

Σύνοψη πολυμεσικού περιεχομένου

Ανίχνευση αλλαγών βίντεο
Εξαγωγή αντιπροσωπευτικών στιγμιotypών
Εξαγωγή χαρακτηριστικών αποσπασμάτων

Ανίχνευση αλλαγών ήχου
Διάκριση φωνής / μουσικής

Αλλαγή πλάνου σε βίντεο

Τύποι αλλαγής : απότομη, βαθμιαία

Τύποι βαθμιαίας αλλαγής : εξαφάνιση, εμφάνιση, διάλυση

Χαρακτηριστικά : χρώμα, κίνηση, ήχος

Αλλαγή πλάνου σε βίντεο

Ιστόγραμμα

Διαφορά

$$\sum_j |H_{t+1}(j) - H_t(j)|$$

Απόσταση χ^2

$$\sum_j \frac{(H_{t+1}(j) - H_t(j))^2}{H_{t+1}(j) + H_t(j)}$$

Τομή

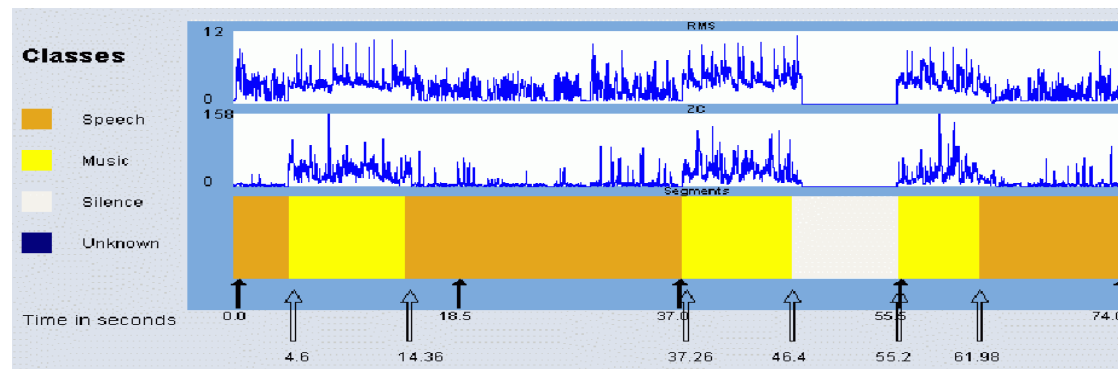
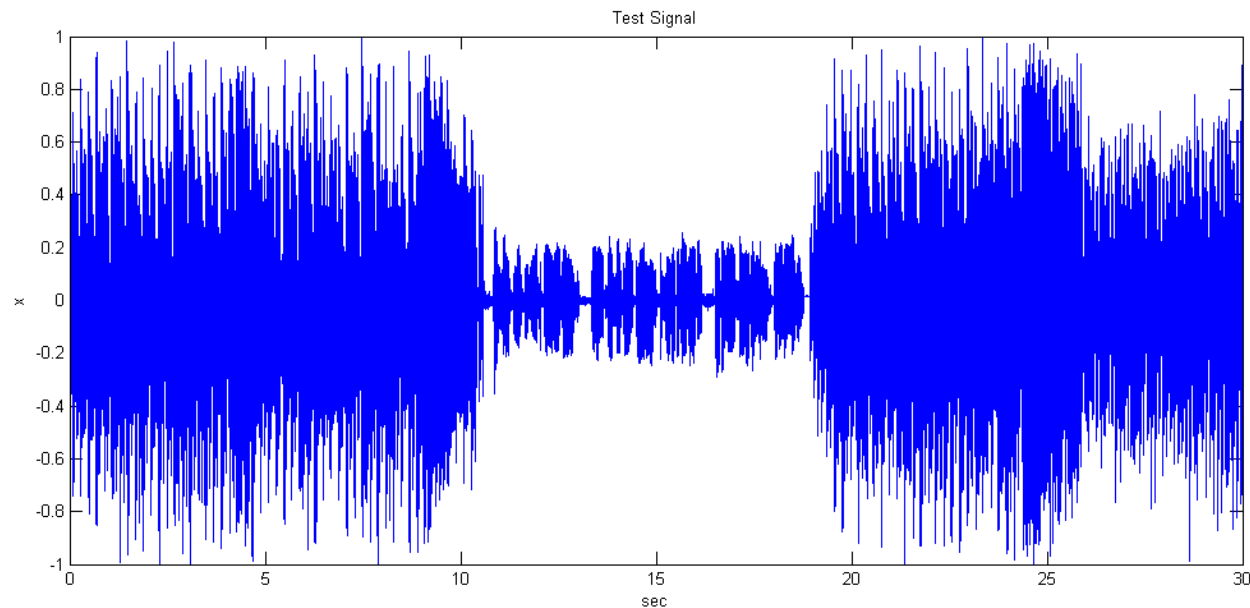
$$\sum_j \frac{\min(H_{t+1}(j), H_t(j))}{\max(H_{t+1}(j), H_t(j))}$$

