A location-tracking interface for ethnomusicological collections

Michela Magas
Goldsmiths Digital Studios
University of London
London SE14 6NW, UK
Email: map02mm@gold.ac.uk

Polina Proutskova
Department of Computing, Goldsmiths
University of London
London SE14 6NW, UK
Email: mas02pp@gold.ac.uk

Annotation in ethnomusicological archives is often incomplete, and searches involving artist's or composer's name often do not apply. We propose a novel system for intuitive exploration of ethnomusicological recordings which combines a geographical visual interface with ethnomusicological annotation, similarity matching and location tracking.

I. INTRODUCTION

Music is the best way to approach a culture and its people: appreciating ethnomusicological recordings may help establish the understanding for combating the ills of ethnocentrism, racism and religious fundamentalism [1],[2]. In this paper we present our research into tools for more efficient exploration of ethnomusicological archives. In the opening section we provide an overview of the current state of ethnomusicological archives; we follow by describing existing systems which can be used as interfaces for ethnomusicological recordings; and we present a novel user interface combining annotation with geographical sampling, audio similarity searches, and geographical tracking of search results. We conclude with an outlook and further research suggestions.

II. THE CURRENT STATE OF ETHNOMUSICOLOGICAL ARCHIVES

Ethnomusicological recordings document the diversity of musical repertories and traditions in a systematic way and constitute an important part of cultural heritage. Many endangered or extinct music cultures can only be preserved for humanity in the form of recordings safeguarded in ethnomusicological archives. Dissemination of ethnomusicological recordings made possible a shift during the Twentieth Century from a global understanding of music based on ethnocentrism and a belief that music evolved torwards European harmony, to one based on knowledge of the fundamentals, culture and underlying philosophies of the societies in which the music was created [2]. Anthony Seeger [1],[3] describes the significance of the archives for the field of ethnomusicology and the public groups interested in using the archives' resources.

Europe has a long standing tradition of documenting musical cultures by means of audio (nowadays also audio-visual) recordings. European institutions hold historical recordings of key importance not only for European cultures but also musical traditions outside Europe. The first ethnomusicological audio recording was made in 1900 by Erich M. von Hornbostel, when Siamese (Thai) musicians were giving a performance in Berlin. The Berlin Phonogram Archive was founded in the same year in order to collect, preserve and study music recordings of world cultures. We can still listen to this first recording today, as well as to many other wax cylinder recordings preserved at the Berlin Phonogram Archive.

There are huge repositories of ethnomusicological recordings residing at various institutions in Europe. The National Sound Archive (Traditional Music Section) in London hosts one of the largest ethnomusicological collections in the world, containing recordings from all around the globe. The Berlin Phonogram Archive and the Phonogram Archive in Saint Petersburg (Russia) are also very large, culturally spread collections. Many countries fund national sound/music archives, which are the main documenting bodies of local and national music cultures of the state (Polish National Sound Archive, Bulgarian Academy of Sciences Music Archive, etc.). Fig. 1 shows the amount of ethnomusicological recordings in some of the largest archives of the world.

Though the policies of European institutions usually allow researchers and general public to access recordings in ethnomusicological archives, their holdings are still very much underused and are not sufficiently exposed to potential users. In many cases it is still necessary to come physically to the archive in order to search its collections and to listen to the recordings [4].

A considerable number of recordings in European archives has now been digitized according to world standards (IASA TC-04). DISMARC (http://www.dismarc.org/) was the first European project to expose catalogues of ethnomusicological archives online. It provides a metasearch of ethnomusicological collections to museums, universities, national archives and broadcasting companies from Germany, Sweden, Finland and the UK. The number of participants is steadily growing. We envision further growth of digital content in ethnomusicological archives as well as large-scale collaborative projects like DISMARC and EUROPEANA (European Digital Library, http://www.europeana.eu/) to make huge distributed reposito-

	Items	Hours (incl. commercial)	Original collections
Berlin Phonogram Archive	Over 150,000	18,000	Over 1,000
US Library of Congress, Archive of Folk Culture	-	Over 100,000	Over 4,000
National Sound Archive of the British Library, World and Traditional Music	300,000	-	370
Archives for Traditional Music, Indiana University	128,550	Over 250,000	2,000

Fig. 1. Amount of recordings in selected ethnomusicological archives

ries accessible online.

