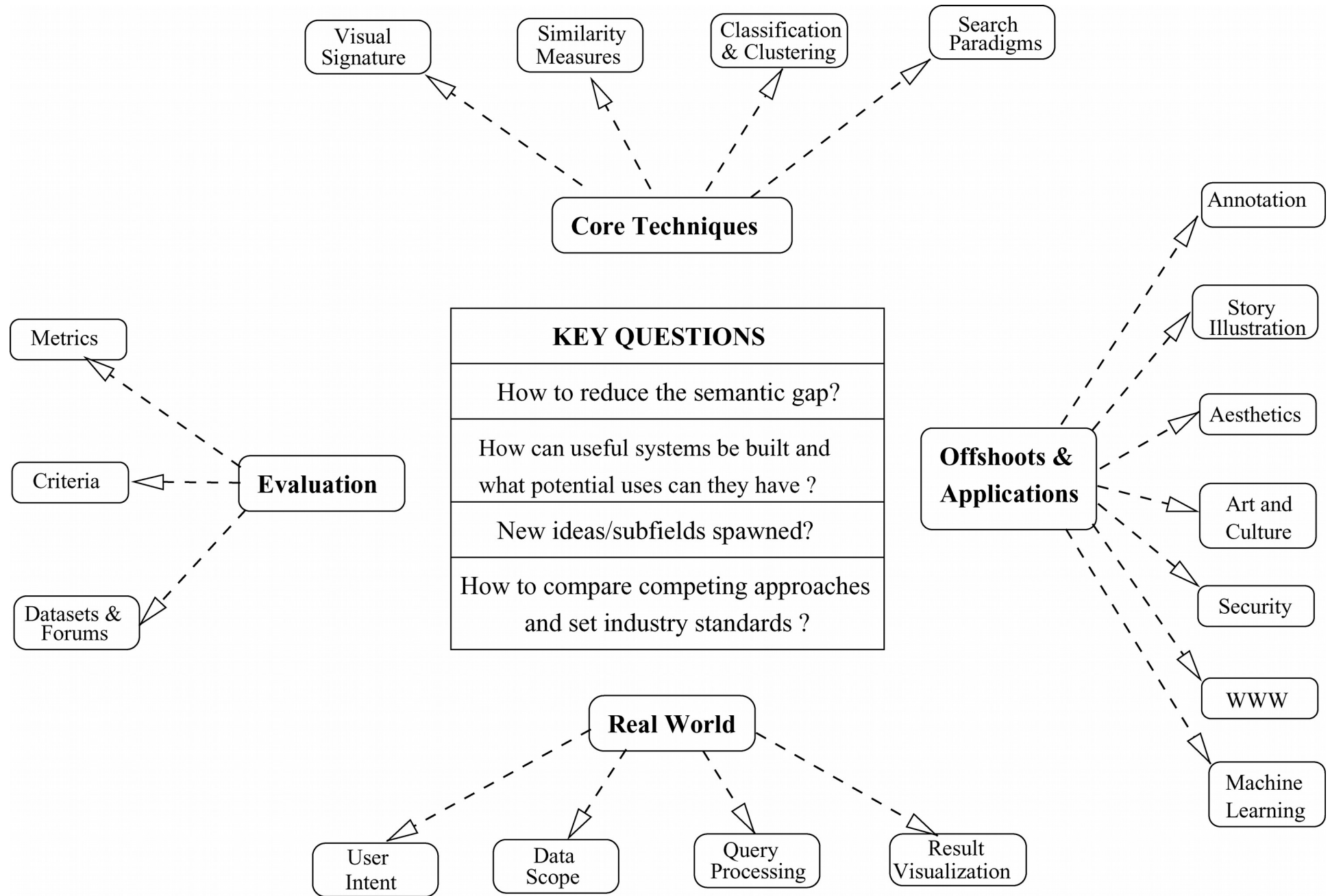
Σχέσεις ομοιότητας

Κατηγοριοποίηση ειδών μουσικής

Διάκριση φωνής / μουσικής

Τμήμα Επιστήμης Υπολογιστών

HY-474

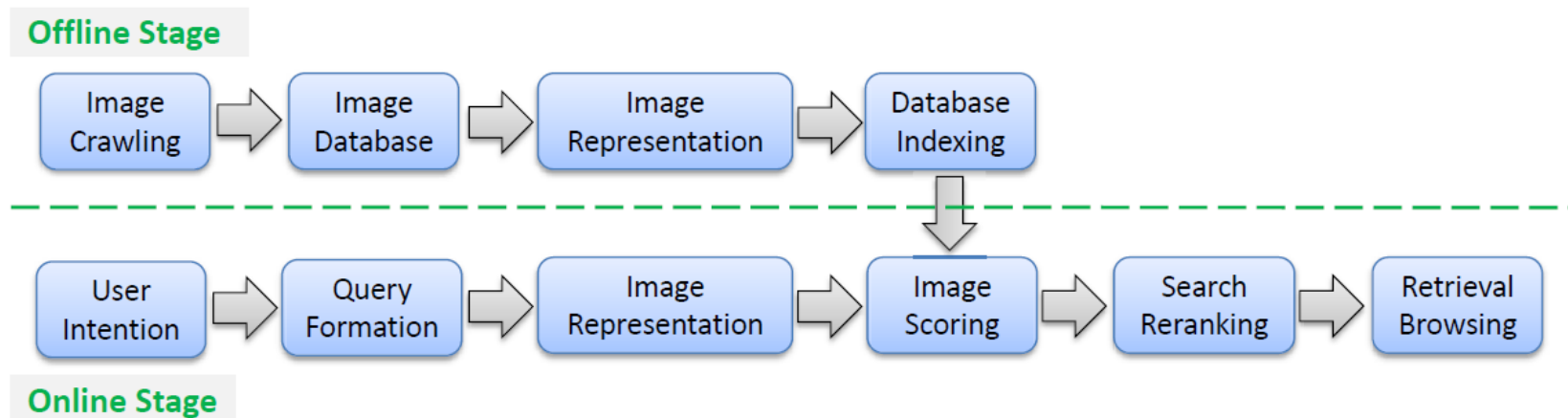


Ανάκτηση περιεχομένου στον πραγματικό κόσμο

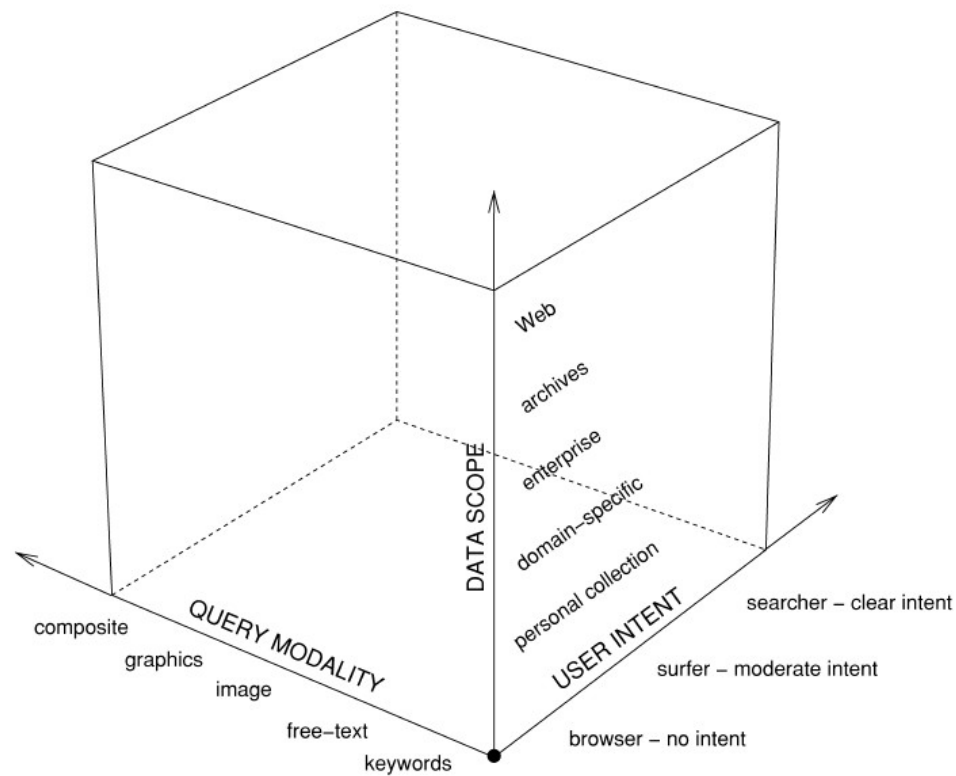
Πρόθεση χρήσης : περιήγηση, κατηγορία, τεκμήριο