Συμπιεσμένο βίντεο

Διανύσματα μετακίνησης

Ιστογράμματα συντελεστών DCT

Αλλαγή ηχητικού περιεχομένου

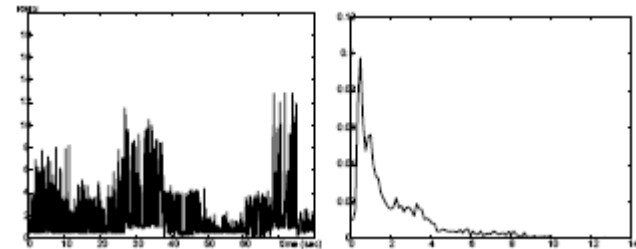
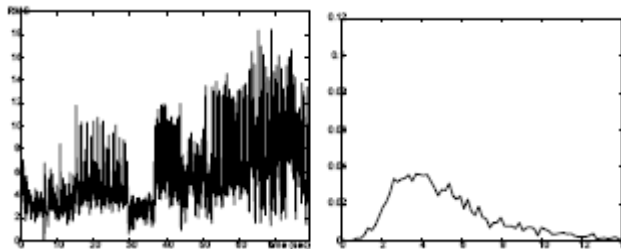


Αλλαγή ηχητικού περιεχομένου

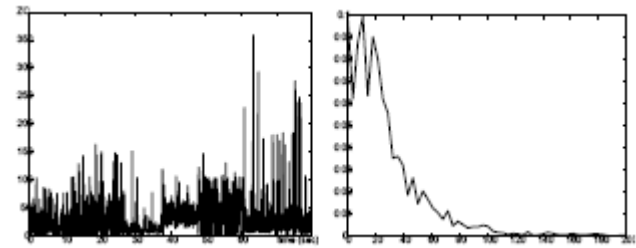
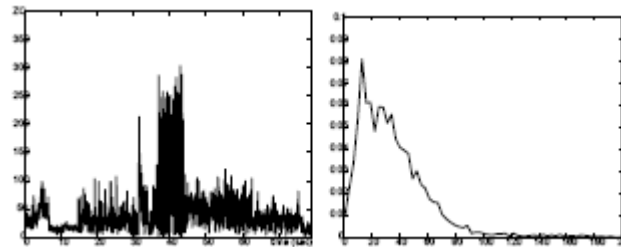
Μουσική