Ethnomusicological archives in Europe, USA and other parts of the world are now reinventing themselves, and there are many obstacles to overcome, including missing infrastructure for large digital libraries, metadata management, effective search strategies and attractive presentation to the user. We address some of these issues in our research described in this paper and suggest scalable solutions.

III. INTERFACES FOR ETHNOMUSICOLOGICAL RECORDINGS

The growth of ethnomusicological resources searchable online requires new scalable retrieval strategies as well as their intuitive graphical representations [5]. Ethnomusicological collections constitute a challenge with respect to both metadata and content management. The audio content of these collections is still a less studied terrain in Music Information Retrieval [6],[7]. Today most of the online ethnomusicological archives' catalogues offer bibliographic text searches that may sometimes include categories specific to music (eg. The British Library Sound Archive Catalogue simple search, http://cadensa.bl.uk/cgi-bin/webcat). Finding salient results is hindered by the incompatibility of controlled vocabularies between archives and metadata schemes.

In ethnomusicological archives search and retrieval differ significantly from music archives containing Western repertories: there is usually no composer because many recordings are of traditional music; the artist is in most cases not known to the user; there is no such thing as album title and the song title may be just the lyrics incipit; genre and functional descriptions may exist, but they are most definitely not compatible between cultures. The first descriptor according to which an ethnomusicological collection is usually searched is the cultural origin of the music [8], which often (but not always) coincides with the geographic location where the recording was made.

Alan Lomax was the first to suggest a user interface for an ethnomusicological collection based on geographic location, using recordings from the Cantometrics database [9]. The Cantometrics project, started in the 1960s and pursued by Lomax until the 1990s, aimed at discovering correlations between singing styles and societal traits such as hierarchy structures and gender relations. For that purpose Lomax and his colleagues gathered and analysed around 7,000 recordings from more than 500 cultures [9],[10]. In his Global Jukebox [11] clickable icons on the world map trigger the audio content of the recording made at the respective location. A data similarity search is based on the 37 Cantometrics parameters, and relies on the manual annotation of each track by expert listeners. The results of this search can be shown on the geographical map or used to map out charts showing correlations between patterns of song and patterns of society [11].

Portals such as the National Geographic Nat Geo Music portal (http://worldmusic.nationalgeographic.com/) or Mondomix (http://www.mondomix.com/) provide a fairly comprehensive overview of music from all over the globe, including audio and video searches, though they are limited to commercial releases. Their interfaces rely on a simple combination of text and media, offering the user a list of countries to choose from, with limited data visualization, interactivity or content management, and few incentives for intuitive exploration.

Visualisations of commercial recordings rely mostly on clustering by genre or mapping of artist relationships (eg. [12],[13]), which make them inappropriate for ethnomusicological recordings. With classification by genre, attempts to classify music often fail because of reasons like ambiguities, subjective judgment and marketing interests [14]. Musicovery (http://www.musicovery.com/, originally developed as Music Plasma or Live Plasma by Frederic Vavrille) provides a colourful visualization of tracks according to genre and links between similar artists. It provides the option of selecting and listening to the genre of 'world' music, though the interface relies on standard music interface data (artist and name of track) leaving no room for ethnomusicological descriptions. It provides no information about the geographical or cultural origin of the tracks, or why particular tracks have been visually linked, and presents no opportunity for data management and annotation.

