Sloud Query-by-Humming Search Music Engine

Valeriy Lobaryev
Sloud Inc.
1900 Campus Commons Drive
Suite 100, Reston, VA 20191 USA
vel@sloud.com

Gene Sokolov
Sloud Inc.
1900 Campus Commons Drive
Suite 100, Reston, VA 20191 USA
gene@sloud.com

Alexandr Gordeyev
National Aerospace University
Chkalov str. 17, Kharkiv,
61070, Ukraine
alex.gordeyev@gmail.com

Abstract

The paper describes Sloud QBH Search Music Engine – a new experimental online query-by-humming search system, based on music information retrieval into notes sequence and two-step search in the indexed melodies database. User's interface of the system is implemented by applying ActiveX technologies. Recognition of the pitch contour is performed in real time by using a color indication on the screen. This allows to improve the quality of humming recognition due to arising visual feedback. Main features of the system are as follows: usage of melodic contour as a hash index; application both music intervals and music duration for approximate matching search, as well as usage of partial matching methods. In the article we have given the data of experimental comparison of its productivity with other similar programs and search systems.

Keywords: QBH, Query By Humming, approximate matching, online systems.

1. Introduction

Query By Humming (QBH) systems are to date one of the components of multimedia search systems of the future, developed within the framework of the MPEG-7 standard [1]. They can be oriented both to professionals in the field of music and to general public. As for the latter, requirements to QBH systems are getting more complicated, since they should allow for humming errors connected with the lack of musical experience [2]. The described system Sloud QBH [3] tries to make accessible online systems, most understandable for the common people.

2. Description of Sloud QBH

2.1 Purpose and Functions

Sloud QBH relates to musical content-based retrieval systems. It is meant for music search by the common people (mainly songs search) by melody hummed into microphone. The result of the search returns to the user in a form of Web-page with a list of found musical compositions, sorted out by relevance.

2.2 The System Architecture

The search system Sloud QBH consists of two parts – client and server, which includes melodies database.

Client part of the engine is a Web-interface, which recognizes and indexes the notes of a hummed melody. Part of the interface, implemented in ActiveX technology, performs the following functions:

a) Provides humming-to-MIDI conversion, i.e. recognizes pitch and duration of the hummed notes;
b) Provides indication of the recognized pitch contour while signing in real time;
c) Allows to play back the recognized melody, converted into MIDI, before sending the query;
d) Indexes the melody and sends a query to Internet-server Sloud QBH;
e) Provides settings (options) of the search and result representation.

Server part of the search engine provides search of musical compositions, which correspond to the query in database and returns the search result to Internet-browser. Block diagram of the search engine is given in Figure 1.

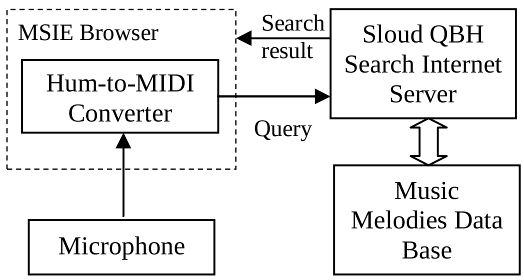

Figure 1. Sloud QBH Search Engine block diagram

The database contains index of melodies and metadata (title of a musical composition, author's name, performer, etc.). Index has a two-level structure. It consists of melodic contour index, which acts as a hash index for quick search, and of full melody index, which contains information about music intervals and melody rhythm. At present the database contains 2.7 thousand of popular songs of various genres and melodic themes of classic music.

2.3 Specifics of System Operation

Analysis of similar systems [4] showed that Sloud QBH has some peculiarities in the recognizing and searching parts dedicated to increase of velocity and search effectiveness.

Voice recognizing part of the system interface works in real time and has a color indication of the recognized pitch contour. Due to this a user directly in the process of humming can see the result of conversion on the indicator. This provides a visual feedback, which contributes to more precise humming and decreases the number of humming errors (see Figure 2.). As is known, humming quality influences much the search effectiveness [5].

