How Query by humming, a Music Information Retrieval System, is Being Used in the Music Education Classroom

Brian Bradshaw*

Abstract.

This study does a qualitative and quantitative analysis of how music by humming is being used by music educators in the classroom. Music by humming is part division of music information retrieval. In order to define what a music information retrieval system is first I need to define what it is. Berger and Lafferty (1999) define information retrieval as “someone doing a query to a retrieval system, a user begins with an information need. This need is an ideal document–perfect fit for the user, but almost certainly not present in the retrieval system’s collection of documents. From this ideal document, the user selects a group of identifying terms. In the context of traditional IR, one could view this group of terms as akin to expanded query.” Music Information Retrieval has its background in information systems, data mining, intelligent systems, library science, music history and music theory. Three rounds of surveys using question pro where completed. The study found that there were variances in knowledge, training and level of awareness of query by humming, music information retrieval systems. Those variance relationships where based on music specialty, level that they teach, and age of the respondents.


I. PURPOSE

The purpose for studying music information retrieval systems, and specifically query by humming is to determine how the technology is being used, if at all in the music education classroom. Coming from both an Information systems and music background, the use of Music Information Retrieval Systems (MIR) holds a special interest for me. There has been a growing number of music archives that has been digitalized. This digitalization is “not only music sheets, but even videos and audio recordings have been translated into their digital counterpart. The preservation of cultural heritage and the possibility of employing the powerful instruments of information technology are the main motivations for this transformation. However, the process of digitization has brought to the fore the problem of organizing and accessing huge amounts of data” (Haus et al, 2004).

Music Information Retrieval (MIR) is a musical database that when used with in conjunction with a “software tool, allows for the contents of the music, for example the pitch or melody, to be analyzed or queried. This type of technology has huge commercial potential, and tools like Shazam, which search for a recording based on a recorded section provided by the user, are extremely popular” (Mangan, 2012) There are two main search types for MIR systems, query by manuscript (musical score) and query by humming. This type of database has an even larger potential as an educational learning tool. There have been several studies completed looking at MIR systems from a university library’s stand point, but none have been completed looking at the use of MIR systems in the music education classroom.

Right now research is being conducted on a large scale “within the Library and Information Science and Computer Science communities to create tools like Shazam, which can search the contents of a recording, instead of the traditional metadata associated with a recording, such as title or composer name”. (Dannenberg, et al., 2007) This type of Music Information Retrieval is known as query by humming. Query by humming is defined as “music information retrieval systems where short audio clips of singing or humming act as queries” (Quin et el 2011). This
is different compared to the traditional search of a data base by textually entering information into a database query. Klone 2008.

Music Information Retrieval Systems (MIR) provides an extremely broad resource for musicians on musical pieces. No longer are educators and students restricted to being able to listen to a piece of music only if the piece is available in the Universities music library. Now they can search YouTube, Spotify and other MIR systems by artist, performer or name of the musical piece. So instead for example of being limited to just one recording of Dmitri Shostakovich’s Symphony No. 5 in D minor by the Cleveland Orchestra, they can find the same symphony performed by other orchestras, bands or by whatever musical group has performed it across the world. You can say for example, using YouTube and other retrieval systems, one could locate Symphony No. 5 in D minor, with certain conductors such as Leonard Bernstein, or principle players such as Yo Yo Ma on cello.

The issue that this study is looking at is how Music Information Retrieval Systems technology is being used in the music education class room. More specifically, what is being looked at is retrieval by humming. Are music educators in the collegiate setting using this technology, and if so how? Output once the music is retrieved can come in many forms such as MP3 or score retrieval. The form of output is not a part of this study, just as long as there is output. How college level educators use that output in and out of the classroom is the subject to be examined. Most studies and articles that have been published in the field of music technologies in the classroom are currently coming from Australia and Europe. These will be noted in chapter 2. There are no current studies looking at how MIR systems are being used in the University level classroom for music education.

The study will use qualitative and quantitative methodology, specifically, a designed survey. The survey will be distributed to professional music educators who are part of the Pennsylvania Music Educators Association (PMEA), Ohio Music Educators Association (OMEA), New York State Music Educators (NYSMTA) and Texas Music Educators Association (TMEA). These organizations are child groups of the parent organization National Association of Music Educators (NAfME).

II. Background

Simon (2000) defines Music Information Retrieval (MIR) is a “blanket term for a diverse range of sub-topics including: computational methods of classification, clustering, and modeling, formal methods and databases, web software, human-computer interaction and interfaces, music perception, cognition and emotion, music analysis and knowledge representation, music archives and libraries, intellectual property rights, sociology and economy of music, and philosophical and ethical issues related to music”.