Collaborative database projects, such as The Freesound Project (http://www.freesound.org/) which focuses on environmental sound [15], and Xeno-Canto (http://www.xeno-canto.org/) which focuses on bird song [16], rely on voluntary contributions of recordings, and use the Google Maps geotagging system as the basis for exploration of recordings from around the world. Clicking on a pin provides a link to a list of recordings related to the chosen location, which in turn provide links to audio samples. Recordings can be downloaded under the Creative Commons license and freely studied and explored. Tags and descriptions can be collaboratively added. Even

	continent	location / culture	singing	solo / group	male / female	instruments
2	Europe	Georgia				
3	N. Europe	England				
4	N. America	Virginia	singing	solo	mixed	guitar
5	N. America	Pueblo Taos	singing	group	male	drum
6	S. America	Interior Amazon, Jivaro				
7	N. Europe	Hebrides, Scotland	singing	solo group	solo female group mixed	no
8	C. Europe	Asturias	singing	group	group female group male	no
9	Afro-America	Andros, Bahamas island	singing	solo group	male	no
10	E. Europe	Daghestan, Caucasus	singing	group	group female group mixed	no
11	C. Africa	Mbuti Pygmies, Equatorial rainforest	singing	group	mixed	no
12	C. Asia	Kazakhstan	singing	solo	male	lute
13	S. Europe	Trapani, Sicily				
14	E. Europe	Molokan, Central Russia				
15	E. Asia	Ainu				
16	N. America	French Canadian				
17	E. Europe	State Siberian Russian Folk Choir	singing	group	group female group male	no
18	W. Africa	Wolof, Senegal				
19	E. Europe	Rhodope, Bulgaria				

Fig. 2. Manual annotation of the Cantometrics Training Tapes dataset.

though they are not designed specifically for ethnomusicological recordings, these collaborative audio-collecting systems offer greater scope for ethnomusicological research.

Geotagging is also used by the British Library Sound Archive online interactive maps which allow researchers to sample, download and study a variety of recordings with related texts and metadata (http://sounds.bl.uk/Maps.aspx). However, none of the above examples make use of the feature extraction and audio similarity tools made available by the latest developments in MIR research.

IV. ENHANCED GEOGRAPHICAL USER INTERFACE

We present a dynamic user interface which offers a novel audio content-based search for ethnomusicological recordings. It makes use of the latest findings on similarity searches in the field of MIR and adds a visual feature for location-tracking the similarity results, thus offering a carefully designed, intuitive gateway for exploring and searching a music collection of any cultural and geographical scope.

The starting point for the design of this interface is a set of ethnomusicological recordings of world music compiled by Alan Lomax as part of the Cantometrics project. We use a collection of 355 recordings from the Cantometrics Training Tapes dataset [9] containing all varieties of singing and playing styles from more than 50 cultural origins. The nature of the collection suggests a search by geo-location. We propose a multi-layered interface which allows for query triggers from a geographical plane located at a level below the audio similarity search interface. The layering of geo-positioning and similarity matching enables the discovery of interesting anthropological and ethnomusicological relationships.

The Cantometrics recordings can be sampled at their geo-locations before a query is triggered. We manually annotated the tracks with 9 attributes: a capella singing, accompanied singing, purely instrumental, choral singing, solo singing, male singing, female singing and mixed male and female singing (Fig. 2). These are represented by simple, colour-coded symbols at their relative location: male or female singer, solo or group, vocal or instrumental. They provide a novel view of the geographic distribution of singing style patterns. In order to preserve the legibility of the symbols and the layers above them, the design of the World map is a minimal, flattened, monochrome version of the Google Maps graphic (Fig. 3).

Global positioning of the individual symbols is done manually, in absence of any geographical coordinates in Lomax's original annotation. This occasionally presents a problem: some locations belong to a different country than they did at the time when the Cantometrics recordings were collected and some of the countries and people's names have since changed. We have tried to keep faithful to the original descriptions as much as possible, though cannot guarantee absolute accuracy.