Πεδίο δεδομένων : προσωπική συλλογή, ειδικού αντικειμένου,
αρχεία, Παγκόσμιος Ιστός

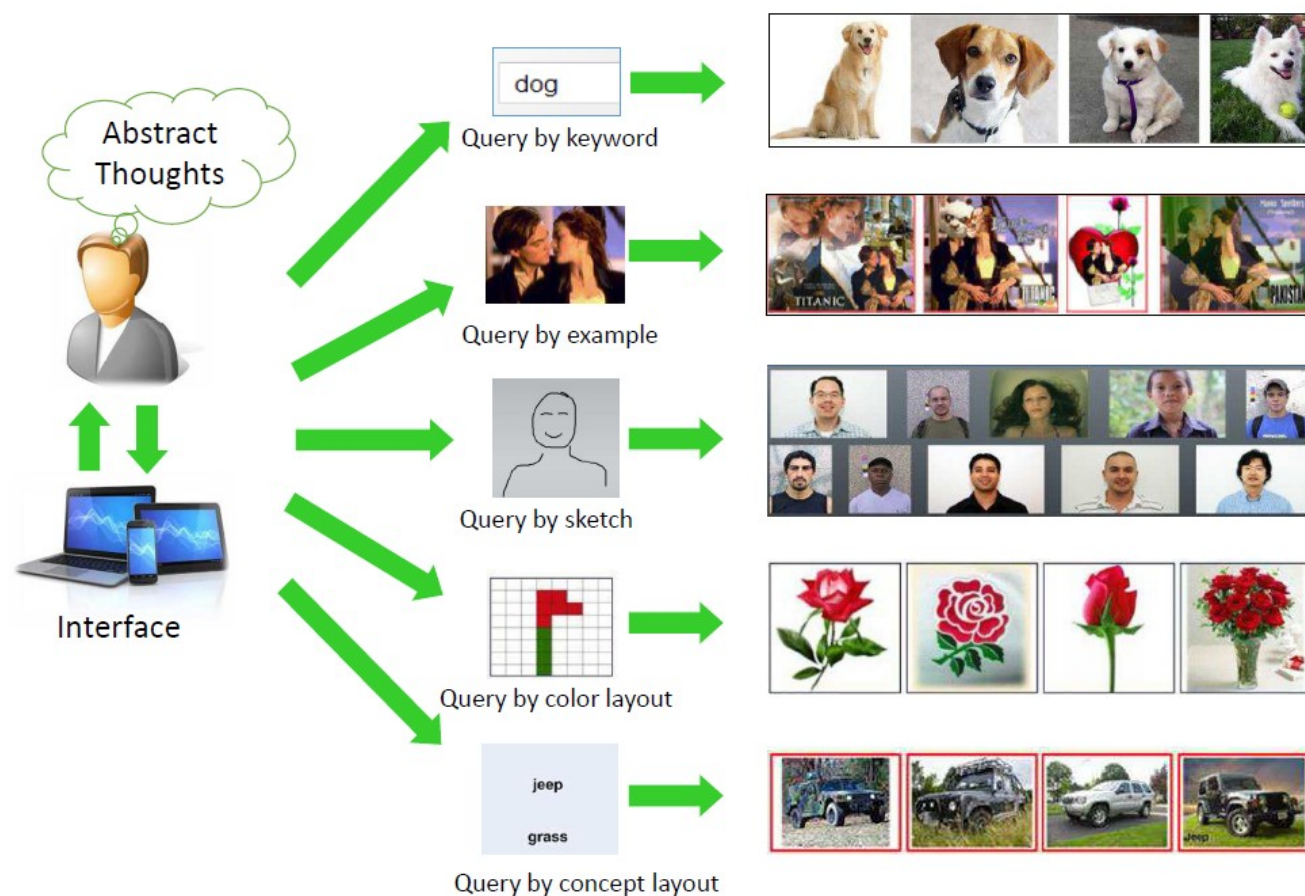
Τρόπος ερώτησης : λέξεις-κλειδιά, κείμενο, λόγος, εικόνα, σκίτσο



Σχήματα αναζήτησης περιεχομένου : χρήστης



Σχήματα αναζήτησης περιεχομένου : χρήστης



Μέτρα ομοιότητας / μάθηση

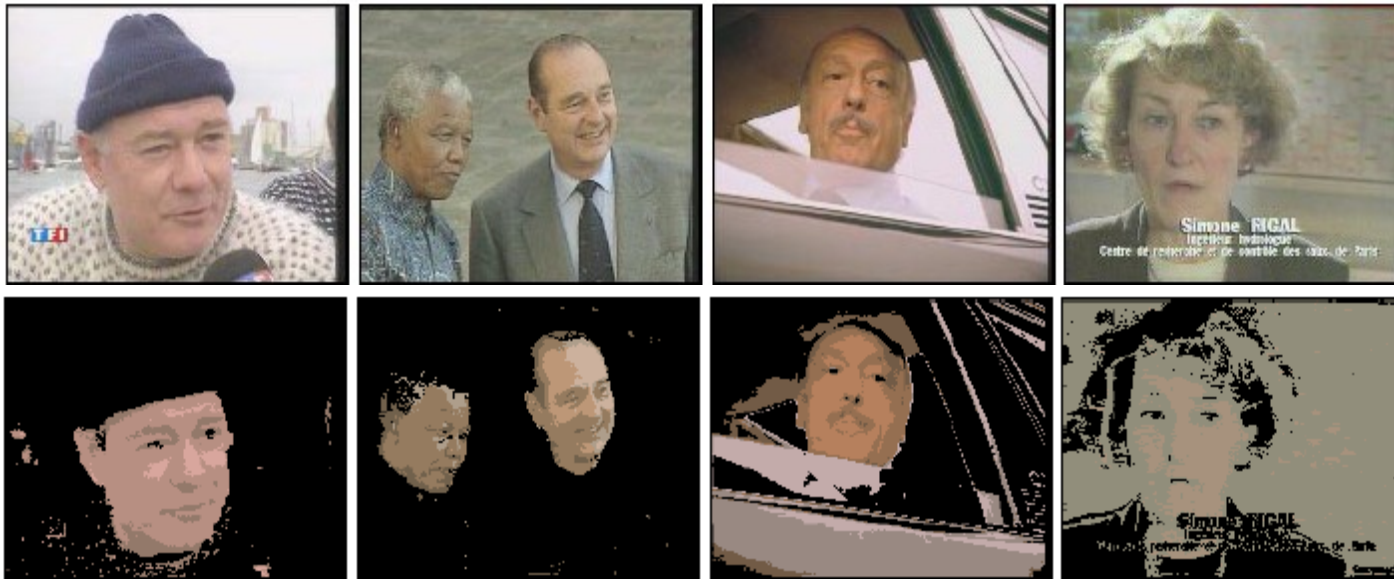
Συμφωνία με τη σημασιολογία
Αντοχή στο θόρυβο
Υπολογιστική επίδοση
Κλίμακα αντικειμένου
Ιδιότητες απόστασης

Ομαδοποίηση (ιεραρχική, συσσωρεύσεις, μίξεις)