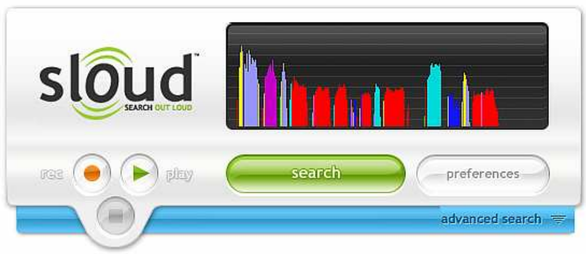

Figure 2. Sloud QBH Search Engine ActiveX window

An advantage of such a structure is also the fact that the query is prepared on the client side of the system. In this case the query is very short. Besides, there is a possibility to evaluate its quality before sending to the server. The system provides for playback of the recognized melody notes in MIDI format. This allows the user to listen to a query and take a decision either to send it to the server or to sing it once again.

In order to hasten the search, on the first stage the system selects a certain subset of musical compositions by melodic contour index. In order to make this fast index we used a 3-position code (-1; 0; +1). On the second stage, the system determines degree of melody matching. It is realized through correlation methods separately for pitch and rhythmic components. Both of these components contain information about melody, but user can remember and reproduce pitch or rhythmic movement of the melody not equally, therefore the system supplies both separate and unified searches by these two components. In a number of cases this allows to reach a higher relevance.

Either while humming or during music recognition mistakes can occur, to them relate: missed and excess notes, mistuning, and inaccuracy of melody rhythm. In order to successfully treat queries, containing such errors, the present system applies a method of partial queries: the main query is divided into several overlaid queries, which are treated separately.

3. Comparison of Online QBH Systems Efficiency

3.1 Features of Tested Systems

Nowadays there are more than twenty experimental QBH systems; a full list of them is given on site MIR System [6]. Their structure can be different, but all of them fulfill the function of search music by melody brought in through microphone by humming or whistling.

Among them it's possible to distinguish several QBH systems with Web-interface, which is often implemented in a form of Java applet. In Table 1 there are several such systems, most similar to the Sloud QBH by the basic functions. Since in all QBH systems the search is implemented in two stages: query processing from singing and in fact search of melody by a set of data, into which singing is converted, the search result essentially depends on the quality of query [10].

Table 1. Features of tested QBH systems

Name	Query Control	Conversion Output	Features Used for Matching	Matching	Access Method	Source
Musipedia (Tuneserver/Melodyhound)	visual	symbolic, gross contour of melody	melody contour	approximate, complete	Java applet	[7]
MelodieSuchmaschine (Fraunhofer)	visual, humming record (WAV)	symbolic, notes pitch and duration	pitch, contour	approximate	Java applet	[8]
Query by Humming at NYU	humming record (WAV)	audio	pitch, contour, key signature, time signature, rhythm	approximate, complete	Java GUI client program	[9]
Sloud QBH Search	visual, converted humming (MIDI)	symbolic, notes pitch and duration	contour, intervals, durations	approximate, partial	ActiveX applet	present paper

It is influenced both by vocal qualities and musical experience of a singer and accuracy of Humming-to-MIDI conversion in the systems with symbolic melody introduction. Usually it is quite difficult to control humming quality connected with musical experience of a singer.

As a rule, the system developers suggest to visually evaluate quality of recognition after the end of humming treatment (see for instance [7]). Sloud QBH, in contrast to similar systems, has a possibility to control humming accuracy in real time, as well as evaluate quality of a query by ear.

3.2 Criteria and Conditions of Test Operation

3.2.1 Testing of recognizing part

Independently on the structure of search system, quality of voice conversion into a set of data, which is treated by a search system, has a greatest influence on the effectiveness of QBH systems. The compared QBH systems (see Table 1) have a different structure and don't let to separately evaluate the efficiency of a query processing subsystem. That is why in order to evaluate quality of Humming-to-MIDI conversion of Sloud QBH, we have conducted a preliminary test with several specialized software, which perform such conversion.

For testing we have selected the following programs: Solo Explorer [11], Wave Goodbye [12], Amazing MIDI [13], TS-Audio-to-MIDI [14], Widi Pro [15], IntelliScore [16]. These software products have been on the market of recognizing software for several years. All of them recognize humming recorded in wav-file and convert it into MIDI sequence.