Music Information Retrieval (MIR) has its background in both data mining and as an intelligent system. As an intelligent system MIR uses data mining techniques to conduct queries and searches. So in order to fully explain what a MIR system is, a basic explanation of what data mining is must first be provided. “Data mining is defined as the automatic extraction of useful, often previously unknown information from large databases or data sets. One of the popular definitions for data mining is the process of exploration and analysis, by automatic or semi-automatic means, of large quantities of data to discover meaningful patterns and rules. This can be simplified as a process of getting information or knowledge from large amounts of data. That is why sometimes data mining is also known as knowledge mining” (Lone, 2014).

From an intelligent system standpoint MIR system is best explained by John Sowa (2002), the researcher gives this easy to understand definition of what an inelegant system does: “People communicate with each other in sentences that incorporate two kinds of information: propositions about some subject, and meta level speech acts that specify how the propositional information is used as an assertion, a command, a question, or a promise. By means of speech acts, a group of people who have different areas of expertise can cooperate and dynamically reconfigure their social interactions to perform tasks and solve problems that would be difficult or impossible for any single individual. The result is a system with a dynamically changing architecture that can be reconfigured in various ways”.

Mangan (2012) describes Music Information Retrieval as being “split into two separate areas. Symbolic representation allows for musical recordings to be searched for using metadata, such as Artist, Composer or Title. The other area is Content based Music Information Retrieval, where the melodies in the music itself can be searched”. Traditional music retrieval approaches, based on keyword and textual metadata, face serious challenges. If the user is familiar with the name of the song or other significant information to describe the music, retrieval is straightforward. However, if one does not know the title, singer, alternative retrieval methods are necessary. Content-based music information retrieval (MIR) is gaining widespread attention and can be very helpful, since it forsakes the need of keyword. Often, it consists of a form of query-by-example, such as through singing, humming or playing a sample of the piece, as a query to the databases.

“In most cases the query to the system consists of a melody or a melodic contour. These queries can either be entered manually or transcribed from a monophonic audio recording of the user humming or singing the desired
The second approach is called Query-by-humming (George Tzanetakis et al.).

Beginning in the mid-1990s, there is a body of field research based in the area of music information retrieval. In 1995, Asif Ghias first proposed the method based on the Query by Humming, Ghias’s method was based on up and down of the melody. From then on, most approaches use pitch sequence to represent the query. For retrieval, dynamic programming, dynamic time warping (DTW) and hidden Markov models were used. Recently locality sensitive hashing (LSH) approaches are used in this field and gets an impressive result. Since 2005, a number of QBH systems have been evaluated in Music Information Retrieval Evaluation eXchange (MIREX). However, none of the research presents evidence of user perspectives on these tools. So my research is going to be focused on the professional musical educator and their perspective on query by humming.

The processing of how a query is conducted is best described by Jang and Lee (2008) they state that any MIR system can be broken into “3 steps of processing: preprocessing, on-line processing, and postprocessing”. Music query by humming is then retrieved by “the songs in the database are in MIDI (Music Instrument Digital Interface) format, which contains all the music elements of a song and is equivalent to the sheet music of the song. The pitch/beat information of a song is extracted from the MIDI’s major track (or vocal track), which is defined as the track that, when played alone, can be identified immediately by a human who is familiar with the song’s melody”. This is important because they identify the key term of MIDI which stands for Music Instrument Digital Interface. MIDI has become the standard in music digitalization regardless of format.

One of the reasons MRI systems are becoming more of a reality is because “the World Wide Web becomes the source and distribution channels of diverse digital music, a large amount of music tracks is accessible to people. Since it is usually difficult and time consuming for a user to find and choose his/her desired music, the music recommender system becomes an indispensable tool. Music recommendation is valuable in many real world applications, such as social music communities, online music stores, and some music devices (e.g. PCs and MP3 players) where music recommendation can be used to generate music playlists”. (Bu et al, 2014) Music educators can create their own play list based on their lecture series. The play list can be distributed via social media or educational programs such as Blackboard.

As a related development, with the continued advances in mobile technologies making leaps and bounds on a daily base, the reality of a mobile MIR system is becoming more prevalent. Rho, Hwang, Park (2010) “wireless local area network (WLAN) and third generation (3G) technologies provide high-speed mobile access to rich multimedia content on the Internet. Therefore, there has been rising demand to access abundant multimedia content using mobile devices” The ability of MIR systems to be on a mobile device will greatly enhance access to MIR systems being successful. If a user is able to pull out their mobile devise in order to identify or retrieve a song they do not know the name of which is a significant benefit.