Ταξινόμηση

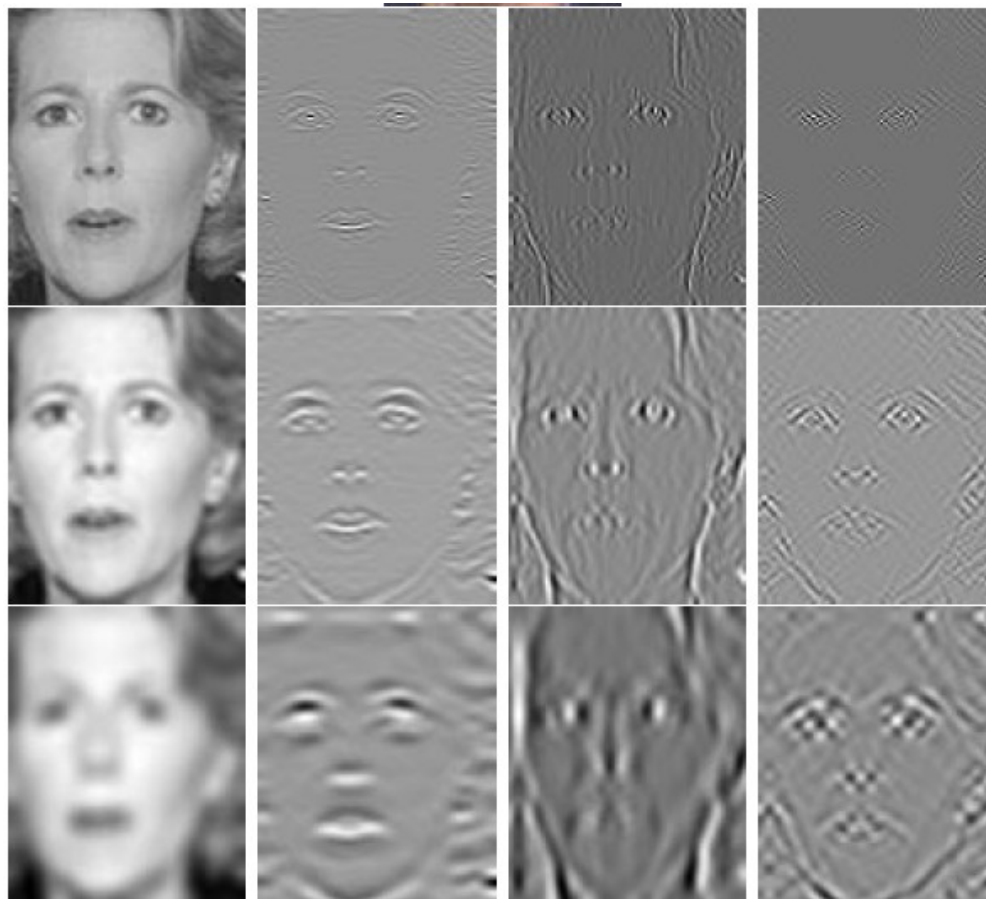
Ανίχνευση προσώπων : χρώμα

Ανίχνευση χρώματος δέρματος

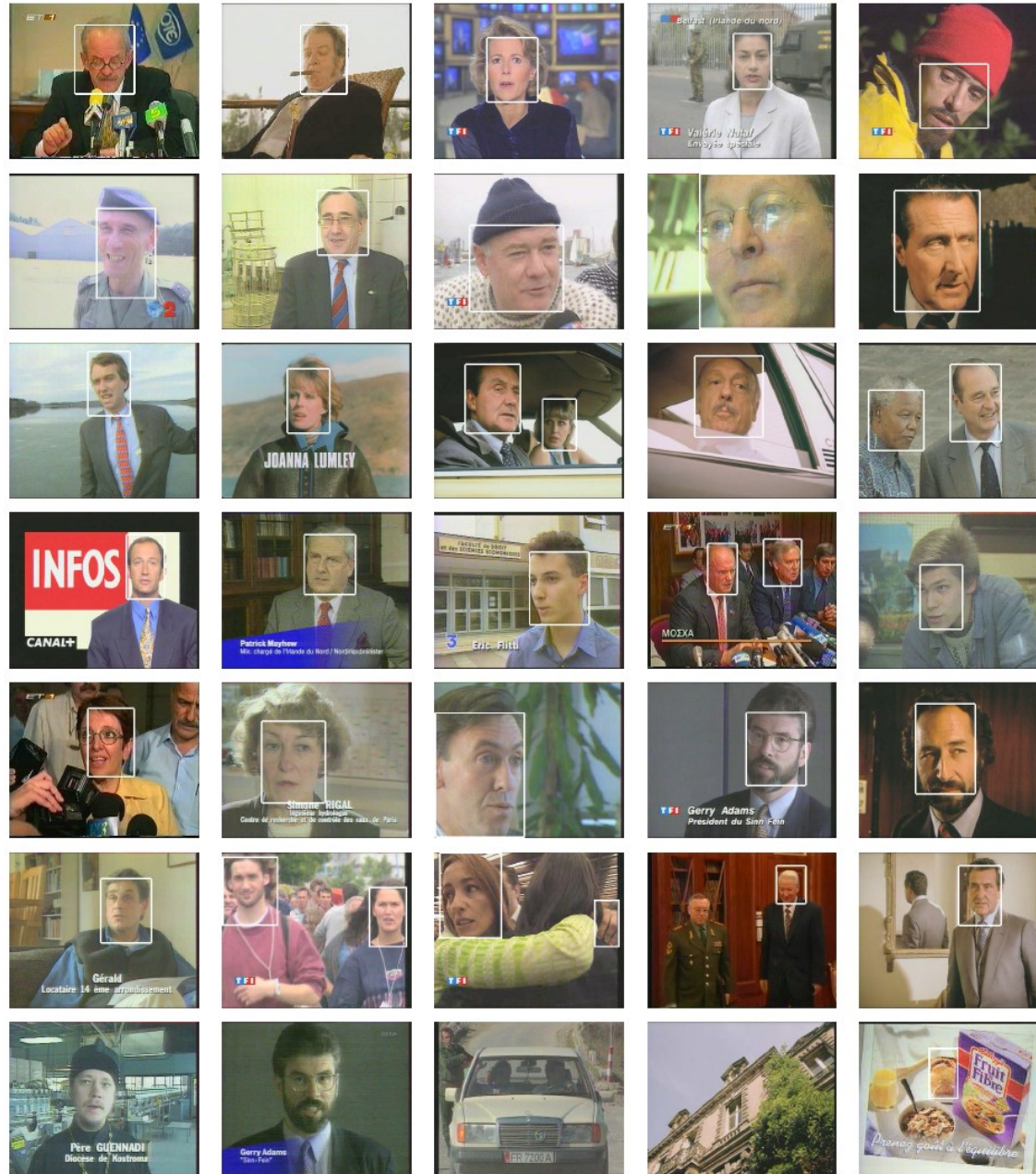


Ανίχνευση προσώπων : υφή

Ανάλυση σε
ζώνες συχνοτήτων

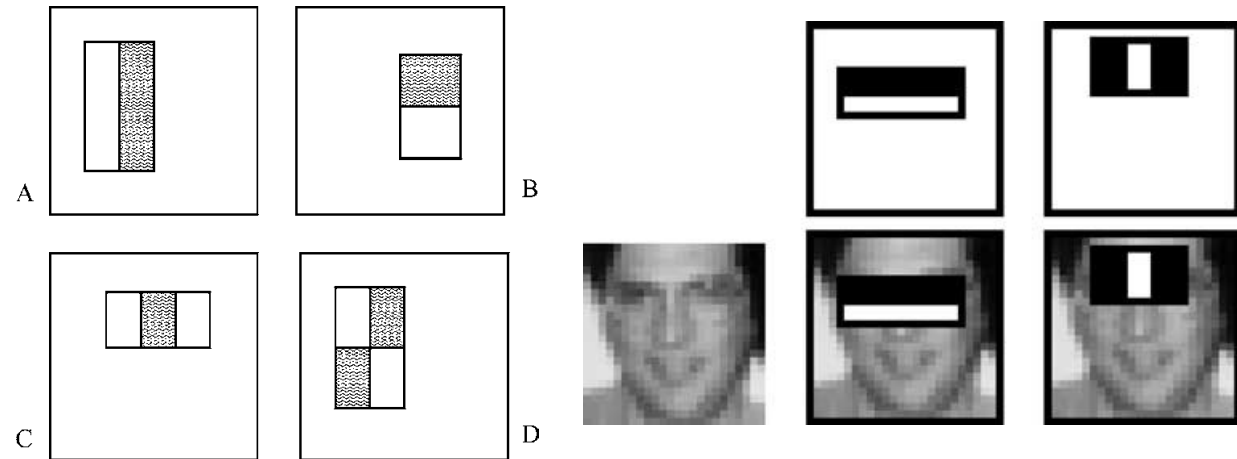


Ανίχνευση προσώπων

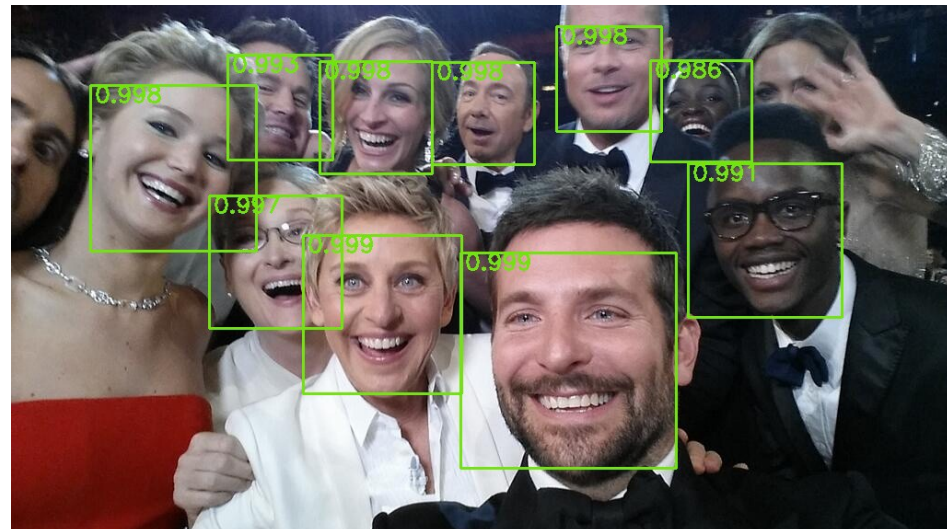