Φωνή

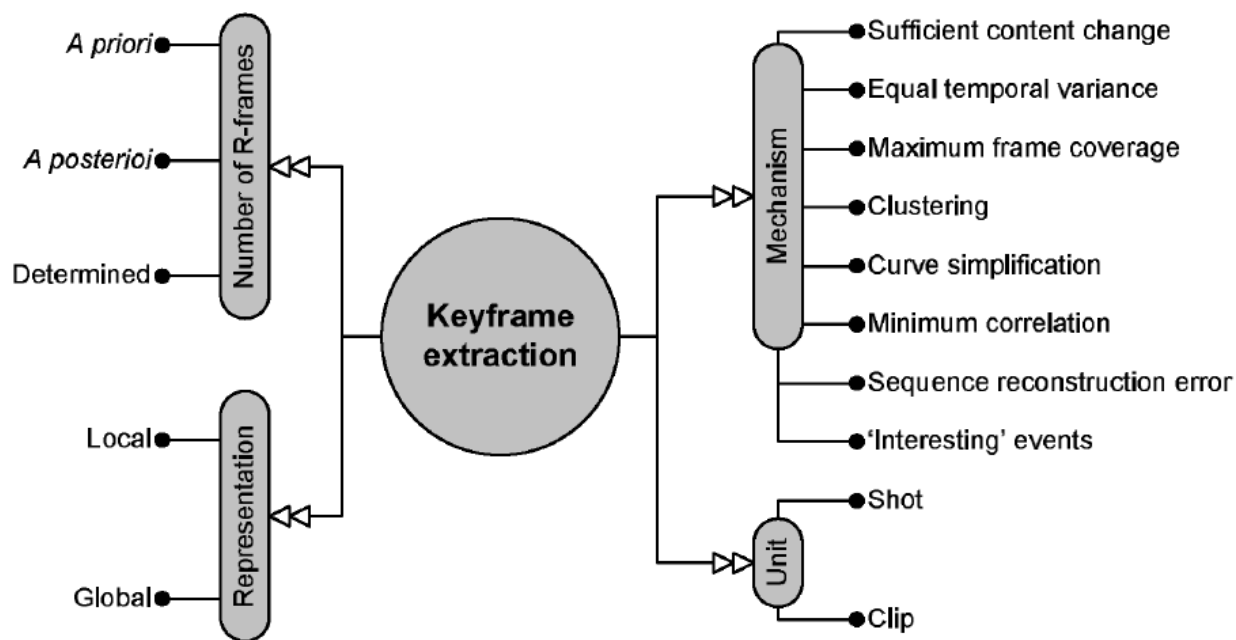
RMS



ZC



Εξαγωγή χαρακτηριστικών στιγμιότυπων



B.T. Truong and S. Venkatesh, **Video abstraction: a systematic review and classification**, ACM Trans. on Multimedia Computing, Communications and Applications, Feb. 2007.

