A framework for managing multimodal digitized music collections

Πλαίσιο διαχείρισης πολύμορφων ψηφιακών μουσικών συλλογών

ΙΟΝΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΕΡΚΥΡΑΣ

ΤΜΗΜΑ ΑΡΧΕΙΟΝΟΜΙΑΣ ΚΑΙ ΒΙΒΛΙΟΘΗΚΟΝΟΜΙΑΣ

ΨΗΦΙΑΚΕΣ ΒΙΒΛΙΟΘΗΚΕΣ



**ΥΠΕΥΘΥΝΟΣ ΚΑΘΗΓΗΤΗΣ**: ΣΑΡΑΝΤΟΣ ΚΑΠΙΔΑΚΗΣ

**ΕΙΣΗΓΗΤΡΙΑ ΕΡΓΑΣΙΑΣ:** ΚΑΤΕΡΙΝΑ ΛΑΜΠΡΟΥ

Περίληψη

Στο έγγραφο αυτό παρουσιάζεται ένα πλαίσιο διαχείρισης ετερογενών –πολύμορφων ψηφιακών μουσικών συλλογών, που περιέχουν οπτικές μουσικές απεικονίσεις ( ανιχνευμένη μουσική φύλλων , ακουστικές καταγραφές ). Ως πρώτη συμβολή , προτείνεται η επεξεργασία, περιλαμβάνοντας την εξαγωγή γνωρισμάτων , την ακουστική ευρετηρίαση και τον συγχρονισμό της μουσικής ( συνδέοντας τα οπτικά με τα ακουστικά ψηφιακά δεδομένα). Ως δεύτερη συμβολή εισάγεται η διεπαφή του χρήστη στην πολύμορφη μουσική απεικόνιση , στην πλοήγηση και στην βασισμένη στο περιεχόμενο ανάκτηση. Το σύστημα αυτό, προσφέρει υψηλή ποιότητα αναπαραγωγής ήχου και χρονικο-συγχρονισμό της ψηφιακής μουσικής φύλλων. Το σύστημα επιτρέπει στο χρήστη να επιλέγει περιοχές μέσα στις ανιχνευμένες σελίδες ενός μουσικού αποτελέσματος, ώστε να βρεθούν παρόμοια μουσικά τμήματα στα ακουστικά έγγραφα. Η διεπαφή του χρήστη και οι λειτουργίες αναζήτησης θα ενσωματωθούν στο σύστημα υπηρεσιών της Εθνικής Βαυαρικής βιβλιοθήκης, ως μέρος του Probado προγράμματος .

Εισαγωγή

Πρόσφατα σημαντικές προσπάθειες ψηφιοποίησης έχουν πραγματοποιηθεί για μεγάλες συλλογές βιβλίων και άλλων έντυπων εγγραφών. Αυτές οι προσπάθειες οδηγούν στην ανάγκη ισχυρών εργαλείων τα οποία επεξεργάζονται, αναλύουν και σχολιάζουν τα ανιχνευμένα έγγραφα, παρέχοντας τη βάση για αποδοτική και αποτελεσματική έρευνα περιεχομένου , πλοήγηση και αναζήτηση στα ψηφιακά δεδομένα. Στην περίπτωση των ανιχνευμένων κειμένων, έχουν προταθεί ποικίλες λύσεις για τη επεξεργασία αυτοματοποιημένων εγγραφών, οι οποίες περιέχουν ένα συστατικό γνώρισμα για την αναγνώριση οπτικού χαρακτήρα, OCR , ώστε να εξάγει το περιεχόμενο του κειμένου από τις εικόνες . Τα πρόσφατα συστήματα μπορούν να αντιγράψουν με πιθανή εξαγωγή λάθους, τα οποία οφείλονται στο επιρρεπές σε λάθη OCR (Optical Character Recognition). Επιπλέον τα συστήματα αυτά, παρουσιάζουν υψηλής ποιότητας ανίχνευση στα ενδιαφερόμενα κείμενα . Το OCR εξάγει αποτελέσματα μόνο στην επεξεργασία εκ των έσω , δηλαδή στα στάδια ανάλυσης και ανάκτησης . Η γενική ιδέα είναι να συνδυαστούν κατάλληλα οι δυνάμεις και των δύο τύπων δεδομένων –απεικονίσεων , ( ανίχνευση και κείμενο ) για την πιο άνετη πλοήγηση στα ανιχνευμένα έγγραφα. Ένα πολύ γνωστό παράδειγμα είναι το Google Book search , project. Παρόλο την πρόοδο στην περιοχή των κειμένων, υπάρχει μια σημαντική έλλειψη στη διαχείριση γενικού ψηφιακού υλικού συμπεριλαμβανομένων εικόνων , βίντεο 3D γραφικών δεδομένων ή audio δεδομένων. Συγκεκριμένα, απαιτούνται εργαλεία τα οποία θα εξάγουν αυτόματα σημαντικές σημασιολογικές οντότητες από τα ανιχνευμένα έγγραφα, δημιουργώντας συνδέσεις μεταξύ σχετικών οντοτήτων.

Αυτό το έγγραφο εστιάζει στο ιδιαίτερο σενάριο των ετερογενών μουσικών συλλογών που περιέχουν ψηφιακή μουσική φύλλων (δεδομένα εικόνας ) και audio δεδομένα. Για κάθε μουσικό έγγραφο στη συλλογή είναι διαθέσιμο ένα πλήρες σύνολο ανιχνευμένων σελίδων του μουσικού αποτελέσματος και τουλάχιστον μια ακουστική καταγραφή ενός μουσικού έργου. Τα έγγραφα αυτά, έχουν οπτική και ακουστική μορφή, ορίζοντας την πολυμορφία τους . Μια πολύμορφη μουσική συλλογή περιλαμβάνεται από τρία επίπεδα.