On query launch, the recorded sample is matched to other Lomax recordings from around the globe and brought to view in the mHashup similarity interface [17]. Our interface can potentially use any audio similarity search engine that is able of finding best matching fragments of audio files in a database. In this example we use the audioDB database management system which employs Euclidean distance in the normed audio shingle space to retrieve similar tracks [18],[19],[20]. For each audio file we extract 20-band Mel Frequency Cepstral Coefficients (MFCC). These are extracted using a short-time Fourier transform with hop size 100ms (2205 samples), window length 185.76ms (8192 samples), FFT length 16384 samples (2.69Hz frequency bins). Feature extraction is performed using fftExtract.

The similarity interface and the geographical interface create layers which can be independently reduced in size or enlarged as required. Layering of query results over their geographical locators allows for an immediate overview of a variety of elements and parameters related to the recordings. The length of a bar corresponds to the relative length of a recording. The position of the matching audio shingle [17] is aligned with that of the queried one, visually highlighted, and allows for further sampling by direct manipulation. All searches are stored, enabling a return to a previously chosen query pathway and a jumping-off point for new searches.
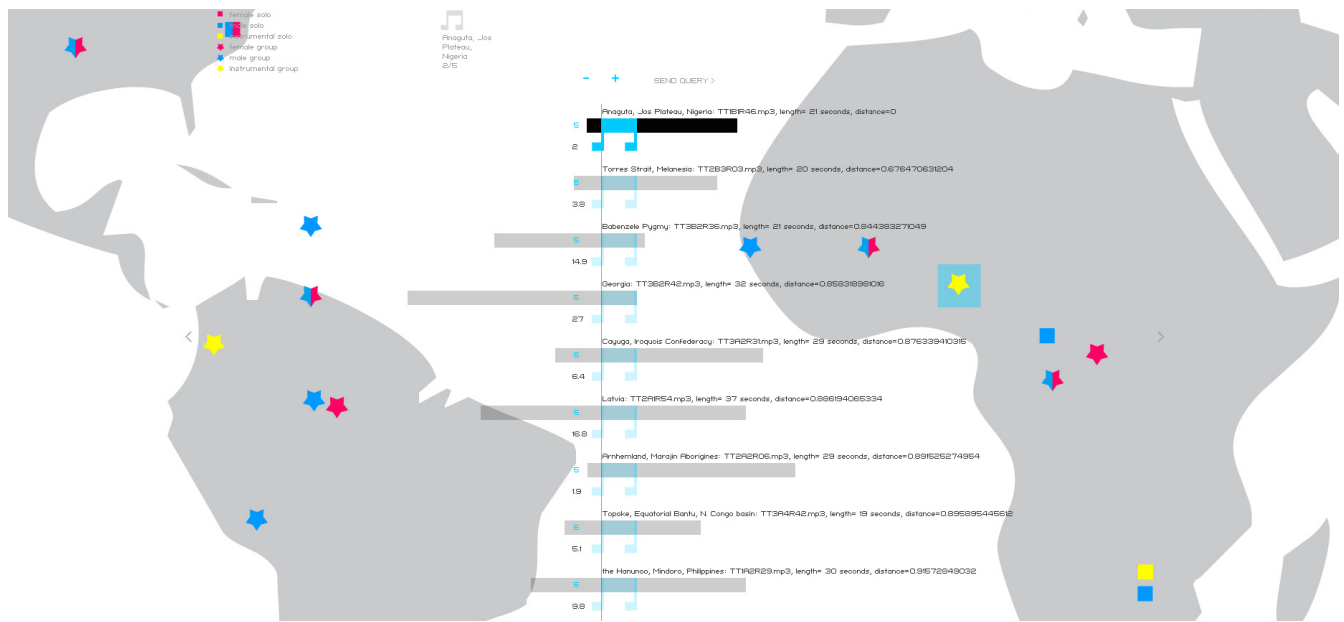
Fig. 3. A location-tracking interface for ethnomusicological recordings. The square location tracker is visible on the right in Nigeria, as the site of the Anaguta, which have been chosen as the initial query for the audio similarity results displayed in the centre. The highlighted query results are displayed on bars corresponding to the relative lengths of tracks, as segments relative to the query length, and are aligned to the query point. Clicking on a query result will play the corresponding audio and launch the location tracker on a trajectory to the corresponding location on the map. Artwork copyright 2009 Michela Magas.