In test we have used two melody samples, hummed by a woman and recorded in format PCM (44.1 kHz, 16 bit, mono). After conversion of test wav-file into MIDI-file by each of the programs, the quantity of the arising errors was counted. To each kind of an error we gave *w* - weight number in accordance with an influence degree, which it can make on the effectiveness of QBH search. We have evaluated conversion accuracy, which we computed as A *= (100% – (sum of errors % of all types multiplied by w))*. The results are given in Table 2. It turned out that on average conversion accuracy of the recognizing part of Sloud QBH is the highest in the test.

3.2.2 Testing of QBH search effectiveness

Evaluation of QBH search effectiveness was carried out for the above-selected systems (see Table 1.) on the same specially prepared test samples. In this test we assumed, that quality of humming recognition of the tested systems is the same.

For the testing we have used records of melodies fragments of popular compositions of the Beatles, performed by five common people (see Table 3). Records format was: PCM 44.1 kHz, 16 bit, mono, normalized by volume. We used melodies of three types: a) melodies hummed clearly, without distortions; b) melodies hummed falsely; c) one melody sample ("Maxwell's Silver Hammer") with an excess note. There was only one requirement – to sing using the syllables [t@] or [d@]. No other requirements such as compliance with musical key or tempo for a melody performance were made. All these testing files were played back through system mixer, imitating humming record for the input of the tested systems through microphone.

Table 2. Accuracy of Humming to MIDI conversion

Kind of conversion error \ Name		Sloud		Solo Explorer		Wave Goodbye		Amazing MIDI		TS-Audio-to-MIDI		Widi Pro		IntelliScore	
sample: Happy Birthday (26 notes)	w	numb.	%	numb.	%	numb.	%	numb.	%	numb.	%	numb.	%	numb.	%
Octave pitch error	0.7	2	7.7	0	0	1	3.8	0	0	0	0	0	0	0	0
Semitone pitch error	0.8	1	3.8	1	3.8	2	7.7	3	11.5	2	7.7	2	7.7	4	15.4
Irrelevant notes	1.0	1	3.8	1	3.8	11	42.3	3	11.5	15	57.7	12	46.2	20	76.9
Missing notes	0.9	0	0	2	7.7	0	0	5	19.2	3	11.5	1	3.8	0	0
Conversion accuracy, %			88		86		49		62		26		44		11
sample: Augustine (25 notes)	w	numb.	%	numb.	%	numb.	%	numb.	%	numb.	%	numb.	%	numb.	%
Octave pitch error	0.7	1	4.0	0	0	2	8.0	0	0	0	0	0	0	4	16.0
Semitone pitch error	0.8	0	0	0	0	0	0	1	4.0	0	0	0	0	0	0
Irrelevant notes	1.0	0	0	3	12.0	6	24.0	11	44.0	4	16.0	13	52.0	7	28.0
Missing notes	0.9	0	0	0	0	2	8.0	1	4.0	2	8.0	2	8.0	0	0
Conversion accuracy, %			97		88		63		49		77		41		61
Average conversion accuracy, %			92		87		56		56		51		43		36

Table 3. Melodies used in test

Tune #	Excerption from	Number of notes	Performer	Singing quality
1	Girl	22	woman	Off-key
2	Let it be	17	man	Clear
3	Michelle	13	man	Clear
4	Yellow Submarine	16	woman	Very off-key
5	Yesterday	18	man	Off-key
6	Girl	23	man	Clear
7	Maxwell's Silver Hammer	15	man	Note insertion

3.3 Test Results

The test results are given in Table 4, where in the column for each tested system a number denotes position of a melody found in the list by relevance position. If the system could not find the melody at all, we put dash in the cell.

As can be seen from the table, if a melody is hummed accurately (2, 3 and 6 melodies), most systems find it and put on first position by relevance. The matter gets on rather worse with the search results, when a query melody is hummed falsely (1, 4 and 5 melodies).