*Επίπεδο αποθετηρίου ( repository layer)*

Αυτό το επίπεδο αποτελείται από τα ψηφιακά μουσικά δεδομένα (ανιχνευμένη μουσική φύλλων και ακουστική καταγραφή). Επιπλέον, αυτό το επίπεδο περιλαμβάνει τα εργαλεία που αναλύουν αυτόματα και τους δύο τύπους δεδομένων . Συγκεκριμένα, αυτά τα εργαλεία χρησιμοποιούνται για την εξαγωγή γνωρισμάτων ( OMR και audio επεξεργασία) , για την ακουστική ευρετηρίαση , τον συγχρονισμό (ευθυγράμμιση ) μουσικής φύλλων και τα ακουστικά δεδομένα. Στόχος και αντικείμενο της επεξεργασίας , αποτελεί η σύνδεση των 2D περιοχών ( μετρήσιμες σε pixels) μέσα στις ανιχνευμένες σελίδες , με τα αντίστοιχα σημασιολογικά χρονικά τμήματα (μετρήσιμα σε δευτερόλεπτα) μέσα σε μια ακουστική καταγραφή.

*Επίπεδο διεπαφής χρήστη*

Αυτό το επίπεδο περιλαμβάνει την πολύμορφη μουσική πρόσβαση και την άμεση αλληλεπίδραση που προκύπτει. Πέρα από την προσφορά τυποποιημένων λειτουργιών για την ακουστική αναπαραγωγή ήχου , την διαχείριση αρχείων και τους ελέγχους αναπαραγωγής ήχου, η διεπαφή Viewer – Result, παρουσιάζει συγχρόνως τις μουσικές αντίστοιχες περιοχές της ανιχνευμένης μουσικής φύλλων κατά την διάρκεια της ακουστικής αναπαραγωγής ήχου. Η διεπαφή μπορεί να χρησιμοποιηθεί με το marking των περιοχών μέσα στα ψηφιακά αποτελέσματα , με την χρήση του ποντικιού , όπου αργότερα θα τεθεί ως query, ώστε να ξεκινήσει η αναζήτηση μουσικής με βάση το περιεχόμενο. Τα ανακτώμενα ακουστικά έγγραφα παρουσιάζονται στη διεπαφή Viewer- Result , η οποία παρέχει τις κατάλληλες λειτουργίες για την πλοήγηση μέσα σε αυτά τα έγγραφα.

*Επίπεδο server*

Αυτό το επίπεδο συνδέει το επίπεδο αποθετηρίου με το επίπεδο διεπαφής χρήστη .Το επίπεδο Server έχει άμεση πρόσβαση σε όλα τα δεδομένα που περιλαμβάνονται και παράγονται στο επίπεδο αποθετηρίου καθώς επίσης, λαμβάνει και χειρίζεται τα ποικίλα αιτήματα που τίθενται στη διεπαφή χρήστη. Το επίπεδο Server είναι ο κεντρικός υπολογιστής της μουσικής Probado. Ο στόχος του είναι να χειριστεί την επικοινωνία με το επίπεδο διεπαφής χρήστη και να σχεδιάσει τις εισερχόμενες ερωτήσεις σε ένα σύνολο διαθέσιμων μηχανών ερώτησης . Κάθε μηχανή ερώτησης προσφέρει μια ιδιαίτερη λειτουργία , πχ. Η ανάκτηση βάση περιεχομένου χρησιμοποιεί την ακουστική αντιστοιχία ή την μεταφορά δεδομένων μέσο streaming.

*Σημείωση*

Η επεξεργασία στο επίπεδο αποθετηρίου γίνεται offline , ενώ στα επίπεδα server και διεπαφής χρήστη γίνεται online.

Αυτόματη επεξεργασία εγγραφών

Σε αυτό το τμήμα περιγράφονται σημαντικές μεθόδους που απαιτούνται για την επεξεργασία, την αντιστοιχία και την ευθυγράμμιση των ποικίλων τύπων μουσικών δεδομένων . Η βασική ιδέα είναι να μετασχηματιστούν οι ανιχνευμένες εικόνες και οι ακουστικές καταγραφές σε μια κοινή απεικόνιση γνωρισμάτων , η οποία θα επιτρέπει την άμεση σύγκριση των δύο διαφορετικών τύπων δεδομένων . Τα γνωρίσματα μουσικής βασισμένα στο χρώμα έχουν αποδειχθεί μια mid-level απεικόνιση μουσικής. Στο σύστημα αυτό, τα εξαγόμενα γνωρίσματα οργανώνονται από ένα κατάλληλο ευρετήριο δομής, το οποίο μπορεί να χρησιμοποιηθεί στην απόδοση αντιστοιχίας μουσικών έργων. Επιπλέον εισάγεται ένας νέος μηχανισμός , ο οποίος επιτρέπει τον προσδιορισμό και τον σχολιασμό ανιχνευμένων σελίδων της μουσικής φύλλων από το διαθέσιμο ακουστικό υλικό . Συνοψίζεται η διαδικασία συγχρονισμού της μουσικής ώστε να γενικοποιηθούν οι ευθυγραμμίσεις που απαιτούνται για την οπτικοποίηση στην σύγχρονη αναπαραγωγή ήχου στη διεπαφή χρήστη.

Χαρακτηριστικά γνωρίσματα μουσικής βασισμένα στο Χρώμα

Για να συγκριθούν και να συσχετιστούν τα ποικίλα μουσικά δεδομένα και σχήματα, χρειάζεται μια κατάλληλη απεικόνιση γνωρισμάτων ώστε να ικανοποιηθούν οι ανάγκες των διαφορετικών μουσικών δεδομένων. Η απεικόνιση γνωρισμάτων πρέπει να είναι ακριβής στις σημασιολογικές παραλλαγές και στα λάθη του μετασχηματισμού. Επιπλέον πρέπει να είναι αρκετά χαρακτηριστική ώστε να συλλάβει τις διακριτικές μουσικές πτυχές ενός μουσικού κομματιού.

