Chapter III
Melodic Query Input for Music Information Retrieval Systems

Richard L. Kline
Pace University, USA

ABSTRACT

This chapter discusses key issues of building and using a system designed to search and query a music collection through the input of the actual or perceived melody of a song. It presents a detailed survey and discussion of studies, algorithms and systems created to approach this problem, including new contributions by the author. Emphasis is placed on examining the abilities and likely errors of those with little or no formal musical training to remember and reproduce melodic phrases, as these must be taken into account for any music information retrieval system intended for use by the general public. The chapter includes an extensive discussion on human humming as an intuitive input method for the musically untrained person and an examination of several music information retrieval systems designed for this type of input query.

INTRODUCTION

An effective music information retrieval (MIR) system requires suitable mechanisms for specifying input queries. For queries of the nature of identifying a tune, theme or melody, providing query input in the form of a melody line or rhythm usually will be the most intuitive and powerful means of specifying such a search. Similarly, collections to be searched in this manner must have some available representation of the melody and rhythm contained therein, such as that presented in written form or in formats such as MIDI (1993), rather than raw or processed digitized audio samples such as CD or MP3. In this chapter we will present many of the most important theories, algorithms, techniques and performed studies for capturing, representing and performing similarity or identification searches using melodic input query strings.

MIR systems requiring the highest search efficiencies generally must demand that the collection
being searched and any given input queries are known to be error-free. There must be confidence in the quality of the songs comprising the search collection and error-free input can be reasonably expected only for systems built for and used by those with substantial musical training, researchers or publishers.

Systems designed for collections or input queries which are known or suspected to contain errors must be built with additional complexity, or a resulting loss of search efficiency, to take real or expected errors into account. These errors can come from many sources. For the collection being searched, errors may be introduced when digitizing manuscripts or when automated transcriptions are performed. For input queries, systems designed for those lacking in music training and those allowing for vocalized melody input, often called query-by-humming systems, are nearly certain to contain input errors. Some of these are due to the way nonmusicians perceive and recall music; others are due to performance errors; and still others due to imperfect transcription of acoustically-recorded vocal or instrumental input queries.

In this chapter we bring together the most relevant studies from the music literature in order to paint a clear picture of the allowances which must be made for input processing and searching of melodic queries. We then describe early and recent research efforts in creating error-tolerant melody-based MIR systems and highlight the best contributions of each. We conclude the chapter with a summary of the best practices presented and extrapolate likely avenues of further exploration in search of more efficient and successful MIR systems.

The discussion presented here refers specifically to music with an identifiable melody line, that is, a sequence of notes readily recognized as the dominant tune within a song that would be naturally reproduced by a listener able to play, sing or hum only a single note at a time. With vocal music, the melody is almost always associated with the notes sung by the primary vocalist. Even complex instrumental orchestral pieces typically have a theme or melody which is often played by a solo instrument and readily identified by casual listeners.

Regardless of whether a specific music collection being searched has been stored in a representational notation or digitized recordings with possible feature extraction, the ability to form musical queries based on a melody, via notation or by vocal or instrumental performance, will remain a powerful and needed option for many such MIR systems.

**OBTAINING AND PROCESSING MUSIC**

Many data formats have been created to store music and these fall into two basic categories. The average consumer is most familiar with music stored as a digital or analog audio recording, such as is found in cassettes, CDs, MP3 files, DRM-encrusted file formats and myriad other physical and digital formats. The recorded sounds can be stored in a single data stream or several; each channel can represent the output intended for a specific speaker upon playback (e.g., stereo or surround sound), or each channel can be devoted to a single instrument or vocalist, as is done for most professional audio recordings. This latter arrangement allows the greatest flexibility in processing the recording, affording the opportunity to select the relative emphasis given to various tracks through mixing before creating the published version.

The second method of storing music is a strictly logical or diagrammatic one. Rather than saving digitized waveforms of the actual acoustic sounds produced, information about the individual notes to be played or sung is recorded instead. The fundamental characteristics of a note include the time it should begin, its duration, its pitch value and the instrument or voice which is to intone the note.
Related Content

MARSYAS-0.2: A Case Study in Implementing Music Information Retrieval Systems
www.igi-global.com/chapter/marsyas-case-study-implementing-music/24427?camid=4v1a

Lyric Recognition and Christian Music
www.igi-global.com/chapter/lyric-recognition-christian-music/31059?camid=4v1a

Alternative Design Goals for a General Music Markup Language
www.igi-global.com/chapter/alternative-design-goals-general-music/24557?camid=4v1a

DART: A Framework for Distributed Audio Analysis and Music Information Retrieval
www.igi-global.com/chapter/dart-framework-distributed-audio-analysis/24435?camid=4v1a