Ανίχνευση προσώπων : μάθηση

Εξαγωγή
χαρακτηριστικών
με μάθηση

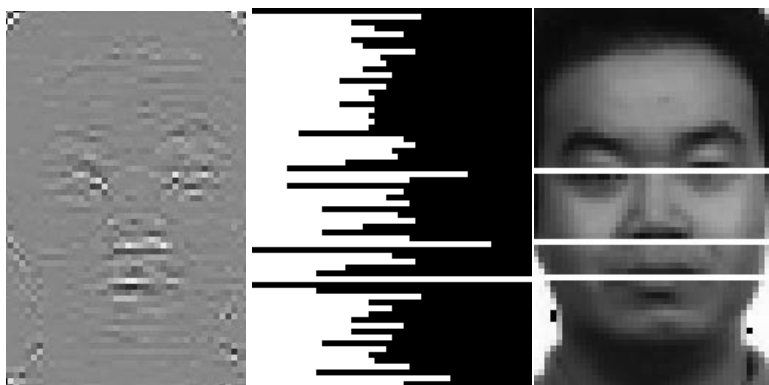


Συνελικτικό
νευρωνικό δίκτυο

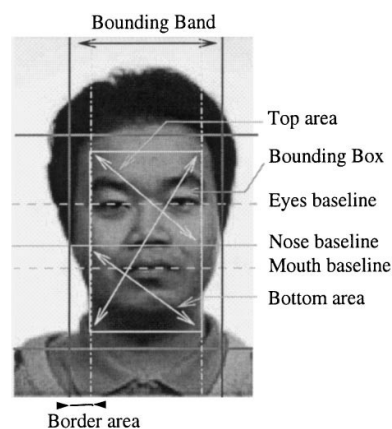


Αναγνώριση προσώπων

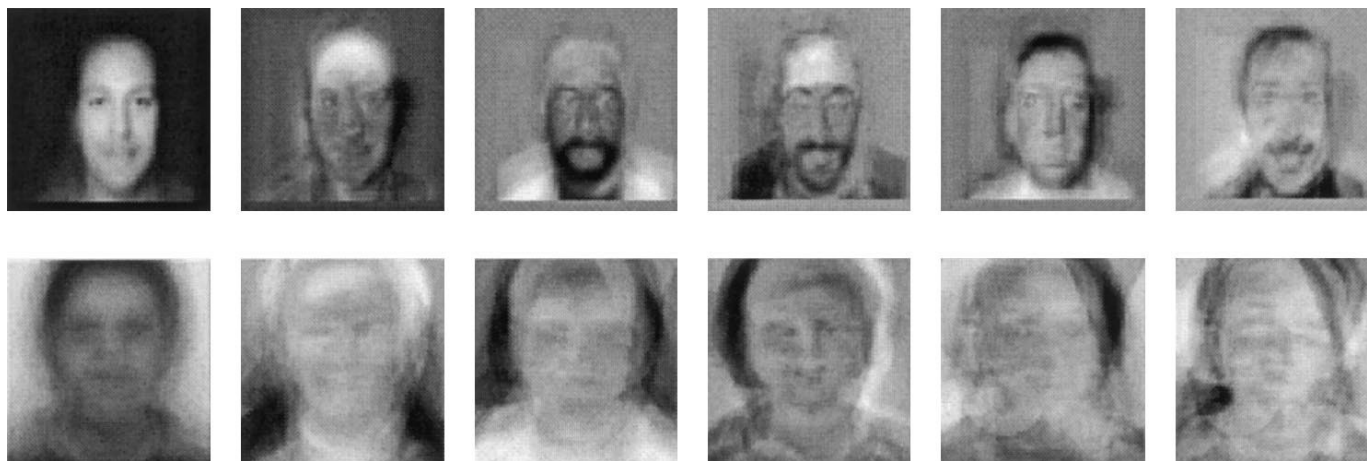
Ανάλυση σε ζώνες συχνοτήτων (διακριτός κυματιδιακός μετασχηματισμός)
Εντοπισμός χαρακτηριστικών