Το χρώμα αντιστοιχεί στις 12 παραδοσιακές κατηγορίες pitch ( τόνοι έντασης ) της κλίμακας μέτρησης και υποδεικνύεται συνήθως από τις ιδιότητες των συλλαβών του τόνου έντασης , C, C# , B …D όπως χρησιμοποιούνται στην Δύση (δυτική μουσική). Τα γνωρίσματα του χρώματος αποτελούν το γνωστό φαινόμενο της ανθρώπινης αντίληψης, ότι ο τόνος έντασης είναι περιοδικός , υπο την έννοια ότι δύο τόνοι γίνονται αντιληπτοί ως παρόμοιοι στο χρώμα ακόμα και αν διαφέρουν κατά μια οκτάβα. Στην περίπτωση των ακουστικών καταγραφών τα γνωρίσματα του χρώματος υποδεικνύουν τη βραχυπρόθεσμη κατανομή ενέργειας και συσχετίζουν την αρμονική πρόοδο του κομματιού βασισμένα στις τεχνικές επεξεργασίας σήματος . Ο μετασχηματισμός μιας ακουστικής καταγραφής σε απεικόνιση χρώματος μπορεί να γίνει χρησιμοποιώντας τις σύντομες τεχνικές Fourier σε συνδυασμό με τις binning στρατηγικές.

Ο μετασχηματισμός της ανιχνευμένης μουσικής φύλλων σε ένα χρωμάγραμμα ( απεικόνιση χρώματος ) , απαιτεί κάποια στάδια. Κάθε ανιχνευμένη σελίδα αναλύεται χρησιμοποιώντας την οπτική αναγνώριση μουσικής , OMR( Optical Music Recognition). Στο σύστημα αυτό χρησιμοποιείται το λογισμικό SharpEye, ώστε να γίνει η εξαγωγή των μουσικών νότων , παραμέτρων ( χρόνοι αρχής, τόνοι έντασης , διάρκειες )και πληροφοριών από την ανιχνευμένη εικόνα. Έπειτα χρησιμοποιώντας ρητές πληροφορίες του τόνου έντασης και χρόνο – συγχρονισμού, ένα χρωμάγραμμα μπορεί να υπολογιστεί ουσιαστικά αναγνωρίζοντας και προσδιορίζοντας τους τόνους έντασης που ανήκουν στην ίδια κατηγορία χρώματος .

*Σημείωση* Ένα χρωμάγραμμα μουσικής φύλλων είναι πιο καθαρό από ένα ακουστικό χρωμάγραμμα όσον αφορα την απεικόνιση του(γνωρίσματα).

Scanned score

Ακουστικό ευρετήριο και ακουστική αντιστοιχία

Η βασική ιδέα για την αυτόματη ανάλυση εγγραφών είναι ότι συνδέοντας δυο διαφορετικούς τύπους δεδομένων ( οπτικά και μουσικά δεδομένα ) στον ίδιο τύπο απεικόνισης χρώματος , αποτελεί την άμεση σύγκριση στο επίπεδο των γνωρισμάτων απέναντι στις δυο περιοχές . Για μια αποδοτικότερη σύγκριση, επέρχεται μια διεργασία των γνωρισμάτων του χρώματος ποσοτικοποιώντας τα διανύσματα του χρώματος, χρησιμοποιώντας τα codebook ( σημασιολογικά εξηγήσιμα διανύσματα ). Σύμφωνα με κάποια διανύσματα codebook τα γνωρίσματα μπορούν να αποθηκευτούν σε κάποιο αρχείο του ευρετηρίου , το οποίο είναι γνωστό ως δομή ευρετηρίου και χρησιμοποιείται στην ανάκτηση προτύπων – κειμένων .

Στο σύστημα αυτό, υιοθετείται η ακουστική αντιστοιχία ως μια βασική μηχανή για την ανάκτηση μουσικής και προσδιορισμού κομματιών από το μουσικό έργο. Η βασική προσέγγιση αντιστοιχίας λειτουργεί ως εξής . Κάθε μουσικό έγγραφο του αποθετηρίου, μετατρέπεται σε μια ακολουθία 12 διαστάσεων των διανυσμάτων του χρώματος. Στα όρια των εγγραφών όλες αυτές οι ακολουθίες χρώματος συνδέονται με μια ενιαία ακολουθία των γνωρισμάτων του χρώματος . Όμοια όταν δίνεται ένα query, μουσικού clip, μετασχηματίζεται σε μια ακολουθία γνωρισμάτων του χρώματος. Αυτή η ακολουθία του ερωτήματος συγκρίνεται με όλες τις άλλες συνακολουθίες που αποτελείται από τα διανύσματα ακολουθίας της βάσης δεδομένων. Αυτές οι συνακολουθίες θα αποτελέσουν τις επιθυμητές αντιστοιχίες για την ανάκτηση βάση περιεχομένου.

Για την μέτρηση των πιθανών χρονικών διαφορών ανάμεσα στο ερώτημα ( music clip) και των αντίστοιχων χρονικών περιοχών μέσα στα έγγραφα (πχ. Διαφορές ρυθμού μεταξύ διαφορετικών ερμηνειών του ίδιου κομματιού μουσικής ) υιοθετείται η τεχνική του πολλαπλού ερωτήματος με ποικίλα χρωμαγράμματα σε διαφορετικά ποσοστά δειγματοληψίας .