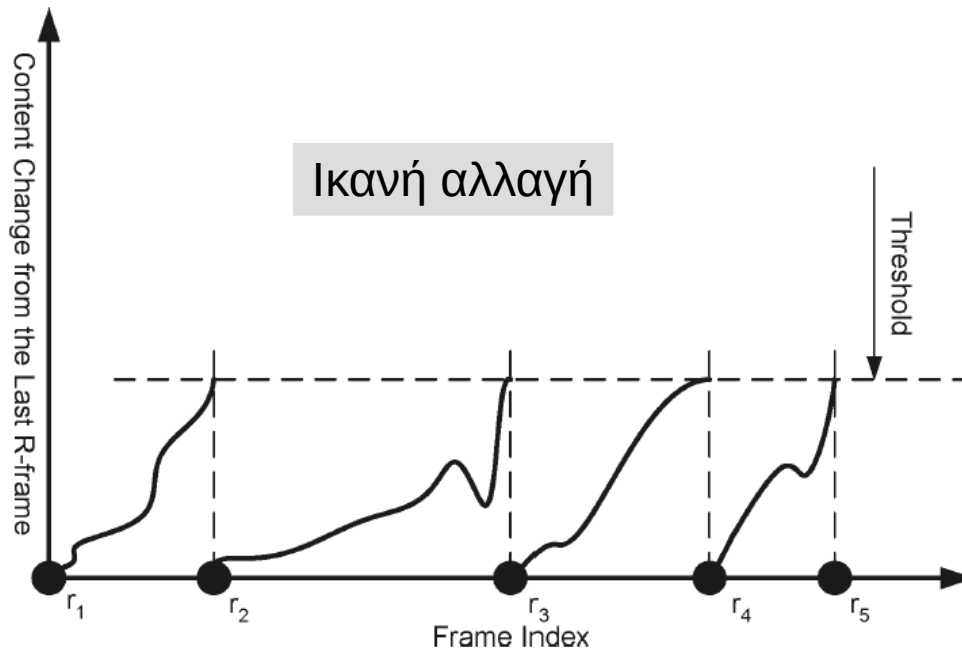
Εξαγωγή χαρακτηριστικών στιγμιότυπων

Μικρή δραστηριότητα κίνησης

Ανομοιότητα μεταξύ πλαισίων



Αλλαγή πλάνου και εξαγωγή στιγμιότυπων



$$r_i = \arg \min_t \{C(f_t, f_{b_i}) > \varepsilon, t > b_i\}$$
$$b_{i+1} = \arg \min_t \{C(f_t, f_{r_i}) > \varepsilon, t > r_i\}.$$

$$\{b_1, b_2, \dots, b_{k+1}\} = \arg \min_{b_i} \sum_{i=1}^k \sum_{j=1}^k |\mathcal{V}(b_i, b_{i+1}) - \mathcal{V}(b_j, b_{j+1})|$$

2ο βήμα : αντιπροσωπευτικά πλαίσια

Εξαγωγή στιγμιότυπων

Μέγιστη κάλυψη

ελάχιστος
αριθμός

$$\{r_1, r_2, \dots, r_k\} = \arg \min_{r_i} \{\mathbf{k} | \mathbf{C}_{r_1}(\varepsilon) \cup \mathbf{C}_{r_2}(\varepsilon) \cup \dots \cup \mathbf{C}_{r_k}(\varepsilon) = \mathbf{V}\}$$

ορισμένος
αριθμός

$$\{r_1, r_2, \dots, r_k\} = \arg \min_{r_i} \{\varepsilon | \mathbf{C}_{r_1}(\varepsilon) \cup \mathbf{C}_{r_2}(\varepsilon) \cup \dots \cup \mathbf{C}_{r_k}(\varepsilon) = \mathbf{V}\}$$

$$\{r_1, r_2, \dots, r_k\} = \arg \max_{r_i} \{|\mathbf{C}_{r_1}(\varepsilon) \cup \mathbf{C}_{r_2}(\varepsilon) \cup \dots \cup \mathbf{C}_{r_k}(\varepsilon)|\}$$

Τα αντιπροσωπευόμενα σύνολα μπορεί να περιλαμβάνουν ασυνέχειες

Εξαγωγή στιγμιότυπων

Ομαδοποίηση πλαισίων

Προεπεξεργασία για μείωση της διάστασης των δεδομένων

Ομαδοποίηση (ιεραρχική ή σωρευτική)

Φιλτράρισμα ομάδων

Επιλογή αντιπροσωπευτικών πλαισίων (κέντρα ομάδων)