The matched recordings can be sampled individually, with a highlighting device tracking the query trajectory from one part of the globe to another. The location tracker aligns itself with the geographical location of each sampled query result before launching on a trajectory to the next one, creating a web of pathways indicating relationships between cultures in distant locations. Thus looking for recordings similar to highly polyphonic choral singing from Georgia in the Caucasus with passages of yodel in the highest part, the location tracker will travel around the world to Mbuti Pygmies in the Central African rainforest singing complicated interlocked patterns also containing yodel elements, to a choir from Montenegro, to Native Americans from the Iroquois Confederacy, back to Georgia and then to Nigerian pan pipe playing. The location tracker is here intended as a visual aid to the user's understanding of the relationships, connections and distances between ethnomusicological recordings and related cultures.

V. PRELIMINARY USER EVALUATION

Preliminary user testing was conducted with four male participants of varying ages: 60, 37, 30 and 27. They were encouraged to explore the interface intuitively, without much briefing or explanation. They were then asked four questions: How easy is it to sample a recording? How easy is it to trigger the audio similarity query? How easy is it to sample and compare the results? How easy is it to zoom in and out of layers? The first three questions received consistently high

scores (Fig. 4), while the zoom function proved harder to use.

On a touch screen, such as on the Apple iTouch, the zoom function is integral to the device's core set of functionalities and regular users are familiar with the related commands. In such an environment a zoom function is expected by the users and therefore does not require special instructions. The task of implementing a zoom function is more complex in the desktop environment, particularly where it relies on the use of a mouse. Here we rely on a trigger by the zoom + and - symbols for its functioning. These need to be visually flagged very clearly as they are not a common feature of desktop interfaces, and users do not necessarily expect to find them or understand that, in the case of our interface, each layer can be scaled separately. Zooming in and out of layers, while not essential to the operation of our interface, ought to be flagged to the user more prominently as it may be a useful addition, providing more freedom for the exploration of the interface.

Useful suggestions have been provided regarding the sending of queries, such as making the 'Send Query' button relate visually to the latest audio shingle sampled, in order to reinforce the connection and indicate that the said shingle can now be used as query for audio similarity matching. We have responded to this suggestion by making the 'Send Query' button mobile: it now aligns itself with the latest audio shingle sampled, thus indicating the said shingle can be used as a query trigger for further searches.

A participant noted that the audio similarity matching occasionally matches silence to audio query content. This is a

Participant's age:	60	37	30	27	Average score
How easy is it to sample a recording?	8	9	10	9	**9**
How easy is it to trigger the audio similarity query?	9	8	8	10	**9**
How easy is it to sample and compare the results?	9	8	6	9	**8**
How easy is it to zoom in and out of layers?	6	3	3	8	**5**

Fig. 4. Preliminary user evaluation. Results are 1 - 10, where 10 is best.

bug of the current feature extraction process and it is to be further explored. Sampling longer audio shingles of more than 15 seconds may cause a delay in the streaming of audio on certain networks. On these rare occasions, a visual cue ought to be provided to indicate to the user that the audio is about to be streamed. An option to play results simultaneously has been requested and can be provided. Similarly, an option can be provided allowing the users to choose whether they prefer to initialise the audio similarity search according to a match by melody or timbre. The usefulness of the above suggestions will be tested with further user evaluation.