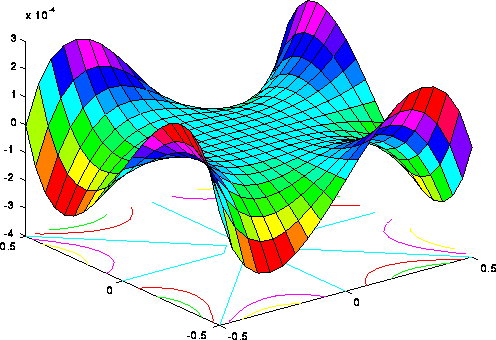
Προσδιορισμός ανίχνευσης και σχολιασμός

Μετά την διαδικασία ψηφιοποίησης , τα ψηφιακά έγγραφα πρέπει να σχολιαστούν (περαιτέρω πληροφορίες) κατάλληλα προτού ενσωματωθούν σε μια ψηφιακή βιβλιοθήκη . Στην περίπτωση των ψηφιακών ακουστικών καταγραφών , ορίζονται μεταδεδομένα όπως : τίτλος , καλλιτέχνης , στίχους , σε κάθε μεμονωμένη καταγραφή. Το έργο του βοηθητικού εγχειριδίου σχολιασμού διατίθεται σε βάσεις δεδομένων που ειδικεύονται σε διαφορετικούς τύπους δεδομένων όπως , Gracenote , DE-PARCON .

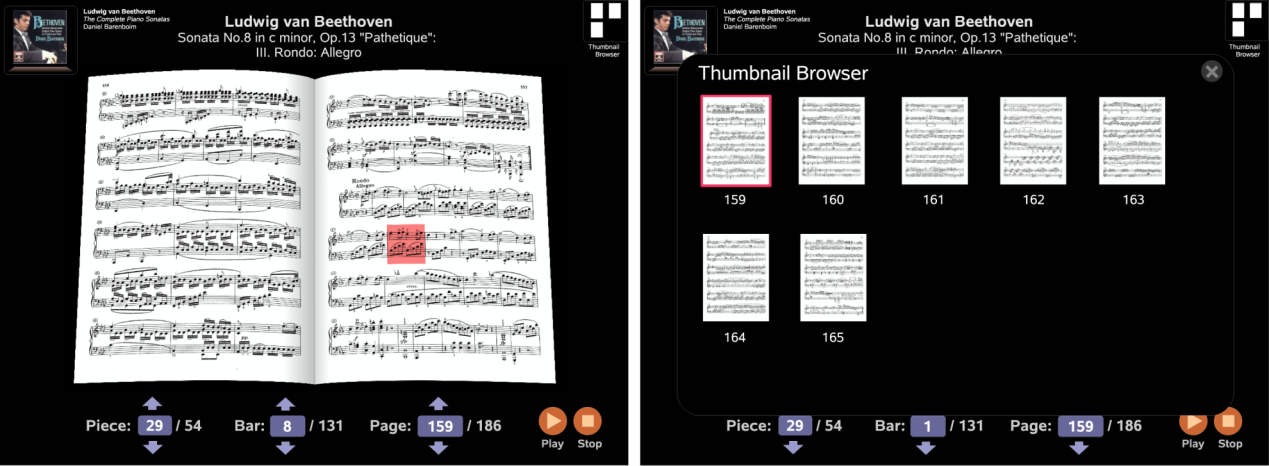
Ο σχολιασμός της ανιχνευμένης μουσικής φύλλων επιτυγχάνεται με μια καινούρια αυτοματοποιημένη διαδικασία. Συγκεκριμένα, στο πιλοτικό αυτό πρόγραμμα, υποθέτουν την ύπαρξη μιας ακουστικής βάσης δεδομένων που περιέχει σχολιασμένες ψηφιακές ακουστικές καταγραφές για όλα τα μουσικά κομμάτια που περιλαμβάνονται στην διαδικασία ψηφιοποίησης μουσικής φύλλων. Έπειτα αυτοματοποιείται ο σχολιασμός των αυτοματοποιημένων σελίδων. Σε ένα στάδιο προεπεξεργασίας μετασχηματίζονται τα ακουστικά έγγραφα σε ακολουθίες από διανύσματα χρώματος και δημιουργείται μια ακουστική δομή ευρετηρίου. Έπειτα, στο στάδιο σχολιασμού κάθε ανιχνευμένη σελίδα μετατρέπεται σε μια ακολουθία διανυσμάτων του χρώματος . Χρησιμοποιώντας την ακολουθία αυτή ως ερώτημα υπολογίζεται η καλύτερη αντιστοιχία μέσα στα ακουστικά έγγραφα . Υποθέτοντας , ότι κάθε σελίδα περιλαμβάνεται σε ένα single μουσικό έργο , η καλύτερη αντιστοιχία αναμένεται να βρεθεί μαζί με μια μουσική αντίστοιχη ακουστική καταγραφή. Η ανιχνευμένη σελίδα μπορεί να προσδιοριστεί από την καλύτερη αντιστοιχία και μετα να σχολιαστεί αυτόματα από τα μεταδεδομένα που συνδέθηκαν ήδη με την αντίστοιχη ακουστική καταγραφή. Αναμένεται ότι η βελτίωση των OMR αποτελεσμάτων και η ταυτόχρονη χρήση των διαφορετικών OMR διαδικασιών θα βοηθήσουν σημαντικά την ποιότητα της αντιστοιχίας .

Ανίχνευση και Ακουστική Ευθυγράμμιση

Μόλις προσδιοριστούν οι ανιχνευμένες σελίδες της μουσικής φύλλων και αντιστοιχιστούν οι ακουστικές καταγραφές συνδέονται αυτόματα τα σημασιολογικά σχετικά γεγονότα με δύο είδη μουσικής απεικόνισης . Σε αυτό το στάδιο υιοθετούνται τεχνικές συγχρονισμού της μουσικής ώστε να συνδεθούν οι περιοχές ( που δίνονται ως συντεταγμένες pixels) μέσα στις ανιχνευμένες εικόνες της δεδομένης μουσικής φύλλων με τις σημασιολογικά αντίστοιχες χρονικές θέσεις μέσα στην ακουστική καταγραφή. Η βασική ιδέα είναι να μετατραπούν οι ανιχνευμένες σελίδες και οι αντίστοιχες ακουστικές καταγραφές μέσα σε ακολουθίες γνωρισμάτων του χρώματος οι οποίες μπορούν μετά να συγχρονιστούν βασιζόμενες σε τυποποιημένες τεχνικές ευθυγράμμισης, όπως η δυναμική χρονική warping. Στην περίπτωση των ανιχνευμένων σελίδων κατασκευάζεται μια ακολουθία γνωρισμάτων για τη μουσική εγγράφου , από κατάλληλη ομαδοποίηση γνωρισμάτων που λαμβάνονται από τις μεμονωμένες σελίδες .