Ελάχιστη συσχέτιση

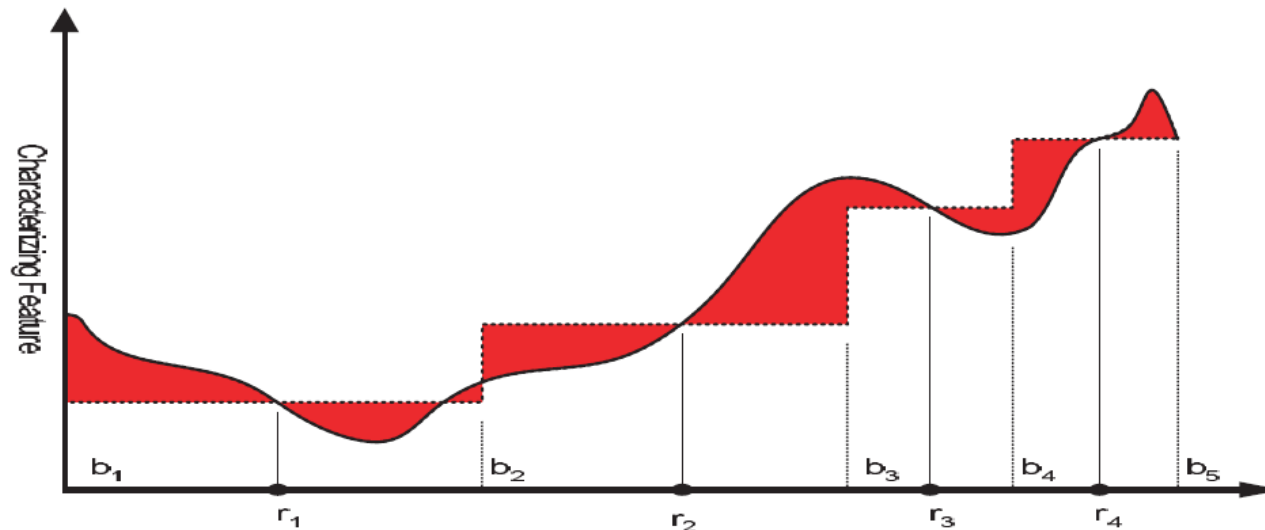
$$\{r_1, r_2, \dots, r_k\} = \arg \min_{r_i} \{Corr(f_{r_1}, f_{r_2}, \dots, f_{r_k})\}$$

μόνο διαδοχικά

$$\{r_1, r_2, \dots, r_k\} = \arg \min_{r_i} \left\{ \sum_{i=1}^{k-1} Corr(f_{r_i}, f_{r_{i+1}}) \right\}$$

Εξαγωγή στιγμιότυπων

Σφάλμα αποκατάστασης



$$\mathcal{E}(\mathbf{V}, \mathcal{R}) = \sum_{i=1}^n \mathcal{D}(f_i, \mathcal{I}(i, \mathcal{R}))$$

