# mHashup: Fast Visual Music Discovery via Locality Sensitive Hashing

Michela Magas
Stromatolite Design Lab
Goldsmiths Digital Studios

Michael Casey[†]
Goldsmiths Digital Studios

Christophe Rhodes[‡]
Goldsmiths Digital Studios

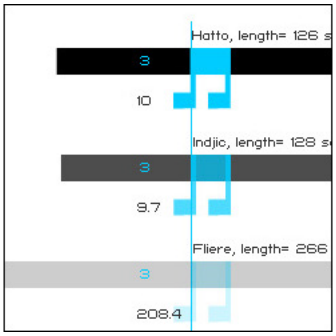

**Figure 1:** *mHashup*

## Introduction

Millions of tracks in music download services pose a problem: how to find "dark media" – those items for which little is known but that users may want to find. The problem is similar for music lawyers and musicologists: how to identify closely matching music fragments in large recorded music collections.

mHashup is a novel visual interface to large music collections, such as today's million-song download services, for discovering musical relationships among tracks. Users engage in direct on-screen query and retrieval of music fragments in an instantaneous feedback flow performed by a locality-sensitive hash table in secondary storage.

## mHashup

The mHashup interface was constructed by analysing user interactions with complex audio search processes and selectively identifying the salient elements of the query formation, search and retrieval processes. A system of visual symbols, representing the audio query and results, with corresponding auditory displays was designed to facilitate enhanced interaction with the search engine. The visual interaction design of mHashup makes content-based search intuitive.

The mHashup visual language consists of track bars with relative lengths corresponding to the temporal extent of each track. Within each mHash bar, the start point and end point of a selected music fragment, called an audio shingle, are highlighted and afford direct manipulation to change the location and extent of the selection. The shingles are short-windowed segments of a track, usually extending over 1s-10s of audio.

The user triggers a search via a search button; the results consist of a display of similar tracks to the user's query with matching shingle locations within the tracks highlighted as new query handles. The track displays are aligned according to the best matching time-point within each track. This affords ease of visual inspection of the time positions of the retrieved tracks with respect to the query track. The user selects a shingle handle within any one of the result tracks to hear the match and by activating the search button the user iterates the search process; this can be performed in a feedback cycle affording rapid traversals of musical geodesics in the audio similarity space.

Matching specific audio content requires temporal features. We use a combination of cepstral coefficients and chroma features extracted from the audio in short time intervals of 100ms. An audio shingle consists of a concatenation of audio features into high-dimensional feature vectors. The system uses Euclidean distance in the normed audio shingle space to retrieve similar tracks; a locality sensitive hashing algorithm (LSH) performs the retrieval with sublinear time complexity in the total number of shingles; 100ms per retrieved track for a database of 1.4 million shingles compared with 2 hours per track using linear distance computation. The degree of closeness of similar tracks is visually represented by the use of greater or lesser transparency in the track display. A memory-store is generated for each query enabling a return to a previously chosen query pathway and a jumping-off point for new searches.

## Conclusion

mHashup facilitates both professional music use cases, such as musicologists and copyright lawyers seeking the origins of sampled music with location markers precisely given for each returned track, and end-user music applications, such as discovery of "dark media" by its relationship to known "hot" items. Applications include locating sampled audio; un-mixing remixes; finding cover songs; and spotting musical influences. mHashup's visual interface reflects the core functionality of a content-based search engine as a visual grammar to be explored by direct manipulation.

## References

CASEY, M., RHODES, C. and SLANEY, M. Analysis of Minimum Distances in High Dimensional Musical Spaces, *IEEE Transactions on Speech, Audio and Language Processing*, accepted for publication.

SLANEY, M. and CASEY, M. 2008. Locality Sensitive Hashing for Music Audio, *IEEE Signal Processing Magazine*.

---

email: michela@stromatolite.com
[†] email: m.casey@gold.ac.uk
[‡] email: c.rhodes@gold.ac.uk