(WARPING)

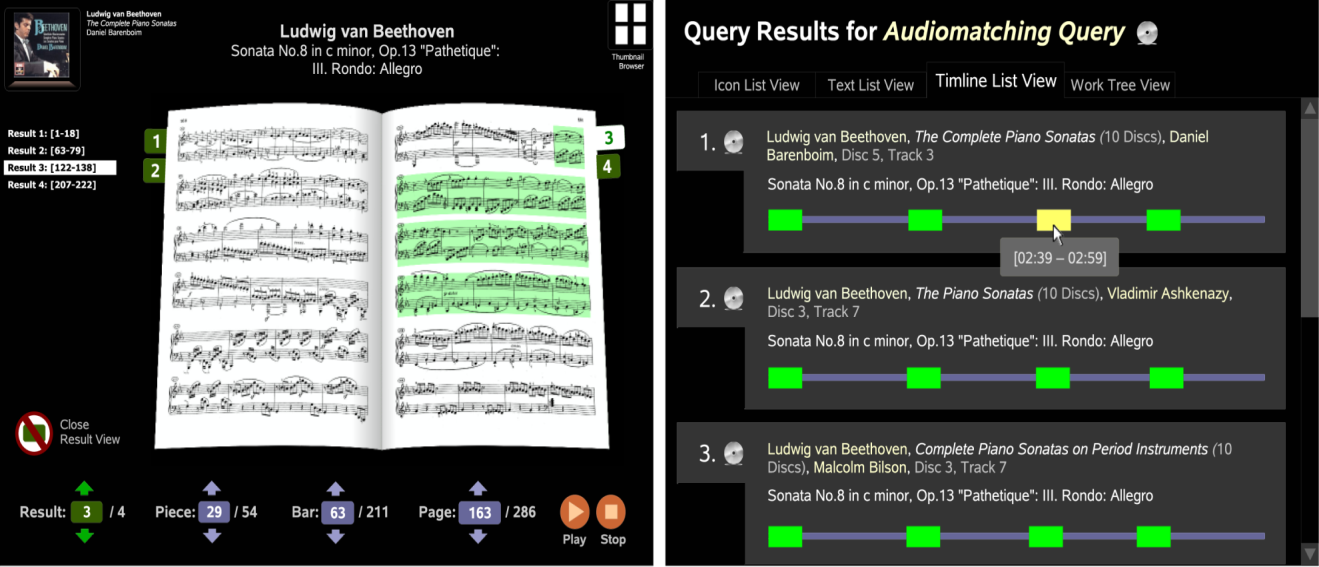
Score and Result – Viewer διεπαφή



Κατά την έναρξη της ακουστικής αναπαραγωγής ήχου, αντίστοιχα μέτρα μέσα στην μουσική φύλλων είναι συγχρόνως τονισμένα , βασιζόμενα στη δομή σύνδεσης που παράγεται από την ανιχνευμένη ακουστική ευθυγράμμιση . Στην παραπάνω εικόνα μια περιοχή στο κέντρο αριστερής σελίδας αντιστοιχεί στο 8ο μέτρο της 3ης movement, και είναι τονισμένη από ένα περιβαλλόμενη από ένα σκιασμένο κουτί. Κατά τη διάρκεια της αναπαραγωγής ήχου οι αριθμημένες σελίδες γυρίζουν αυτόματα. Η διεπαφή Viewer-Score διαχειρίζεται ολόκληρο ανιχνευμένο scorebook (ψηφιακό βιβλίο με παρτιτούρες ), και επιτρέπει την πλοήγηση μέσα σε αυτό, χρησιμοποιώντας τους αριθμούς που βρίσκονται κάτω από τις ανιχνευμένες σελίδες . Με την μηχανή αναζήτησης THUBNAIL φαίνεται ένα τοπικό πλαίσιο των σελίδων γύρω από την τρέχουσα θέση αναπαραγωγής ήχου, που μπορεί να χρησιμοποιηθεί για πλοήγηση. Αν πάνω από μια καταγραφή είναι διαθέσιμη , για το ενεργό κομμάτι μουσικής , ο χρήστης μπορεί να αλλάξει ανάμεσα σε αυτά, (λίστα με εικονίδια) κάνοντας marking πάνω σε αυτά. Μια λειτουργία, για την ανάκτηση μουσικής βασισμένη στο query έχει ενσωματωθεί στη διεπαφή Viewer- Score . Συγκεκριμένα, ο χρήστης μπορεί να επιλέξει συγκεκριμένες περιοχές μέσα στη μουσική φύλλων , χρησιμοποιώντας το δείκτη ποντικιού. Με το σωστό marking στην επιλεγμένη περιοχή μια ερώτηση (query ) μπορεί να εκδοθεί.



Για παράδειγμα, στην παραπάνω εικόνα , τα πρώτα 17 μέτρα του Beethoven (Rondo) επιλέχθηκαν ως ερώτημα . Η δομή σύνδεση που παράγεται από την ακουστική ανιχνευμένη ευθυγράμμιση, ορίζει την επιλεγμένη περιοχή μουσικής φύλλων στο αντίστοιχο χρονικό διάστημα της ακουστικής καταγραφής που προσδιορίζεται στο στάδιο προεπεξεργασίας . Στο παράδειγμα αυτό, τα πρώτα 17 μέτρα του Rondo, αντιστοιχούν στα δευτερόλεπτα 1-20 , μιας ερμηνείας του Barenboim. Μια ακολουθία γνωρισμάτων του χρώματος τότε εξάγεται από την ακουστική καταγραφή. Ακολούθως η ακουστική αντιστοιχία, χρησιμοποιείται ως ερώτημα της ακολουθίας γνωρισμάτων στο ακουστικό ευρετήριο.