Table 4. Search performance in tested QBH systems

Tune #	Relevance of search result			
	Melodie Suchmaschine	Musipedia	Query by Humming at NYU	Sloud QBH Search
1	1	-	-	1
2	1	-	-	1
3	1	1	-	1
4	4	5	-	10
5	3	-	-	1
6	1	1	-	1
7	1	-	-	1

But if the number of notes in a musical phrase of a query differs from the number of notes in a musical phrase of the original in the database, than only two systems cope with such a query. Obviously it is connected with the peculiarities of that part of search systems, which can demand precise coincidence of notes quantity [17].

4. Summary and Conclusions

In this work we have conducted a comparison research of the productivity and robustness of online QBH systems to queries of different quality. Following the conducted test we can conclude that the worked out Sloud QBH system by search effectiveness can be compared with the best of them.

In future we plan to improve hash index in order to increase robustness of the system, as well as enlarge music database up to 10 thousand records (songs), this will allow using the search engine in action. Also it is significant that such database should contain popular (favorite) melodies for the last 20 years, in order to be demanded among the most part of Internet users.

References

[1] B. S. Manjunath, Philippe Salembier, and Thomas Sikora, editors. *Introduction to MPEG-7*. Wiley, 1st edition, 2002.

[2] R. B. Dannenberg, W. P. Birmingham, and G. Tzanetakis et. al. “The MUSART Testbed for Query-By-Humming Evaluation” in *ISMIR 2003: Proceedings of the Fourth International Conference on Music Information Retrieval*, Baltimore: Johns Hopkins Univeristy, 2003, pp. 41-50

[3] “Sloud Search Service,” [Web site] 2006, [2006 April 10], Available: http://www.sloud.com/

[4] Lie Lu, Hong You and Hong-Jiang Zhang. "A New Approach to Query by Humming in Music Retrieval," in *Proc.of IEEE International Conference on Multimedia and Expo (ICME01)*, 2001, pp. 776-779.

[5] S. Pauws. "CubyHum: A fully operational Query by Humming System," in *Proceedings of the Third International Conference on Music Information Retrieval, ed. Michael Fingerhut (Paris: IRCAM -- Centre Pompidou, 2002)*, 2002, pp.187-196.

[6] “MIR Systems,” [Web site] 2006, [2006 April 15], Available: ttp://mirsystems.info/index.php?id=mirsystems

[7] “MIR Systems: Musipedia,” [Web site] 2006, [2006 April 5], Available: http://mirsystems.info/index.php?id=432

[8] “MIR Systems: MelodieSuchmaschine (Fraunhofer),” [Web site] 2006, [2006 April 5], Available: http://mirsystems.info/index.php?id=418

[9] “MIR Systems: QBH at NYU,” [Web site] 2006, [2006 April 5], Available: http://mirsystems.info/index.php?id=434

[10] R.B. Dannenberg and N. Hu. "Understanding Search Performance in Query-by-Humming Systems,". in *Proc.ISMIR 2004, Spain*, 2004, pp. 205 - 211.

[11] “Wav to midi conversion software – Solo Explorer,” [Web site] 2006, [2006 April 10], Available: http://www.recognisoft.com/

[12] “Wave Goodbye – Wave to Midi Conversion,” [Web site] 2006, [2006 April 10], Available: http://www.btinternet.com/~irshatwell/WaveGoodbye/intro.html

[13] “AmazingMIDI. WAV to MIDI converter for music transcription,” [Web site] 2006, [2006 April 11], Available: http://www.pluto.dti.ne.jp/~araki/amazingmidi/

[14] “TallStick Software: TS-AudioToMIDI,” [Web site] 2006, [2006 April 11], Available: http://www.audioto.com/

[15] “WIDI Recognition System,” [Web site] 2006, [2006 April 11], Available: http://www.widisoft.com/

[16] “IntelliScore Polyphonic. WAV to MIDI Converter,” [Web site] 2006, [2006 April 11], Available: http://www.intelliscore.net/.

[17] R. Typke, F. Wiering and R.C. Veltkamp. “A Survey of Music Information Retrieval Systems,” in *Proc. ISMIR 2005, London*, 2005, pp.153-160.