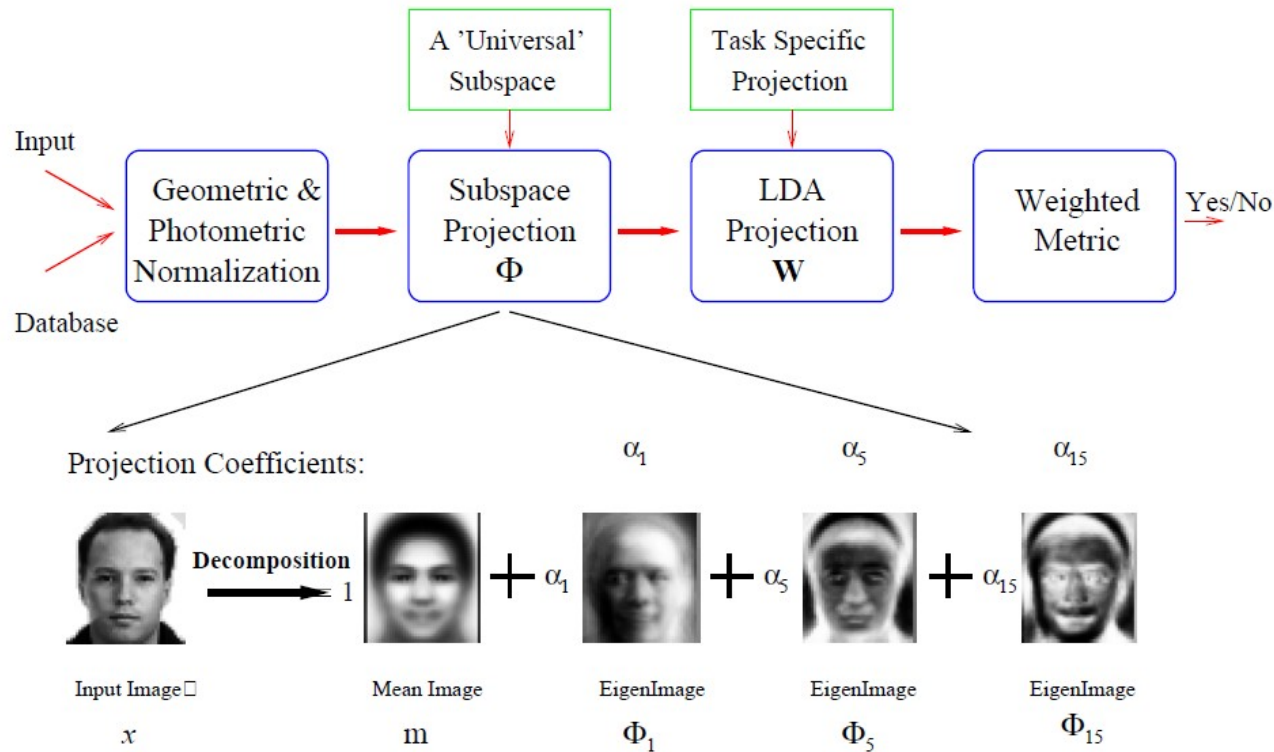
Ευθυγράμμιση



Ιδιοπρόσωπα



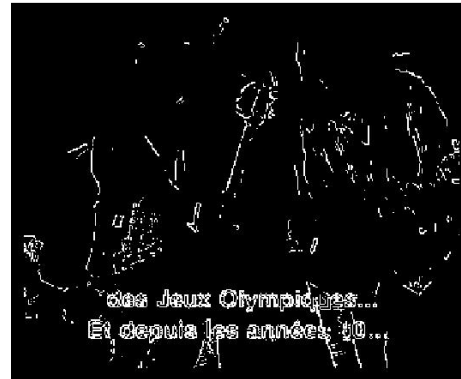
Αναγνώριση προσώπων



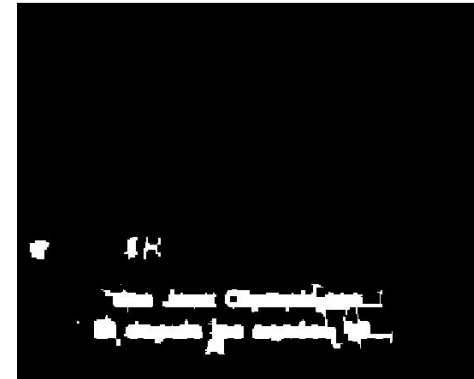
Εντοπισμός και αναγνώριση κειμένου



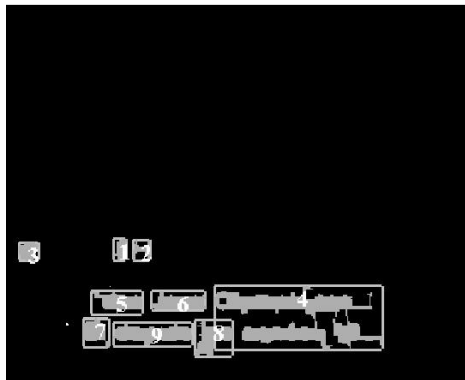
αρχική εικόνα



ανίχνευση ακμών



εξαγωγή από το φόντο



ομαδοποίηση



ανάλυση περιοδικότητας



αποτέλεσμα

Εντοπισμός και αναγνώριση κειμένου



Pur Malt

3615

Buckler

3615

Programme disponible
en cassette VHS.
Tapez 3615 code ARTE
ou écrivez à :

"LA SEPT VIDEO"
BP 630
F-60732 Ste Geneviève cedex 09

Die VHS-Kassette ist in
französischer Fassung erhältlich.
Schreiben Sie an "LA SEPT VIDEO"
Mehr dazu im Videotext auf Seite 107. **arte**

AVEC KAOMA ET CHICO ROBERTA

Pur Malt

3615

Buckler

3615

Programme disponible
en cassette VHS.
Tapez 3615 code ARTE
ou écrivez à :

"LA SEPT VIDEO"
BP 630
F-60732 Ste Geneviève cedex 09

Die VHS-Kassette ist in
französischer Fassung erhältlich.
Schreiben Sie an "LA SEPT VIDEO"
Mehr dazu im Videotext auf Seite 107. **arte**

Cette saveur, c'est Buckler.

Pur Malt

3615

Buckler

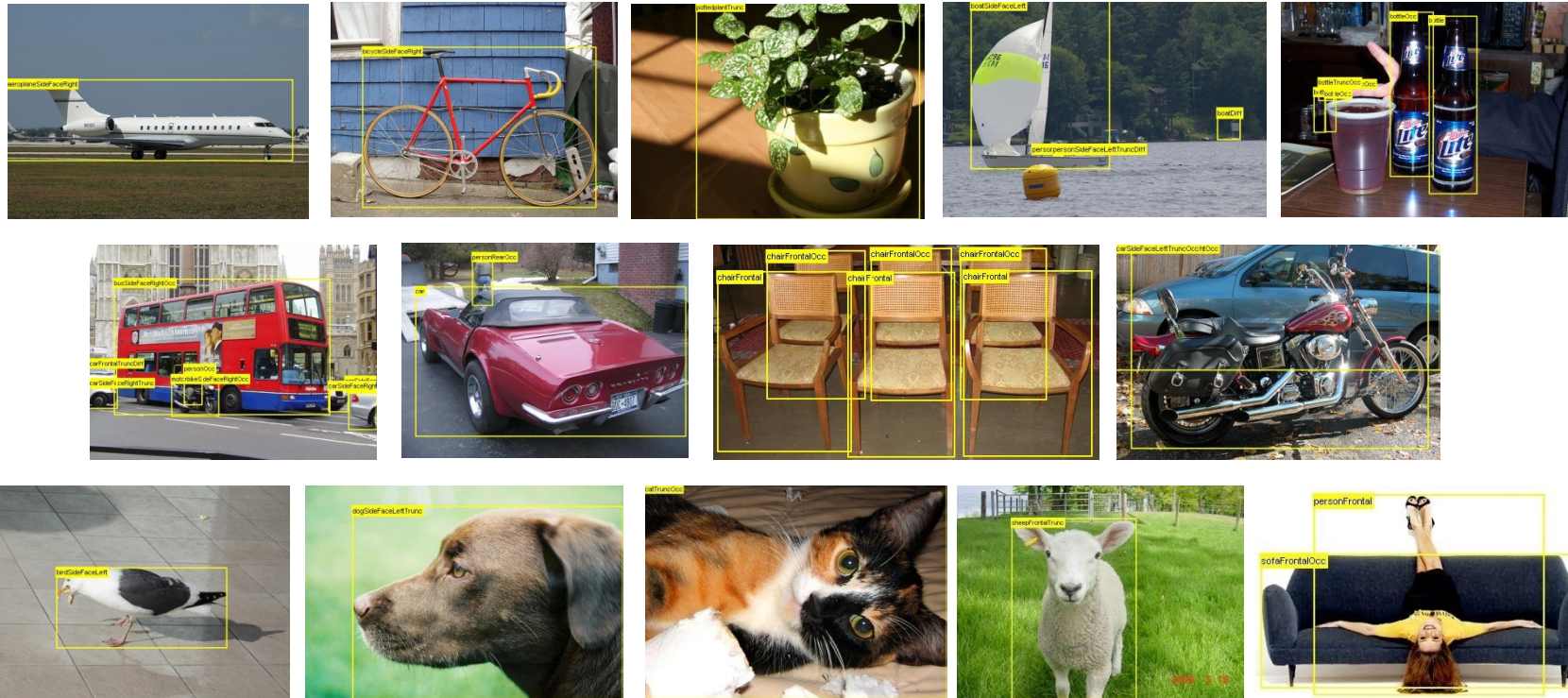
3615

Programme disponible
en cassette VHS.
Tapez 3615 code ARTE
ou écrivez à :

"LA SEPT VIDEO"
BP 630
F-60732 Ste Geneviève cedex 09

Die VHS-Kassette ist in
französischer Fassung erhältlich.
Schreiben Sie an "LA SEPT VIDEO"
Mehr dazu im Videotext auf Seite 107. **arte**