There is an established need for MIR systems to be used in the music education classroom along with other technologies. Social media can be used by both educators and students to share works and expand the student’s repertoire. Through the use of social tagging, cell phones and multimedia recordings, co-operative learning can take place between students and educators. The use of MIR systems for the purpose of sharing with others is one of the largest uses of the systems. The sheer vastness of digital music systems is probably one of the best examples of “Big Data”. The ability of members of social groups to be able to share digital music that they like or think other members may like is what has given the rise to MIR systems popularity.

Shazam is a MIR system that has a vast educational potential. “A short explanation of how Shazam works. The company has a library of more than 8 million songs, and it has devised a technique to break down each track into a simple numeric signature—a code that is unique to each track. The main thing here is creating a ‘fingerprint’ of each performance, says Andrew Fisher, Shazam’s CEO. When you hold your phone up to a song you’d like to ID, Shazam turns your clip into a signature using the same method. Then it’s just a matter of pattern-matching Shazam searches its library for the code it created from your clip; when it finds that bit, it knows it’s found your song”. (http://www.slate.com/articles/technology/technology/2009/9/10/that_tune_named.html) Practice first is product by MusicFirst that enables music students to practice any instrument or voice with a web based interface. The program is polyphonic capable which means the software is capable of picking up all of the notes a piano or guitar plays. It also responds to tone and pitch corrections. Students can be recorded online and teachers can review and give feedback.

Music educators traditionally been slow to adopt modern technologies. As the review of literature demonstrated, MIR is a technology that has multiple benefits to both the educator and the student. As a music history tool and score retrieval software it could be invaluable. For example, a student can be at a concert, simply by using programs like Shazam on their phone, they can find the name of the piece,
how wrote it and even the arranger.

III. Method

Research Question 1: What are the issues facing music educators in relation to music information retrieval systems?

Research Question 2: What is the Level of awareness of music educators of content based systems like Query by Humming?

Research Question 3: What are music educators’ perspectives on the use of content based systems like Query by Humming?

The research for this project falls within that definition of social research. It will be conducted through a survey of music educators in Pennsylvania and Ohio. According to Creswell (2013) epistemology is a research approach which is focused on ways of defining knowledge, and the boundaries and limits of that knowledge. From an epistemological viewpoint, the philosophy adopted for this study was that of interpretivism, as it is chiefly concerned with the thoughts and opinions of the research subjects.

In this study, the researcher conducted a pre-test of a small group to ensure that the survey is asking the questions that need to be answered. Once I received the results back from the pre-test, adjustments were made to the survey. Two rounds of surveys where completed. The first round of the survey the majority of questions will be quantitative, followed by open ended qualitative questions. Quantitative research is defined as a “collection of numerical data, as exhibiting a view of the relationship between theory and research”. (Bryman, 2012) Qualitative research is defined as “non-numeric forms of data collection and analysis represented by words. It investigates the why and how of decision making while at the same tries to understand the human behavior and reasons attached to it. It is used by many researchers to observe and create meaning to the social aspect of life”. (Eisner, 1991) The use of both methods for the purpose of this study is expected to yield the best results.

The second round of the survey was an open ended question. The questions where based on the results of the findings of the first survey. The second survey is qualitative in nature and this further re-enforces the argument for interpretivism. Qualitative style questions are better suited than closed end quantitative questions, because they can better examine the issues and challenges in using a music information retrieval system. The purpose and motivations behind the use of IMR systems or lack and understanding of their use will also be looked at by a qualitative methods approach.

The survey itself is broken down into several different question areas. The areas are designed to obtain the background of those taking the survey, and then to address the research questions. The table below links all research questions to the survey questions. Question Pro was used to deliver the survey electronically via email.

Since there has been no prior academic study of how music information retrieval is being used in the music classroom, hope is that the findings of this study can be used to make suggestions for positive use. Positive use by educators is emphasized, because there is the potential for a negative use in the classroom by students. For example, if a music by humming query was conducted by a student, through their cell phone, the possibility for the student cutting required identification exercises or even cheating on an exam is raised. Chances are that even many experienced educators are not aware of MIR systems in order to use the programs in their classrooms.