Probado project

To Probado project είναι ένα συνεταιριστικό γερμανικό ψηφιακό πρόγραμμα βιβλιοθηκών που χρηματοδοτείται από τον Deutsche Forschungsgemeinschaft. Το πρόγραμμα αυτό άρχισε τον Φεβρουάριο του 2006, και έχει δοκιμαστική διάρκεια έως και 5 χρόνια. Σκοπός του είναι να ενσωματωθούν τα non textual έγγραφα πολυμέσων σε ένα πλαίσιο υπηρεσιών των κατάλληλων βιβλιοθηκών . Οι δευτερεύουσες υποχρεώσεις του προγράμματος , είναι η περαιτέρω ανάπτυξη και εφαρμογή του. Συγκεκριμένα , χρησιμοποιεί μεθόδους υποστήριξης αυτόματης επεξεργασίας των εγγράφων αυτών όπως, ανάκτηση – σχολιασμός – αναζήτηση – αποθήκευση .

The metadata model for the probado music subject area

Το probado-πρότυπο είναι βασισμένο στις λειτουργικές απαιτήσεις για τα βιβλιογραφικά αρχεία (FRBR). **FRBR** είναι μια πρόταση της διεθνούς ομοσπονδίας των ενώσεων και των οργάνων βιβλιοθήκης για μια βελτιωμένη δομή των βιβλιογραφικών μεταδεδομένων. Το έγγραφο FRBR έχει δημοσιευθεί το 1998, και έλαβε σημαντική προσοχή κατά τη διάρκεια των τελευταίων ετών. Η νέα καταχώρηση RDA («περιγραφή και πρόσβαση των πόρων») που αναπτύσσεται αυτήν την περίοδο από τη μικτή οργανωτική επιτροπή για την αναθεώρηση AACR στις αρχές του 2009 θα ενσωματώσει τις έννοιες FRBR.Το πρότυπο FRBR είναι κατάλληλο για τη μουσική επειδή είναι κεντροθετημένο και μπορεί να διακρίνει μεταξύ μιας εργασίας και των εκφράσεών του. Επομένως, διάφορα καινοτόμα προγράμματα στην περιοχή βιβλιοθηκών μουσικής είναι βασισμένα στο πρότυπο FRBR, όπως το πρόγραμμα Variations 2 και Music Australia στην εθνική βιβλιοθήκη της Αυστραλίας .

Το έγγραφο FRBR χρησιμοποιεί ένα πρότυπο οντότητα-σχέσης για την περιγραφή της προτεινόμενης δομής δεδομένων των βιβλιογραφικών αρχείων. Οι οντότητες διαιρούνται σε τρεις ομάδες, όπου η ομάδα 1 περιγράφει τις διανοητικές και βιβλιογραφικές μονάδες, η ομάδα 2 περιγράφει τους πράκτορες που περιλαμβάνονται , και η ομάδα 3 περιγράφει τα διαφορετικά θέματα.

Η μουσική Probado OPAC

Η αναζήτηση στα μεταδεδομένα των αποθηκών μουσικής Probado είναι παρόμοια με την παραδοσιακή αναζήτηση σε OPACs των μεγάλων βιβλιοθηκών μουσικής. Είναι μια βασισμένη στο κείμενο αναζήτηση με μια βασική μορφή αναζήτησης (με μόνο έναν εισαγμένο τομέα) και μια προηγμένη μορφή. Στην περίπτωση προηγμένης μορφής , ο χρήστης μπορεί να διευκρινίσει τις οντότητες και τις ιδιότητες των οποίων οι τιμές πρέπει να ταιριάξουν με το κείμενο που ο χρήστης εισάγει στους τομείς εισαγωγής. Για την παρουσίαση των αποτελεσμάτων αναζήτησης, δεν είναι ικανοποιητικό να εξασφαλίσει στον χρήστη έναν μη δομημένο κατάλογο εγγράφων. Για αυτό τον λόγο η FRBR-δομή των μεταδεδομένων πρέπει να συντηρηθεί. Π.χ., η Probado μουσική OPAC πρέπει να διαφοροποιήσει τον προσδιορισμό ενός μουσικού έργου από το προσδιορισμό μιας ιδιαίτερης καταγραφής.

Βασισμένη στο περιεχόμενο αναζήτηση

Για την βασισμένη στο περιεχόμενο αναζήτηση, το αποθετήριο προσφέρει τρεις μηχανές ερώτησης . Εκείνες οι μηχανές ερώτησης διευκολύνουν την βασισμένη στο κείμενο αναζήτηση στις μετά κειμένου, καταγραφές των φωνητικών διαδρομών του κομματιού (λυρικά ποιήματα), την βασισμένη στη μελωδία αναζήτηση (ερώτηση στις μουσικές νότες από τα αποτελέσματα), και την ακουστική ανάκτηση, βασισμένη στην ακουστική προσέγγιση.