Ανίχνευση αντικειμένων

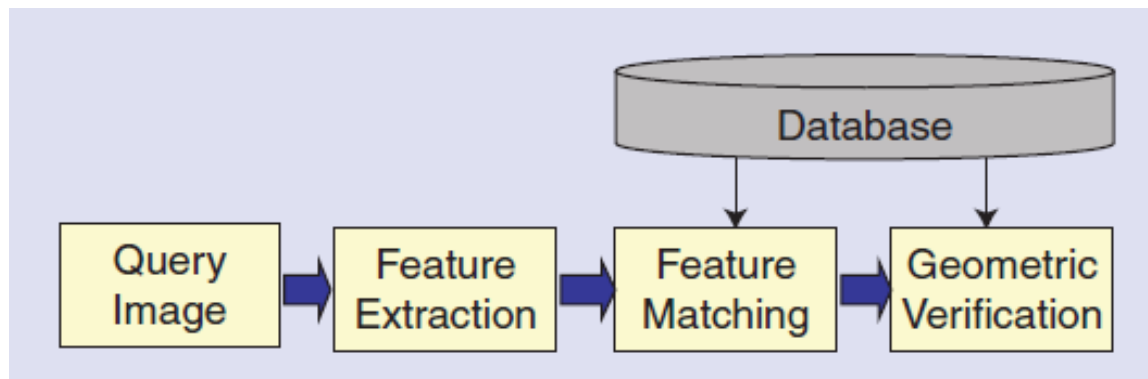


Ανάλυση / Στατιστική μοντελοποίηση / Εκπαίδευση

Περιγραφείς περιεχομένου

Χαμηλού επιπέδου (χρώμα, υφή, σχήμα) **MPEG-7**

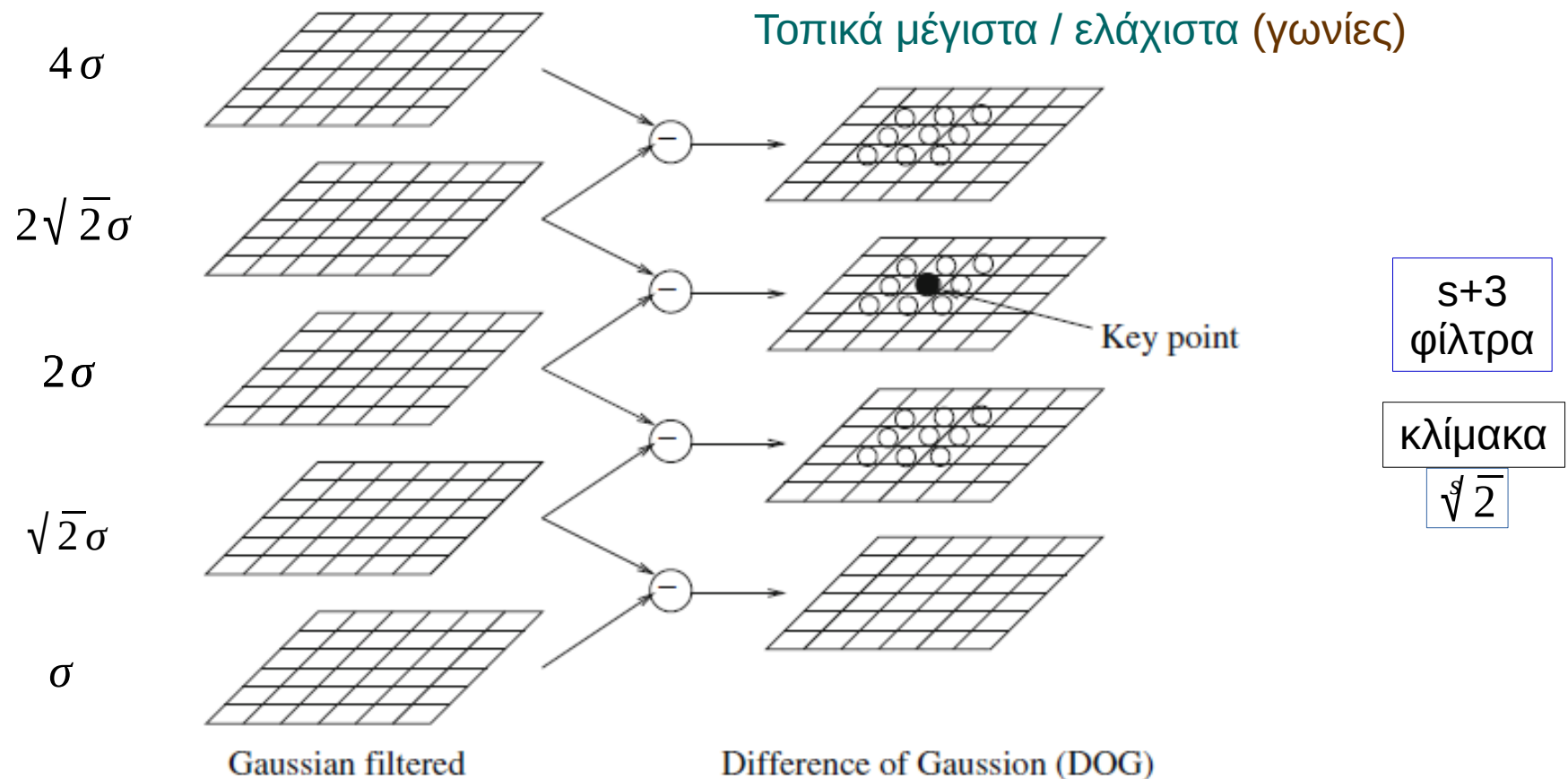
Τοπικά χαρακτηριστικά αναλλοίωτα
σε γεωμετρικούς μετασχηματισμούς



Τοπικά χαρακτηριστικά (πολλαπλές κλίμακες)

Κατασκευή πυραμίδας
πολλαπλής ανάλυσης Gauss

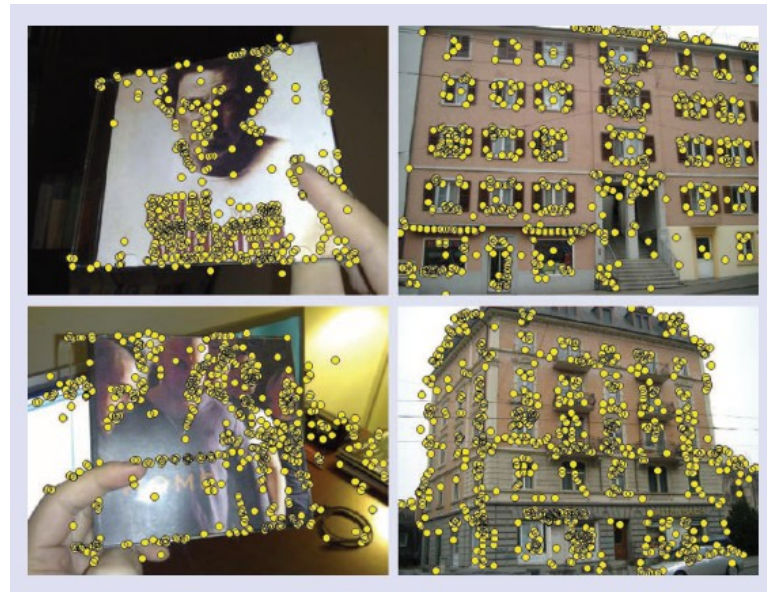
Ανίχνευση σημείων-κλειδιά
Κατεύθυνση κλίσης



Τοπικά χαρακτηριστικά (πολλαπλές κλίμακες)