IV. Population and Sample

The goal of the study was to have 100% of the participants. Given the size of the target population this will be a good representation of the overall population. The respondents will be full time and part time music educators at 4 year colleges and universities. Since two year institutions do not typically offer music degrees, these type of institutions will not be used for this study. The participants of the survey will be members of Pennsylvania Music Educators Association (PMEA), Ohio Music Educators Association (OMEA), New York Music Educators Association and Texas Music Educators Association. According to PMEA they have close to 4,500 members made up of educators, music publishers, manufactures and retailers of music instruments. OMEA according to their web site (www.omea-ohio.org), the Ohio Music Education Association is the third largest of the 52 federated state affiliates of NAfME, The National Association for Music Education. OMEA members are involved at all levels of music education, serving more than 1.8 million public school students in 612 public school districts and 51 colleges and universities in Ohio. The PMEA and the OMEA are state level affiliates of the National Association for Music Education (NAfME). The survey will be distributed by NAfME via email to PMEA and OMEA members. The average educational level of those participating in the study will be master’s degree educated instructors.

This study was conducted in the fall of 2015 for survey distribution and collection. During the fall of 2015 and spring of 2016 for the interpretation and reporting of the data. The survey was distributed via email from the NAfME
national office. The email went out to PMEA and OMEA members in Pennsylvania and Ohio. Through its membership NAfME was able to target PMEA and OMEA members that work in higher education. The members received an email explaining the survey. Contained within the email a link to the survey in Question Pro.

V. Data Analysis

A descriptive analysis was completed on the quantitative survey questions. Descriptive analysis are used to summarize numbers in ways that are useful and meaningful. A general linear model was used to analyze the data and show the relationships between the data that was collected. Mathematical statistical analysis software was used to analyze the raw data and to discover a correlation between the numerical findings.

A qualitative analysis was completed on the survey’s open ended responses. The results were analyzed for themes and patterns. The data went through a reduction process in order to identify and focus on what is meaningful. All the data were carefully reviewed for a content and thematic patterns. After identifying themes and/or content patterns the information will be assembled, organized, and compressed into a display that facilitates conclusions. Then finally conclusions were drawn from the data and reported.

VI. Results

Question pro was used to digitally distribute the survey. There were 226 views of the survey, of which 85 started the survey, and there were 72 completions. The completion rate was 84% with 13 dropouts. The average time to complete the survey was 5minutes. The type of devices that were used to complete the survey are: 84% pc’s (21% used windows 8, 42% used MAC, and 37% used other windows pc’s). 9% used smart phones with 25% android and 75% IPhones, then 7% used tablets with 83% using IPads and 17% using android based systems.

A broad range of ages where represented by the respondents. The age group with the highest percentage of respondents (44.29%) was those holding doctorate degrees. The other groups where (40.00%) master’s degree, and (15.71%) bachelor’s degree. The majority of the respondents teach on the bachelor’s degree level (71.83%). This is followed by master’s level (14.08%), associate’s degree (11.27%) and doctorate degree programs (2.82%). The largest group of the respondents have their educational background in music education (40.28%), followed by music performance (37.5%), Music Theory (8.33%), music History (5.56%) and music technology (1.39%). This seems to be a normal distribution of age for the educational environment.

The survey revealed that in most cases, even when given an example, the respondents did not know what a music information retrieval system is. This suggested a lack of training, and education for music educators on music information retrieval systems. Another finding was that while several of the respondents indicated that it seemed like a nice technology, they would not know how to implement it in the music classroom. For example, one respondent stated: “Contact professors who teach aural skills” which is part of music theory and, “the ability to effectively run query by humming could be useful as an aural skills tool”. Another stated “I think a person from the industry should come in and do a demonstration”. To add to this the respondents did not know what the technology in music education (TI:ME) is. This is the group that is supposed to pioneer and champion the use of technology for music education.

When an explanation was given to the respondents, the majority noted that query by humming could be very useful in the classroom. For example, the technology could be used in tests, to give results faster, to help teach music history and theory. Several respondents stated that they use music information retrieval systems like YouTube to help teach music performance, or to introduce students to examples of great performers and classical performances. One of the biggest opportunities that the respondents saw for MIR systems is to promote self-learning. They pointed out that students can use query by humming based software programs to self-discover more about musical pieces that they hear in public. They can then use social media to share those findings with other students. Such collaboration can further enhance the self-learning process.