Ανάλυση και ευρετηρίαση εγγράφων βασισμένη στο περιεχόμενο

Λόγω του απρόσιτου περιεχομένου των λυρικών ποιημάτων και του υλικού – score σε συμβολική ψηφιακή μορφή, τα ανιχνευμένα αποτελέσματα υποβάλλονται σε επεξεργασία από εφαρμογές της οπτικής αναγνώρισης χαρακτήρα OCR, και την οπτικής αναγνώρισης μουσικής, OMR. Αν και τα εξαγόμενα λυρικά ποιήματα και το υλικό- Score, παρεκκλίνουν από τα σωστά δεδομένα, λόγω του OCR-OMR , είναι επιρρεπή σε λάθη το υλικό είναι εντούτοις χρήσιμο για τη διαδικασία αναζήτησης και ανάκτησης. Χρησιμοποιώντας τις κατάλληλες εφαρμογές καταχώρησης (ευρετηρίου) , τα λυρικά ποιήματα και η μελωδία(που περιλαμβάνεται στο ευρετήριο) δημιουργούνται από τα OMR – OCR . Οι ακουστικές καταγραφές υποβάλλονται σε επεξεργασία από έναν ακουστικό καταχωρητή του ευρετηρίου, που εξάγει τα ακουστικά χαρακτηριστικά γνωρίσματα βασισμένα στο χρώμα και αντιπροσωπεύουν την αρμονική πρόοδο ήχου, ενός κομματιού. Από εκείνα τα χαρακτηριστικά γνωρίσματα, ένα ευρετήριο χαρακτηριστικών γνωρισμάτων δημιουργείται. Είναι δυνατό να εξαχθεί μια απεικόνιση βασισμένη στο χρώμα από τα δεδομένα του OMR. Προσθέτοντας , τα αποτελέσματα των γνωρισμάτων σε ένα ευρετήριο ακουστικών γνωρισμάτων , η ακουστική αντιστοιχία συγχρονίζεται με τις ακουστικές καταγραφές και το αποτέλεσμα – Score.

Αποθετήριο μουσικής Probado της Εθνικής Βαυαρικής βιβλιοθήκης στο Μόναχο της Γερμανίας

Ένα πρωτότυπο αποθετήριο για τις συλλογές των στοιχείων μουσικής είναι αυτήν την περίοδο είναι η οργανωμένη Εθνική Βαυαρική βιβλιοθήκη BSB στο Μόναχο της Γερμανίας . Το BSB είναι καλός υποψήφιος για ένα αποθετήριο μουσικής επειδή το τμήμα μουσικής του αποτελείται από μια μεγάλη συλλογής εγγράφων, συμπεριλαμβανομένων 347.000 εγγράφων μουσικής φύλλων, 36.000 χειρογράφων μουσικής, 78.000 ακουστικών καταγραφών, και 132.000 βιβλίων (μουσικής και μουσικολογίας ).

ΣΧΕΤΙΚΗ ΕΡΓΑΣΙΑ – RELATED WORK

Το πρόγραμμα Variations 2 στο πανεπιστήμιο της Ιντιάνα (Indiana) στο Μπλούμινγκτον ( Bloomington) παρέχει μια φιλική προς το χρήστη πρόσβαση στη μουσική με μορφή ανιχνευμένων μουσικής και ήχου φύλλων. Το πρότυπο μεταδεδομένων του είναι επίσης βασισμένο στο πρότυπο FRBR. Οι προοριζόμενοι χρήστες είναι σπουδαστές στη μουσική και στη μουσικολογία. Δύο ερευνητικά προγράμματα στον τομέα των ψηφιακών βιβλιοθηκών και των αρχείων μουσικής έχουν αρχίσει πρόσφατα και χρηματοδοτούνται από την Ευρωπαϊκή Ένωση: Το πρόγραμμα EASAIER που συντονίζεται από το Πανεπιστήμιο του Λονδίνου Queen Mary, θα δημιουργήσει τις νέες και καινοτόμες μεθόδους πρόσβασης στα ηχητικά αρχεία. Στο πρόγραμμα DISMARC μια κοινοπραξία 10 συνεργατών προγράμματος, θα καταστήσει τις ακουστικές συλλογές προσβάσιμες , όπου ο ιδιοκτήτης του εκάστοτε περιεχομένου παρέχει την πρόσβαση σε αυτό (listenable archives).

Συμπεράσματα και μελλοντικά σχέδια εφαρμογής

Σε αυτό το έγγραφο παρουσιάστηκε ένα πλαίσιο ψηφιακής βιβλιοθήκης για την διαχείριση συλλογών μουσικής ανιχνευμένων φύλλων και σχετιζόμενων ακουστικών καταγραφών . Ξεκινώντας με ένα έργο για την επεξεργασία εγγράφων, οι τεχνικές συνεισφορές αφορούν τις μεθόδους για αυτόματο προσδιορισμό σελίδων της ανιχνευμένης μουσικής φύλλων και την ακόλουθη ευθυγράμμιση (αντιστοιχία ) της μουσικής φύλλων με τις ακουστικές καταγραφές. Για την διευκόλυνση πλοήγησης στη πολύμορφη μουσική, παρουσιάστηκε η διεπαφή Viewer –Score, για τον χρόνο – συγχρονισμό και την αναπαραγωγή ήχου της ανιχνευμένης μουσικής φύλλων και αντίστοιχων ακουστικών καταγραφών . Επιπλέον, παρουσιάστηκε ένας νέος μηχανισμός, για την ανάκτηση μουσικής βασισμένης στο περιεχόμενο, επιλέγοντας άμεσα, περιοχές ως ερωτήματα από την μουσική φύλλων. Τα αποτελέσματα των ερωτημάτων εμφανίστηκαν στη διεπαφή Viewer – Result , επιτρέποντας την audio αναζήτηση και την πολύμορφη πλοήγηση στη διεπαφή Viewer – Score. H τρέχουσα συλλογή δοκιμών του πιλοτικού προγράμματος , αποτελείται από 32 piano sonatas ( 101 αρχεία audio, movements μεμονωμένες ) του Ludwig Van Beethoven . Κάθε ένα από αυτά τα κομμάτια είναι διαθέσιμη μια ανιχνευμένη έκδοση των αντίστοιχων φύλλων μουσικής και λαμβάνεται από μια έκδοση του G. Henle Verlag , ανερχόμενη στο συνολικό αριθμό των 604 ανιχνευμένων σελίδων. Για κάθε movement το λιγότερο μια ακουστική καταγραφή είναι διαθέσιμη. Το ποσοστό προσδιορισμού για τις μεμονωμένες ανιχνευμένες σελίδες που χρησιμοποιούν την αντιστοιχία μουσικής βασισμένη στο χρώμα είναι 82,5 %. Αναμένεται να αυξηθεί αυτό το ποσοστό χρησιμοποιώντας δεδομένα της OMR( Optical Music Recognition) , στην διαδικασία αντιστοιχίας .