Διαλογή



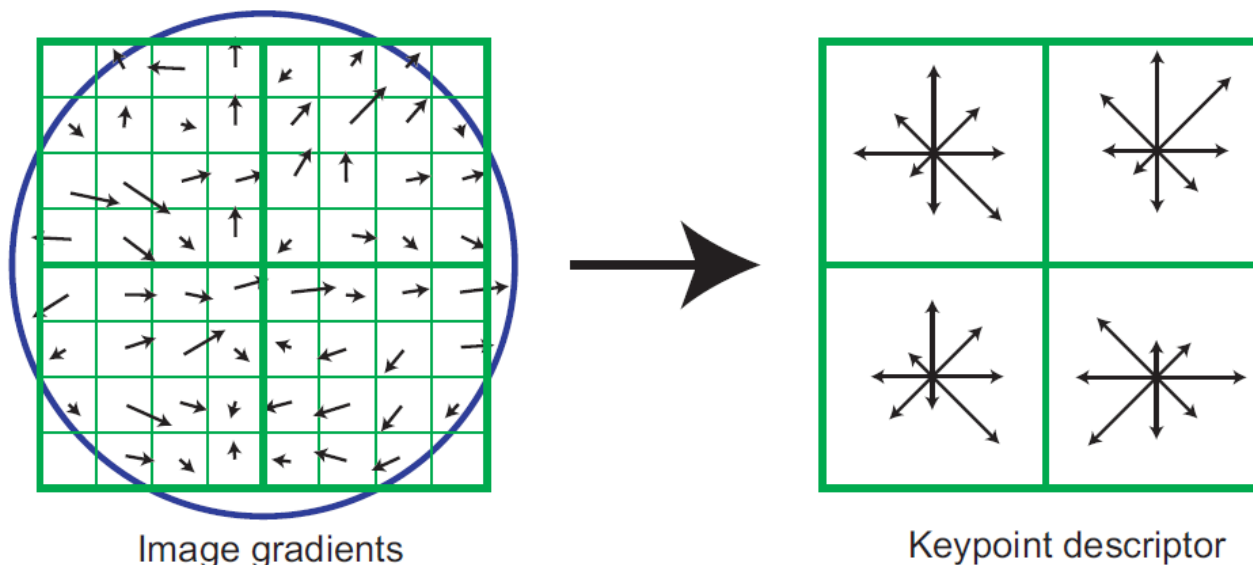
Περιγραφή χαρακτηριστικών σημείων

Ιστόγραμμα κλίσης

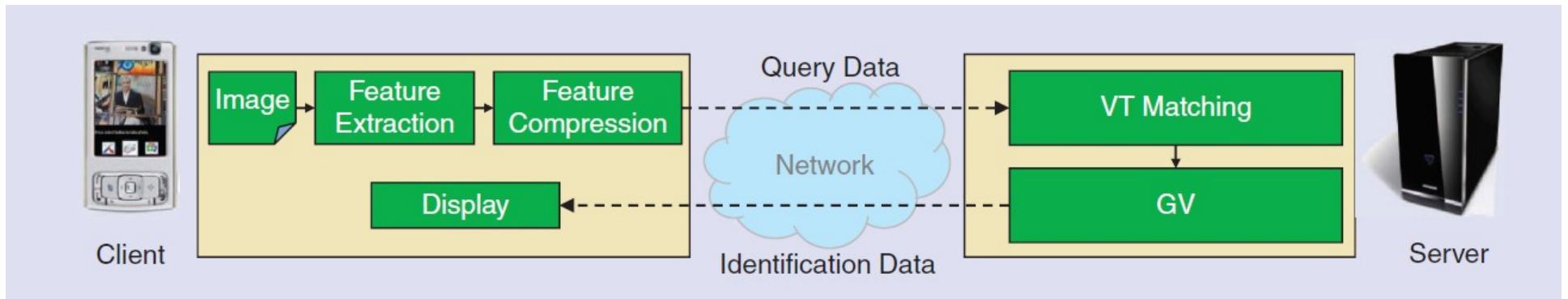
Μπλοκ 16x16 γύρω από κάθε χαρακτηριστικό σημείο

Υποδιαίρεση σε 16 υπο-μπλοκ μεγέθους 4x4

Για κάθε υπο-μπλοκ υπολογίζεται το ιστόγραμμα κατευθύνσεων της κλίσης
Σχηματίζεται διάνυσμα $8 \times 4 \times 4 = 128$ τιμών που συνιστά την περιγραφή

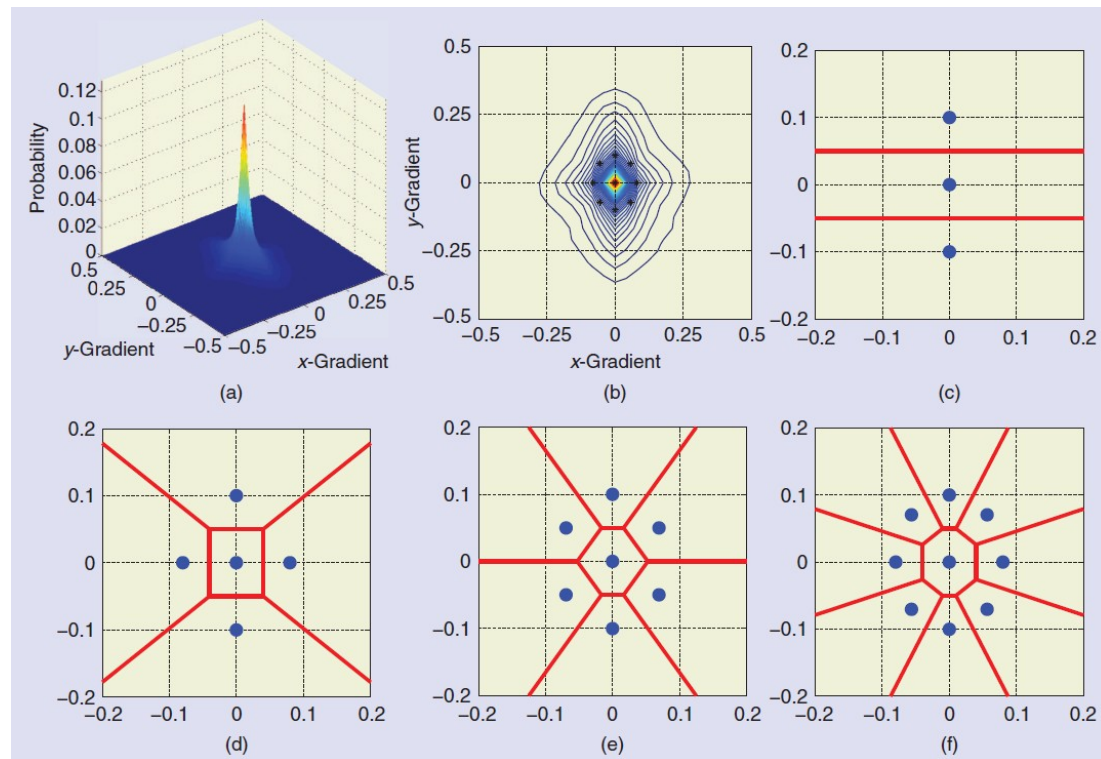


Κινητές συσκευές



Κινητές συσκευές : συμπιεσμένο ιστόγραμμα κλίσης

Υπολογισμός της κλίσης σε σημεία-κλειδιά



Χαρτοφύλακας οπτικών λέξεων

Βασισμένος σε χαρακτηριστικά SIFT

(α) για αντικείμενο ή (β) για πλαίσιο από βίντεο

Ομαδοποίηση χαρακτηριστικών SIFT για σχηματισμό

οπτικών λέξεων περιγραφής αντικειμένου ή πλαισίου

Διανυσματικός κβαντισμός για τη δημιουργία κωδικών λέξεων (βιβλίο)

Η ομαδοποίηση και η αντιστοίχιση ενός μεγάλου αριθμού

περιγραφών SIFT συνιστά υπολογιστική πρόκληση

Οι οπτικές λέξεις είναι πλούσιες σε ενθυλάκωση βασικών οπτικών

χαρακτηριστικών, παρά την αναπόφευκτη αβεβαιότητα

Ένα μικρό σε μέγεθος βιβλίο κωδικών λέξεων μπορεί να έχει περιορισμένη διακριτική ισχύ για μεγάλες βάσεις εικόνων και βίντεο.

Ένα μεγάλο σε μέγεθος βιβλίο κωδικών λέξεων

έχει επίσης προβλήματα, επειδή όμοια χαρακτηριστικά μπορεί να κβαντιστούν σε διαφορετικές κωδικές λέξεις.