A significant number of respondents did not see how the technology could be used to impact music education. Many chose not to answer open ended questions related to its potential impact on music education. 40.91% of the respondents had not heard of music information retrieval systems before participating in the survey. 13.64% of the respondents thought that MIR systems are only good for
entertainment, and 10.61% do not consider them very accurate for searching. That only left 34.85% that thought that Mir systems are useful in the music education classroom. There is a direct relationship with the lack of vision in the use of MRI systems, and the use of CD’s, tapes and LP albums in the classroom, with 61.43% of the respondents stating they still use these media. Only 22.73% of the respondents say that they use MRI systems such as YouTube in conjunction with traditional recorded material.

<table>
<thead>
<tr>
<th>X Variable</th>
<th>Y Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>If so do you use these tools</td>
</tr>
<tr>
<td>Age</td>
<td>What is your opinion of these tools</td>
</tr>
<tr>
<td>Age</td>
<td>Do you find these tools useful</td>
</tr>
<tr>
<td>Age</td>
<td>Do your students use Qbh/Qbp systems</td>
</tr>
<tr>
<td>Age</td>
<td>Do you use recorded material in conjunction with MIR systems</td>
</tr>
<tr>
<td>Education level that you teach</td>
<td>What is your opinion of these tools</td>
</tr>
<tr>
<td>Education level that you teach</td>
<td>Do you use recorded material in conjunction with music information retrieval systems</td>
</tr>
<tr>
<td>Primary focus of music education</td>
<td>What is your opinion of these tools</td>
</tr>
<tr>
<td>Primary focus of music degree</td>
<td>Do your students use Qbh/Qbp systems</td>
</tr>
<tr>
<td>Primary focus of music degree</td>
<td>Do you find these tools useful</td>
</tr>
<tr>
<td>Education level</td>
<td>What is your opinion of these tools</td>
</tr>
</tbody>
</table>

Fig.1. The list of variables.

There was a positive correlation between the variables “what is your primary focus of your music degree” and “if so do you use these tools”. This suggests that there is a correlation between what type of music degree concentration the respondents have and the use of music information retrieval systems. There was a slightly positive correlation between variables x and y in the linear regressions. Figure 1 lists the variables that had a slightly positive correlation. There was one outlier in each category in the linear regressions that kept it from being a 100% positive correlation.

A small amount of respondents identified another use for query by humming music information retrieval systems, as an input for music notation software. Query by humming can be used to record a live player or players. Using this recording the player can see exactly what they have played as musical notation in the form of digital notation, in programs like Finale or Sibelius. This can be used to teach technique, by showing the student what they are doing wrong or doing well. In this way, it could also be used in distance education.

A general lack of knowing what a music information retrieval system is and how a query by humming is completed led to a limitation of the study. YouTube is a music information retrieval system that is widely used. There is a real lack of understanding that YouTube is a data base that is meant as a music information retrieval system, or that a query is a search into a database. That search can be completed by entry into a keyboard or depending on the software a query by humming may be achieved. It does not appear that respondents understood that query by humming is in addition to humming, speech and music through a microphone instead of a keyboard.

VII. Recommendations for Future Research

This study identified several areas of future research. One of the biggest is to identify on a national level just what is being taught in music technology on the higher education level. The organization of Technology In Music Education (TI:ME) needs to take the lead in identifying music technology to be taught, and used as a teaching tool in the music classroom. Repeating this study with a student population would yet be another area of additional research.

If the study where to be replicated, it could be done by evaluating music information retrieval systems generally, and then focusing specifically on query by humming, or specifically looking at how YouTube is in the classroom. Many educators believe that YouTube is the most significant technology ever to happen in music education. The impact on the technology in classroom education cannot be overlooked. YouTube gives students the ability to look up songs by different artist, across musical genera. They can compare and contrast how different artist interprets the same song. The students and instructors are not limited to selections that are contained in the schools musical library. Instead they are exposed to what people have uploaded from around the world.

A future study could be done on collaborative learning and the use of YouTube and several sites throughout the world. The Instructors and students could play, record, share and receive feedback from different universities. This could become especially valued in a world music class. The
shard experiences of the instructors and students could be studied. Ethnographies and case studies could be completed of the students and instructors involved in the project.

REFERENCES


[29] Wiering, “Cognition-based Segmentation for Music...
How Query by Humming, a Music Information Retrieval System, is Being Used in the Music Education Classroom


Authors

Brian Bradshaw: Brian has his BS in music from Clarion University, BA in telecommunications from Westminster College, MS in Communications/CIS from Clarion University, MS in Higher Education Administration from Robert Morris University, and his Doctorate in Information Systems and Communications from Robert Morris University. He has taught K-12 instrumental music and computer science. Brian is currently a professor of computer science at Thiel College, Greenville Pa. Brian is retired from the US Army in which he had served as a specialist in psychological warfare.