παρεμβολή για τα χαρακτηριστικά

Εξαγωγή στιγμιότυπων

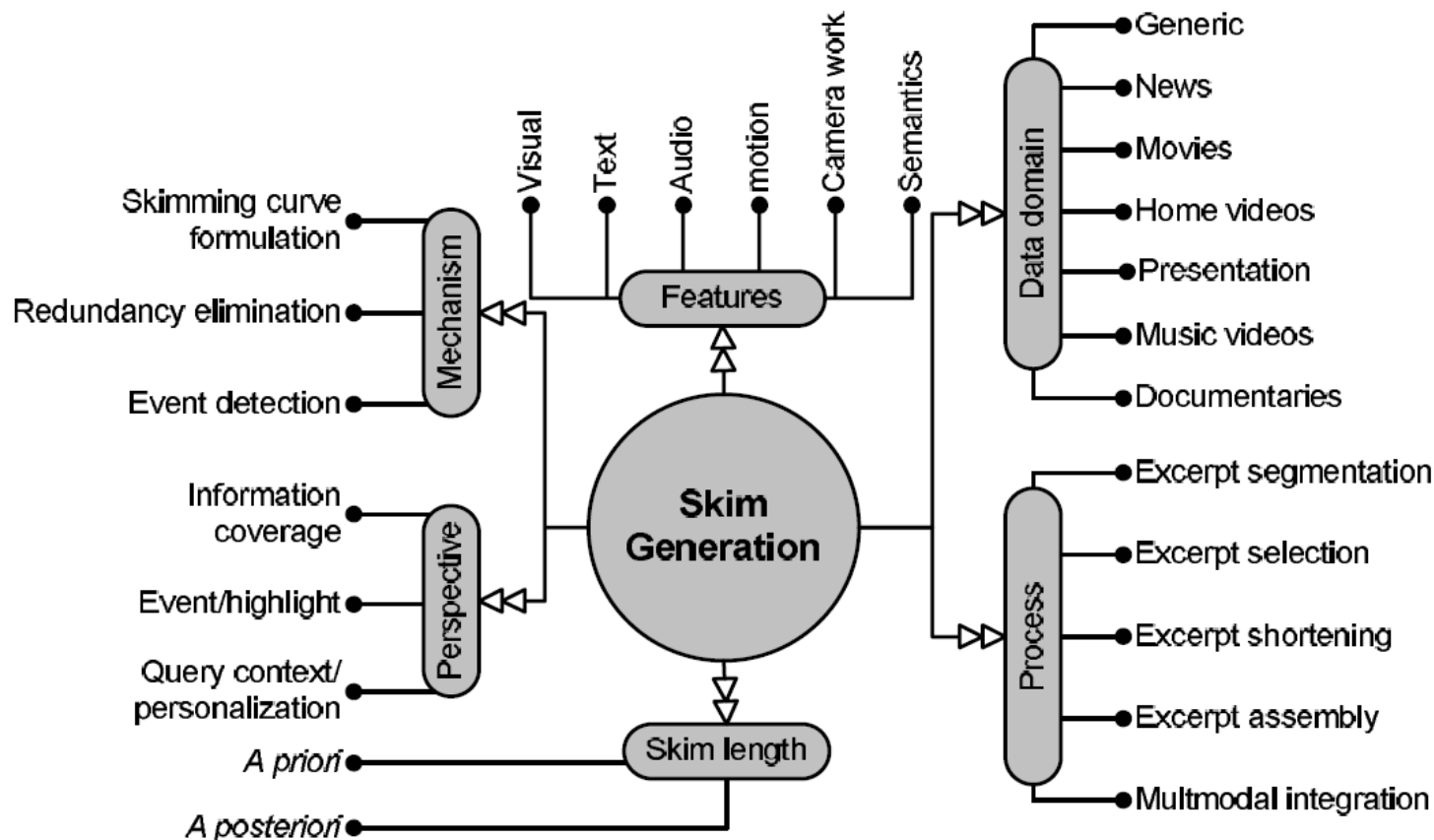
“Ενδιαφέροντα” συμβάντα

Δραστηριότητα κίνησης (εναλλαγές)

Ακινησία, πλούσιο χωρικό περιεχόμενο, παρουσία προσώπων

Μικρή διασπορά σε κοντινό ορίζοντα και διαφορά από απομακρυσμένα

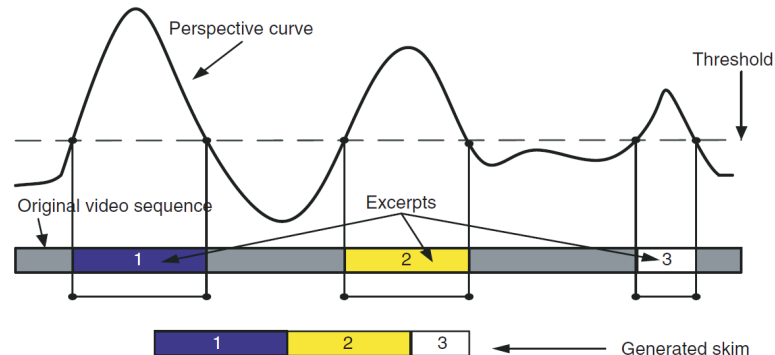
Εξαγωγή χαρακτηριστικών αποσπασμάτων



Εξαγωγή χαρακτηριστικών αποσπασμάτων Σκοπός και τεχνικές

Πληροφοριακός / συντόμευση
Σημαντικά / ενδιαφέροντα συμβάντα
Πλαίσιο αναζήτησης περιεχομένου / εξατομίκευση

