

ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ

ΤΜΗΜΑ ΕΠΙΣΤΗΜΗΣ ΥΠΟΛΟΓΙΣΤΩΝ

ΠΑΡΟΥΣΙΑΣΗ / ΕΞΕΤΑΣΗ ΜΕΤΑΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Χωνιανάκης Σταύρος

Μεταπτυχιακός Φοιτητής

Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

Επόπτης Μεταπτ. Εργασίας: Αναπλ. Καθηγητής Αθανάσιος Μουχτάρης

Παρασκευή, 20/1/2017, 10:00

Αίθουσα B108, Τμήμα Επιστήμης Υπολογιστών, Πανεπιστήμιο Κρήτης

**“Αυτόματη οργάνωση περιεχομένου προερχόμενου από πληθοπωρισμό βάσει
συσχετίσεων στο ηχητικό υλικό”**

ΠΕΡΙΛΗΨΗ

Με τη συνεχή αύξηση των smart-phones και των φορητών ηλεκτρονικών συσκευών, όλο και περισσότεροι από εμάς ασχολούμαστε με τη διαδικασία καταγραφής οπτικοακουστικού υλικού από δημόσιες εκδηλώσεις που παρακολουθούμε. Αυτό το οπτικοακουστικό υλικό (User Generated Content ή εν συντομία UGC) δεν έχει μόνο αξία για αυτόν που το καταγράφει, αλλά είναι χρήσιμο και για τους ραδιοτηλεοπτικούς παραγωγούς και διαχειριστές πολυμέσων που σχετίζονται με την επαγγελματική κάλυψη αυτών των γεγονότων. Για παράδειγμα, το UGC μπορεί να εμπλουτίσει το επαγγελματικής προέλευσης υλικό με διαφορετικές γωνίες λήψης αλλά και να παρέχει κάλυψη για μέρη της εκδήλωσης που δεν έχουν συλληφθεί από τον επαγγελματικό φακό. Ωστόσο, χρειάζεται αρκετή προεργασία ώστε αυτό το περιεχόμενο να μπορεί να είναι αξιοποιήσιμο για το σκοπό που αναφέραμε. Για παράδειγμα, καθώς το UGC δεν διαθέτει ακριβή πληροφορία σχετικά με τη θέση και την ώρα εγγραφής, θα απαιτείτο πολύ μεγάλος χρόνος και προσπάθεια από ένα χρήστη να αναζητήσει συγκεκριμένο

περιεχόμενο ή από ένα παραγωγό να εντοπίσει και να συγχρονίσει έναν αριθμό από βίντεο τα οποία επικαλύπτονται χρονικά.

Ευτυχώς, η αυτόματη οργάνωση τέτοιων συλλογών είναι εφικτή αξιοποιώντας τις συσχετίσεις που υπάρχουν στο ηχητικό υλικό που προέρχεται από διαφορετικές καταγραφές. Αυτή η διεργασία είναι δε ιδιαίτερα αποτελεσματική όταν αντί για το ηχητικό σήμα αυτό καθεαυτό οι συσχετίσεις εξετάζονται με βάση το ηχητικό «αποτύπωμα» (fingerprint) που εξάγεται από κάθε ηχητική καταγραφή. Τα ηχητικά αποτυπώματα έχουν ως στόχο να κωδικοποιήσουν τα μοναδικά χαρακτηριστικά ενός αρχείου ήχου, παρέχοντας σημαντική ευρωστία στο θόρυβο και την παραμόρφωση.

Σε αυτήν την εργασία χρησιμοποιούμε τις συσχετίσεις μεταξύ ηχητικών αποτυπωμάτων για να εντοπίσουμε και να συγχρονίσουμε αρχεία ήχου που προέρχονται από το ίδιο ηχητικό γεγονός. Πραγματοποιούμε μια εκτενή σύγκριση ανάμεσα σε διαφορετικές τεχνικές ηχητικής αποτύπωσης που έχουν προταθεί στο παρελθόν αλλά προτείνουμε και μια νέα τεχνική, αναδεικνύοντας τα οφέλη που αυτή παρέχει στην οργάνωση του περιεχομένου που προέρχεται από αθλητικό γεγονός. Προτείνουμε μια διεργασία για τον εντοπισμό των συσχετίσεων μεταξύ ηχητικών αποτυπωμάτων που είναι εύκολο να γενικευθεί και παρουσιάζουμε μια προσέγγιση εμπνευσμένη από τη θεωρία γράφων για τον εντοπισμό και την ομαδοποίηση αρχείων ήχου που επικαλύπτονται χρονικά.

Chonianakis Stavros

M.Sc. Thesis

Computer Science Department

University of Crete

Master's Thesis Supervisor: Associate Professor A. Mouchtaris

Friday, 20/1/2017, 10:00

Room B108, Computer Science dept., University of Crete

“Automatic organization of user generated content based on audio correlations”

ABSTRACT

With the proliferation of smart-phones and portable electronic devices, more and more of us become engaged in the process of capturing and sharing audiovisual content from public events that we attend. Such User Generated Content (UGC) can be very valuable to the broadcasters and producers associated to the professional coverage of such events,

as it may enrich the footage or provide coverage for parts of the event that have not been captured by the professional equipment. Yet, it is not trivial to organize this content in a way that it can be usable for the said purpose. For example, as user generated content lacks metadata which is informative about the exact location and time of recording, it would require enormous time and effort from the professional editor to manually search for videos referring to a particular segment of the event, or to group and temporally align multiple videos overlapping in space and time.

Fortunately, as several works have demonstrated in the past, it is possible to automatically organize such content by exploiting the correlations in the audio streams available in the UGC. Rather than working with the raw audio data, such correlations are much more efficiently revealed based on fingerprints extracted from each user generated audio recording. Audio fingerprinting finds use as the means for providing a compact and concrete content based signature, by retaining the maximum acoustically relevant information, showing significant robustness to variations with respect to the audio format, the induced noise and distortion in each audio recording.

In this thesis, we use audio fingerprint cross-correlation as the means to detect and synchronize temporally overlapping user generated audio recordings of the same event and we evaluate our tools based on two datasets that we have acquired ourselves; a musical concert and a football match. We perform an extensive comparison based on different state of the art fingerprinting techniques and we propose a novel fingerprinting algorithm with significantly better organization performance for the case of the athletic event. We propose a generalizable scheme for fingerprint cross-correlation and we present an approach inspired by graph theory for clustering recordings from the same temporal segment of the public event into the same group.