Το προτεινόμενο έργο και οι διεπαφές του χρήστη είναι μέρος της οργάνωσης αποθετηρίων της μουσικής Probado , που ξεκίνησε στην Εθνική Βαυαρική βιβλιοθήκη , στο Μόναχο της Γερμανίας . Για το συγκεκριμένο αποθετήριο μουσικής περιλαμβάνεται μια μεγάλη συλλογή κλασικής και ρομαντικής μουσικής piano sonatas , των HAYDN, MOZARD, BEETHOVEN, SCHUBERT, SCHUMANN, CHOPIN, LISZT, BRAHMS. Παρόλο που το προτεινόμενο τεχνικό έργο για την αυτόματη επεξεργασία εγγράφων είναι πλήρως λειτουργικό υπάρχουν περιορισμοί με το βασικό υλικό μουσικής. Το πιο σημαντικό είναι ότι η αντιστοιχία ανιχνευμένης μουσικής φύλλων και ακουστικών καταγραφών ως ιδανική , υπήρξε ως ΄΄ υπόθεση ΄΄ και όχι ως πραγματική. Μέθοδοι για την διαχείριση τύπων διαφορετικών δομών που περιλαμβάνουν τον μερικό συγχρονισμό της μουσικής, δεν έχουν αξιολογηθεί ακόμα και αποτελούν αντικείμενο έρευνας αυτή την περίοδο, 2009.

Βιβλιογραφία

1. Wang, P., Bunke, H.: Handbook on Optical Character Recognition and Document

Image Analysis. World Scientific, Singapore (1997)

2. Witten, I.H., Moffat, A., Bell, T.C.: Managing Gigabytes, 2nd edn. Van Nostrand

Reinhold (1999)

3. Baeza-Yates, R.A., Ribeiro-Neto, B.A.: Modern Information Retrieval. ACM Press,

Addison-Wesley (1999)

A Framework for Managing Multimodal Digitized Music Collections 345

4. Ohta, M., Takasu, A., Adachi, J.: Retrieval methods for English-text with missrecognized

OCR characters. In: ICDAR 1997: Proceedings of the 4th International

Conference on Document Analysis and Recognition, Washington, DC, USA, pp.

950–956. IEEE Computer Society, Los Alamitos (1997)

5. Harding, S.M., Croft, W.B., Weir, C.: Probabilistic Retrieval of OCR Degraded

Text Using N-Grams. In: Peters, C., Thanos, C. (eds.) ECDL 1997. LNCS,

vol. 1324, pp. 345–359. Springer, Heidelberg (1997)

6. Google Inc.: Google Book Search (2007), http://books.google.com/

7. Kurth, F., M¨uller, M., Fremerey, C., Chang, Y., Clausen, M.: Automated Synchronization

of Scanned Sheet Music with Audio Recordings. In: Proc. ISMIR, Vienna,

Austria, pp. 261–266 (September 2007)

8. Krottmaier, H., Kurth, F., Steenweg, T., Appelrath, H.J., Fellner, D.: PROBADO

- A Generic Repository Integration Framework. In: Proceedings of the 11th European

Conference on Digital Libraries (September 2007)

9. Bartsch, M.A., Wakefield, G.H.: Audio thumbnailing of popular music using

chroma-based representations. IEEE Trans. on Multimedia 7(1), 96–104 (2005)

10. Hu, N., Dannenberg, R., Tzanetakis, G.: Polyphonic audio matching and alignment

for music retrieval. In: Proc. IEEE WASPAA, New Paltz, NY (October 2003)

11. M¨uller, M.: Information Retrieval for Music and Motion. Springer, Heidelberg

(2007)

12. Choudhury, G., DiLauro, T., Droettboom, M., Fujinaga, I., Harrington, B.,

MacMillan, K.: Optical music recognition system within a large-scale digitization

project. In: Proc. ISMIR, Plymouth, MA, USA (2000)

13. Byrd, D., Schindele, M.: Prospects for improving OMR with multiple recognizers.

In: Proc. ISMIR, Victoria, Canada, pp. 41–46 (2006)

14. Jones, G.: SharpEye Music Reader (2008), http://www.visiv.co.uk/

15. Kurth, F., M¨uller, M.: Efficient Index-based Audio Matching. IEEE Transactions

on Audio, Speech, and Language Processing 16(2), 382–395 (2008)

16. Gracenote: WWW (2008), http://www.gracenote.com/

17. Krajewski, E.: DE-PARCON Softwaretechnologie (2008),

http://www.de-parcon.de/

18. Arifi, V., Clausen, M., Kurth, F., M¨uller, M.: Synchronization of music data in

score-, MIDI- and PCM-format. Computing in Musicology 13 (2004)

19. Dunn, J.W., Byrd, D., Notess, M., Riley, J., Scherle, R.: Variations2: Retrieving

and using music in an academic setting. Special Issue, Commun. ACM 49(8), 53–58

(2006)

20. Diet, J., Kurth, F.: The Probado Music Repository at the Bavarian State Library.

In: Proc. ISMIR, Vienna, Austria, pp. 501– 504 (September 2007)

21. IFLA Study Group on the Functional Requirements of Bibliographic Records:

Functional Requirements for Bibliographic Records; Final Report. Saur, Munich

(1998), <http://www.ifla.org/VII/s13/frbr/frbr.pdf>