VI. FUTURE WORK

The interface would benefit from the planned addition of our system for automatic annotation of digital audio recordings [21]. This system automatically extracts basic musical attributes, independently of their cultural origin or musical style, and helps in the addition of consistent metadata to large ethnomusicological repositories, thus enabling applications to large-scale collaborative projects.

Every time a new song is added to the database, feature extraction and automatic annotation would be initiated and reflected in the visual interface by the corresponding symbols and metadata. Automated annotation rarely gives absolutely reliable answers, thus it is no substitute for the work of an archivist or a researcher and must be clearly distinguished from reliable human annotation. Where the archivists and librarians find the metadata inconsistent or incomplete, they ought to be able to edit it by direct manipulation of the interface.

Further musical parameters could be extracted automatically: amount of percussivity, pulse stability, amount of repetition, presence of specific singing production patterns such as yodel, drone or overtone singing. The variety of musical style patterns can then be represented graphically showing their spread around the world.

The enhanced geographic interface could be further automated by including a semi-automatic geographical location component. Geographical coordinates based on the given metadata could allow for the automatic positioning of a recording on the map. Only when the cultural/geographic description cannot be automatically located or is ambiguous, would the manager/archivist be prompted to search for coordinates manually.

The function of the location tracker suggests the idea of 'culture-hopping', which can be further explored and enhanced with the addition of relevant anthropological information. Every time the location tracker lands in a particular location, information about the related culture could be provided: emphasizing, for example, whether this is an historic or current culture, whether it is tribal, linked to agriculture or urban.

Online integration may be very beneficial at this point. People interested in particular recordings might be better informed about their cultural context than the archivists are. Social tagging could be used to counter the errors generated in automatic annotation: an online editorial system would allow registered users to add, edit and negotiate information on catalogue entries, which could then become available to all users. Combined with archivists' expertise and moderation, this approach would enable archives to close gaps in annotation and offer hands-on activities to their user communities.

VII. CONCLUSION

In expectation of further growth of digital content and online exposure of ethnomusicological collections we have presented a dynamic graphic user interface combining graphic representation of ethnomusicological annotation with an audio content search, and providing a location-tracking visualization of ethnomusicological collections of large scope. Benefits of combining metadata with audio content search and retrieval have often been stressed in MIR literature [22],[23]. Ethnomusicological archives may benefit more than any other content holders from this approach: on the one hand cultural context is especially important in ethnomusicology and audio similarity can only become meaningful in conjunction with cultural information; on the other hand audio similarity search can help to close gaps in missing, poor or inconsistent metadata.

According to preliminary evaluation, the interface fulfilled all of the essential usability requirements and received positive comments. The users were able to explore the collection intuitively, without prior training. Minor amendments and additions suggested will be introduced in future work to enhance the interface's usability further.

Exposing an ethnomusicological collection to a similarity search highlights the huge variance in the music data with respect to cultural origin, musical style and structure, singing production, instruments played as well as recording quality.

This opens possibilities for further research into the extent the similarity search strategy used (both low level audio features and the pattern matching mechanism) corresponds to listeners' perception of similarity, and particularly how this perception varies culturally.

Our combined location-tracking and audio content similarity search can facilitate cross-cultural research in ethnomusicology, as in Victor Grauer's hypothesis about the evolutionary relationship between the pygmy/bushmen singing style and various interlocked pan pipe playing styles [24]. It offers the user the opportunity for creative exploration and unexpected discoveries: a novel way to discover music cultures of the world.

REFERENCES

[1] A. Seeger, "The role of sound archives in ethnomusicology today," *Ethnomusicology*, vol. Spring/ Summer, 1986.

[2] R. Reigle, "Humanistic motivations in ethnomusicological recordings," in *Recorded Music - Philosophical and Critical Reflections* (M. Dogantan-Dack, ed.), Middlesex University Press, 2009. with CD.

[3] A. Seeger, "Ethnomusicologists, archives, professional organisations, and the shifting ethics of intellectual property," *Yearbook for Traditional Music*, pp. 87–107, 1996.