Ανάκτηση εικόνων : μάθηση

Νευρωνικό δίκτυο

Μάθηση μετά από εκπαίδευση
αναπαράσταση περιεχομένου
κριτήριο ομοιότητας (κατηγορίες)

Γεφύρωση του σημασιολογικού κενού

Προκαταρκτική εκπαίδευση δικτύου σε μεγάλη βάση δεδομένων

Προσαρμογή του δικτύου σε ορισμένες κατηγορίες



Δεικτοδότηση

Αραιός πίνακας :

οι γραμμές αντιστοιχούν σε εικόνες και

οι στήλες υποδηλώνουν οπτικές λέξεις

Σε μια ανάκτηση, ελέγχονται μόνο εκείνες οι εικόνες που έχουν κοινές οπτικές λέξεις με την εικόνα του ερωτήματος

Κατακερματισμός του χώρου των χαρακτηριστικών, ώστε να

βρίσκονται σε κοντινές περιοχές του χώρου όμοιες εικόνες

Τα μεγάλης διάστασης χαρακτηριστικά κωδικοποιούνται σε

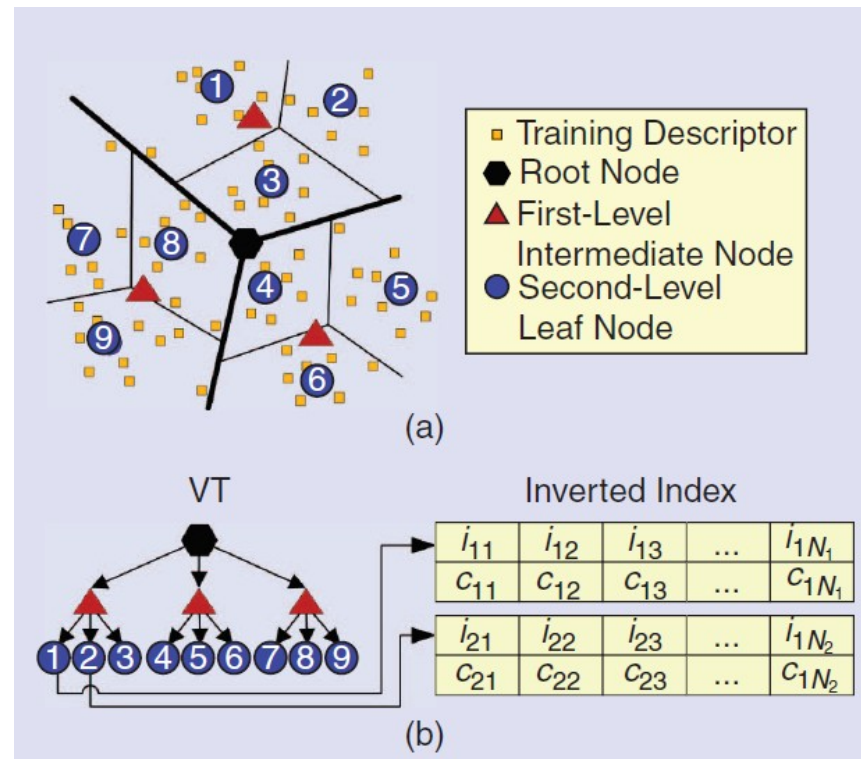
χαμηλής διάστασης δυαδικούς κώδικες για

αποτελεσματική αναζήτηση ομοιότητας

Σημασιολογικά παρόμοια δεδομένα πρέπει να έχουν κοντινούς δυαδικούς κώδικες

Δεικτοδότηση / αναζήτηση

Ιεραρχικό δένδρο λεξιλογίου
Ψηφοφορία για το ευρετήριο



Γεωμετρική επαλήθευση

Αντιστοίχιση συμβατή με μοντέλο (στερεού) γεωμετρικού μετασχηματισμού



B. Girod et al., **Mobile visual search**, *IEEE Signal Processing Magazine*, 2011

Απόσταση χαρακτηριστικών

Ομοιότητα εικόνων

$$\begin{aligned} S(\mathcal{X}, \mathcal{Y}) &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} k(x, y) \\ &= \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \phi(x)^T \phi(y) \\ &= \Psi(\mathcal{X})^T \Psi(\mathcal{Y}). \end{aligned}$$

$$D(I_q, I_m) = \left(\sum_{i=1}^N |q_i - m_i|^p \right)^{\frac{1}{p}}$$

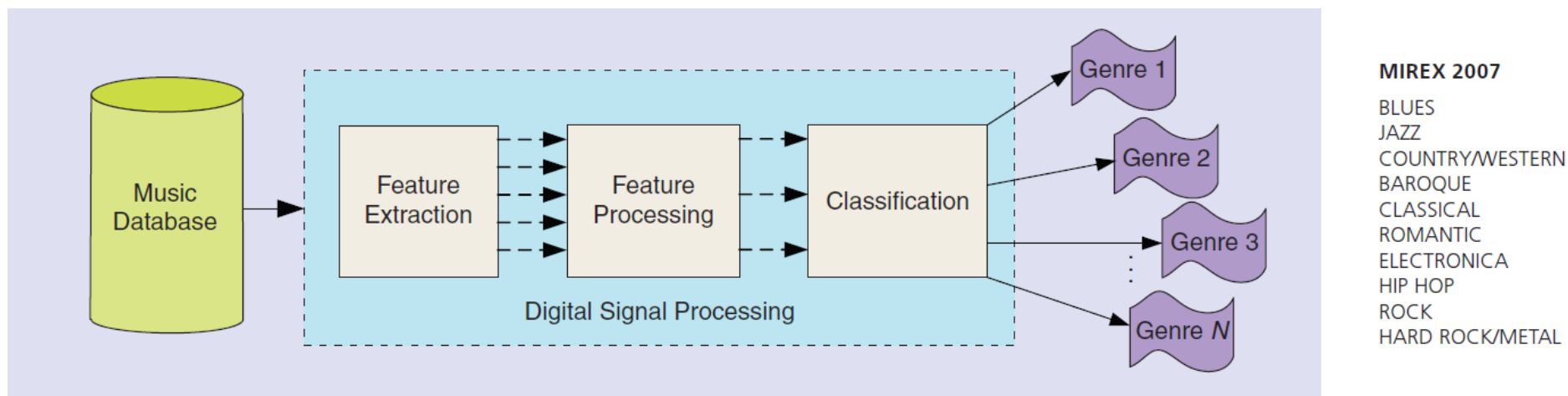
Βαρύτητα οπτικών λέξεων

Συχνότητα χαρακτηριστικού (Term frequency)

Ανεστραμμένη συχνότητα αρχείου
(Inverse document frequency)

$$\log\left(\frac{N}{nt}\right)$$

Ταξινόμηση μουσικής σε είδη



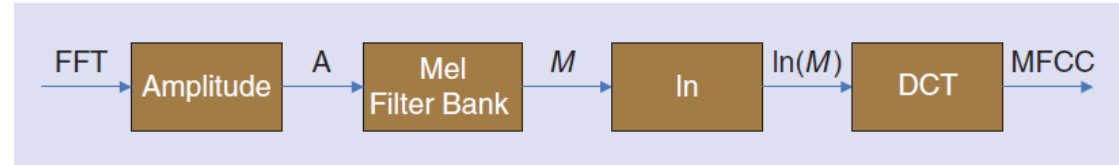
H. Blume et al., **Huge music archives on mobile devices**, *IEEE Signal Processing Magazine*, 2011.

Αναζήτηση βασισμένη σε : κείμενο, σύνοψη, παράδειγμα, μουσικό κομμάτι
Διερεύνηση σε μουσικά αρχεία

Εξαγωγή χαρακτηριστικών μουσικής

Ρυθμός αλλαγής προσήμου $r_{zc} = \frac{1}{2(N-1)} \sum_{n=0}^{N-2} |\text{sgn } x(n+1) - \text{sgn } x(n)|$

Συντελεστές συχνότητας



Θεμελιώδης συχνότητα

Διάκριση φωνής / μουσικής