Αφαίρεση πλεονάζουσας πληροφορίας
Ανίχνευση κορυφαίων στιγμών / συμβάντων
Υπολογισμός ενδιαφέροντος (ορισμός)



Εξαγωγή χαρακτηριστικών αποσπασμάτων

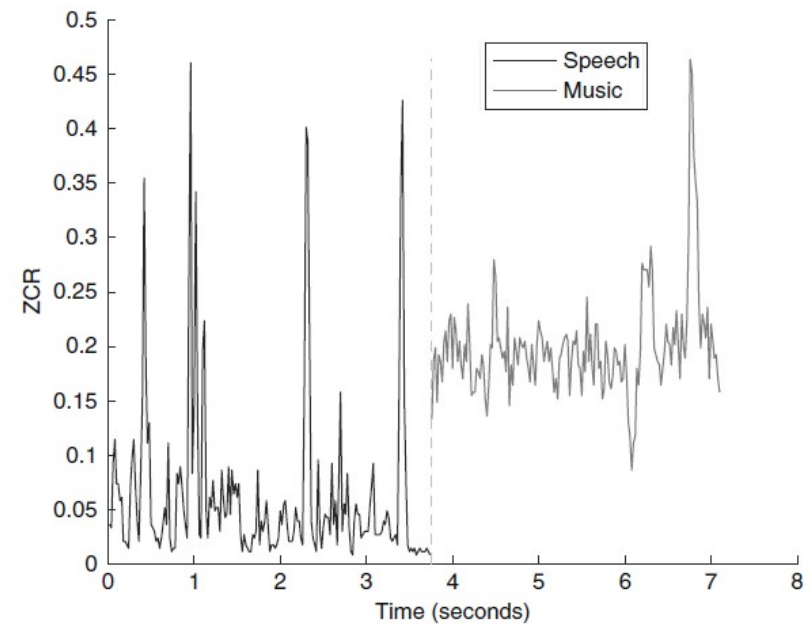
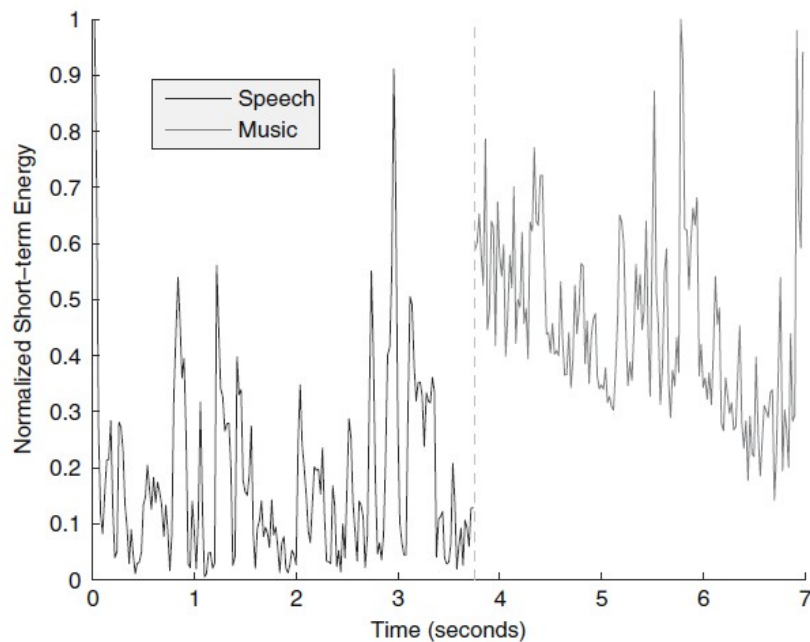
Περιγραφείς και διαδικασία

Επικρατούντα χρώματα, υφή με χωρική κατανομή
Ιστογράμματα χρωμάτων
Μετρήσεις χρωματικών αντιθέσεων
Ανίχνευση / αναγνώριση κειμένου
Ανίχνευση προσώπων
Δραστηριότητα
Κίνηση κάμερας