[4] P. Proutskova, "Data infrastructure for ethnomusicological archives – current situation and future perspectives", *Journal of International Society of Sound and Audiovisual Archives*, pp. 45–54, july 2008.

[5] P. Proutskova and M. Magas, "Beyond the metadata, new intelligent audio content search for large music collections", presented at the *Unlocking Audio 2* conference, British Library, London, UK, 16-17 Mar., 2009.

[6] G. Tzanetakis, A. Kapur, W. A. Schloss, and M. Wright, "Computational ethnomusicology", *Journal of Interdisciplinary Music Studies*, vol. 1, no. 2, pp. 1–24, 2007.

[7] J. S. Downie, "Music information retrieval," *Annual Review of Information Science and Technology*, vol. 37, pp. 295–340, 2003.

[8] M. Mengel, "ethnoArc, linked european archives for ethnomusicological research," *Tech. Rep. D4*, Berlin Phonogram Archive, February 2007.

[9] A. Lomax, "Cantometrics: An Approach To The Anthropology Of Music", The University of California, 1976. accompanied by 7 cassettes.

[10] A. Lomax, "Folk Song Style and Culture", New Brunswick, NJ: Transaction Books, 1968.

[11] A. Lomax et al, "Global Jukebox", available at http://www.culturalequity.org/video/ce_videos_global_jukebox.jsp.

[12] D. Gleich, M. Rasmussen, K. Lang and L. Zhukov, "The World of Music: SDP layout of high dimensional data", in *Info Vis*, 2005.

[13] O. Goussevskaia, M. Kuhn, M. Lorenzi and R. Wattenhofer, "FromWeb to Map: Exploring the World of Music", *IEEE/WIC/ACM Int. Conf. on Web Intelligence*, 2008.

[14] O. Hilliges, P. Holzer, R. Klüber and A.Buts, "AudioRadar: A metaphorical visualization for the navigation of large music collections", in *Lecture Notes in Computer Science*, 2006.

[15] X. Serra, "Technologies to support the collaborative production of sounds: the example of Freesound.org", presented at the *Unlocking Audio 2* conference, British Library, London, UK, 16-17 Mar., 2009.

[16] W.-P. Wellinga and B. Planqué, "Xeno-canto: web-based sound collecting and identification", presented at the *Unlocking Audio 2* conference, British Library, London, UK, 16-17 Mar., 2009.

[17] M. Magas, M. Casey and C. Rhodes, "mHashup: Fast Visual Music Discovery via Locality Sensitive Hashing", *ACM SIGGRAPH*, 2008.

[18] M. Casey and M. Slaney, "Song Intersection by Approximate Nearest Neighbour Retrieval", *Proc. International Conference on Music Information Retrieval (ISMIR)*. Victoria (BC), 2006.

[19] C. Rhodes and M. Casey, "Algorithms for Determining and Labelling Approximate Hierarchical Self-Similarity", in *Proceedings of the International Conference on Music Information Retrieval*, Vienna, Austria, 2007.

[20] M. Casey, C. Rhodes, and M. Slaney, "Analysis of Minimum Distances in High Dimensional Musical Spaces", IEEE Transactions on Audio, Speech and Language Processing, 2008.

[21] P. Proutskova and M. Casey, "You call that singing? toward MIR tools for multi-cultural collections of ethnomusicological recordings," *Proceedings of the International Symposium on Music Information Retrieval*, 2009 (Under submission).

[22] J.-J. Aucouturier, F. Pachet, P. Roy, and A. Beurive, "Signal + context = better classification," *Proceedings of the International Symposium on Music Information Retrieval*, 2007.

[23] P. Proutskova, "Musical memory of the world - data infrastructure in ethnomusicological archives," *Proceedings of the International Symposium on Music Information Retrieval*, 2007.

[24] V. Grauer, "Echoes of our forgotten ancestors", in *The World Of Music*, vol. 2, 2006.