Ηχητικά εφέ
Χαρακτηριστικοί ήχοι
Ανίχνευση / αναγνώριση φωνής

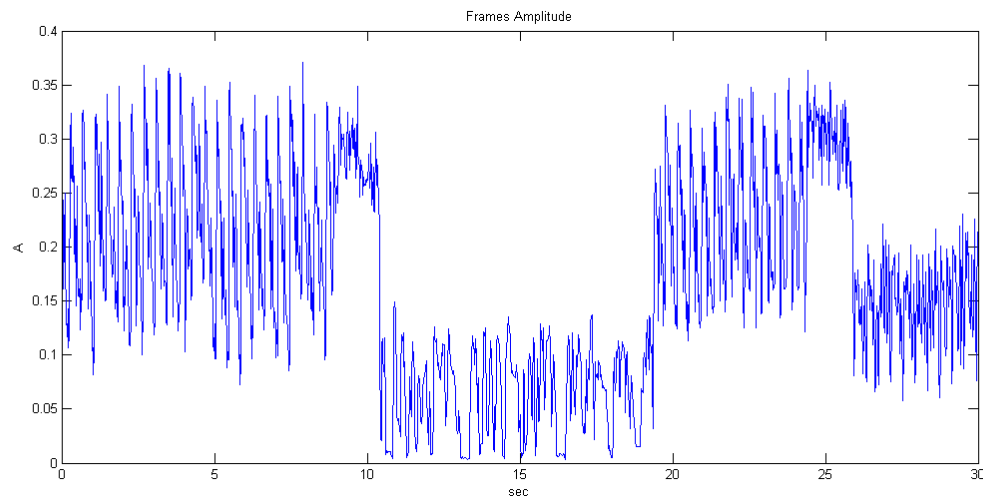
Τμηματοποίηση (πλάνα, σκηνές)
Επιλογή αποσπασμάτων
Συντόμευση αποσπασμάτων
Συναρμογή

Διάκριση φωνής / μουσικής

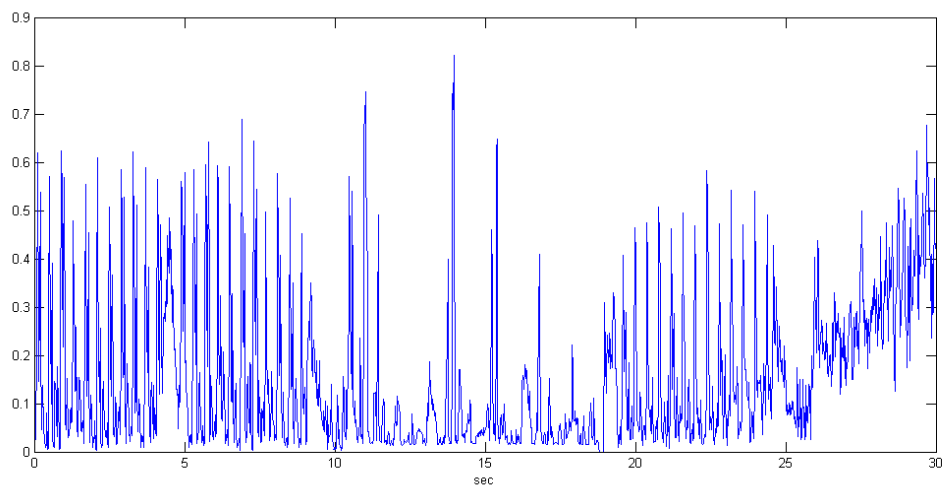


A. Pikrakis, et al, **An overview of speech/music discrimination techniques in the context of audio recordings**, Springer, 2008.

Διάκριση φωνής / μουσικής



Μέσο πλάτος



Μέσος ρυθμός
εναλλαγής προσήμου

Διάκριση φωνής / μουσικής

