

SYSTEM AND METHOD FOR RECOGNIZING BROADCAST PROGRAM CONTENT

FIELD OF THE INVENTION

5 The present invention relates to a system and method for recognizing broadcast program content. The system and method are particularly suited, but not limited to recognize broadcast program content such as music which a communication device user is tuned to and will be described in this context.

BACKGROUND ART

10 The following discussion of the background to the invention is intended to facilitate an understanding of the present invention only. It should be appreciated that the discussion is not an acknowledgement or admission that any of the material referred to was published, known or part of the common general knowledge of the person skilled in the art in any jurisdiction as at the priority date of the invention.

15 Current music recognition systems and methods typically involve the use of music recognition engines. Such music recognition engines typically employ some form of music recognition algorithm. A known music recognition algorithm typically obtains an audio sample, compares the audio sample with the entries in its database, and returns certain information available on the
20 identified audio sample.

Current music recognition systems remained largely as a stand-alone function or as an online service accessible by multiple users. To enhance its capability to handle simultaneous requests from multiple users, online music recognition systems are generally implemented as a server farm with load balancing; that
25 is, the music recognition system will have many instances of itself replicated in the server farm which share in the load when there is a deluge of simultaneous requests. Such implementation, however, assumes high bandwidth network connectivity. Online recognition engines, despite having a

high volume of requests, would still provide decent response times because they enjoy substantial Internet bandwidth.

However, when the current music recognition systems are implemented in a mobile service context, the model of server farm may not work as well. This is because of the limited bandwidth of mobile networks, which are usually GSM networks. In a mobile service supporting millions of users, the server load may become too heavy for the recognition engine to handle, thus resulting in unacceptable user experience and/or actual service failure. Such clogs and bottlenecks will also adversely affect non-data use of the network. Poor quality of service contribute strongly to higher subscriber churn, which in this age of near market saturation and stiff competition, could lead to a mobile network operator's demise.

In addition to the above, current prior art system in the context of mobile services requires the user to capture a snippet of this song as an audio clip and send the same to the song recognition engine. An audio clip, no matter how brief and regardless of the compression technique, is of considerable size.

The present invention seeks to provide a content recognition system and method that alleviates the above mentioned drawbacks.

SUMMARY OF THE INVENTION

This invention was developed to address the challenge of reducing overhead that comes with music recognition requests, thus avoiding clogs and bottlenecks in the relatively low-bandwidth GSM network.

In accordance with a first aspect of the present invention, there is provided a broadcast program content recognition system comprising at least one receiver adapted to sample broadcast program content from a broadcast source;
a content recognition engine for recognizing and storing the sampled broadcast program content; and

a content database in data communication with the content recognition means; the content database adapted to return information relating to the broadcast program content upon receipt of a query from a client device.

Preferably, the client device is a mobile device adapted to receive the broadcast program content.

Advantageously, the query is a SMS query or a HTTP post query. The query comprises the time stamp of the broadcast program content and an identification associated with the broadcast source.

Preferably, the client device is configured to automatically send passive information on the station ID and the time stamp to the content database on a regular time interval. In such a configuration, the content database is further in data communication with an application content manager adapted to process the passive information received to customize broadcast program content to the client device.

Preferably, the passive information could be sent through via SMS, MMS, IP, proprietary messaging, or other available wireless connectivity such as Wi-fi, Bluetooth or Near Field Communication (NFC).

Preferably, the system further comprises a profiling database in data communication with the content database, wherein information from the content database and profile database is adapted, aggregated and consolidated to arrive at certain user-specific conclusion.

In accordance with a second aspect of the present invention, there is provided a broadcast program content recognition system comprising at least one receiver adapted to sample broadcast program content and a content recognition engine for recognizing and storing the sampled broadcast program content; wherein in event where the content recognition engine is unable to recognize the sample broadcast program content, the content recognition means splits the unrecognized sample into at least a first and a second sequential portions and appends the first portion or the second portion to a previously recognized sample.

Preferably, the system is adapted to iteratively split and append the unrecognized sample until either a terminating condition is reached or the appended first or second portion is recognizable.

- 5 Preferably, the system is adapted to mark the unrecognized sample as a failed sample.

In accordance with a third aspect of the present invention there is provided a method of recognizing broadcasted program content comprising the steps of:

- a. receiving a sample of broadcasted program content;
10 b. determining if the received sample is recognizable;
c. splitting the received sample into a first and a second sequential portions if is determined not to be recognizable; and
d. appending the first portion to a previously recognizable sample.

Preferably, the method includes the step of repeating steps (b.) to (d.) until the
15 appended sample is recognizable.

BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

- 20 Fig. 1 is a schematic representation of a content recognition system in accordance with a first embodiment of the present invention.

Fig. 2 is a schematic representation of a content recognition system in accordance with a second embodiment of the present invention.

- 25 Fig. 3 is a flow diagram illustrating the sampling algorithm according to embodiments of the present invention.

Fig. 4 is an illustration of a sample falling outside the content duration, which will result in a failure to be recognized by the content recognition system.

Fig. 5 illustrates splitting of an unrecognizable sample in various scenarios.

DESCRIPTION OF EMBODIMENTS OF THE INVENTION

In accordance with a first embodiment of the present invention there is a broadcast program content recognition system 10. The content recognition system 10 comprises a plurality of broadcast receivers 14 each adapted to receive broadcasted program content from one or more broadcast source 12; a content recognition engine 16; and a content database 18. The broadcast sources 12 are typically broadcast stations. For the purpose of illustration, the broadcast sources 12 are frequency-modulated (FM) frequency broadcast stations. Each broadcast station 12 broadcasts program content in a different FM frequency bandwidth. For the purposes of illustration in this embodiment, the broadcast content from each broadcast sources is music, although it is easily appreciated that the broadcast content may be other audio content including audio advertisements etc.

Each receiver 14 is in communication with its corresponding broadcast source 12. It is to be appreciated that each receiver 14 may be co-located at the same area of the broadcast source or may be distributed geographically. Each receiver 14 is adapted to sample the broadcasted music from its corresponding broadcast source 12 continuously. Each receiver 14 is configured to sample broadcast program content at a sampling time t .

The music recognition engine 16 comprises a program which may be a third-party software application as known to a person skilled in the art (for example, SoundHoundTM). The music recognition engine 16 is adapted to receive and process the music samples from each receiver 14. The processing performed by the music recognition engine 16 comprises parsing and identifying the sample within a fixed recognition time.

The content database 18 is constantly populated by the content recognition engine 16. The data transmitted to the content database 18 from the recognition engine 16 may be in any desired format. For example, the content database 18 may have its own set of decoder which decodes the code from the recognition engine 16 and translates the same to pertinent information

such as a song title, genre, artist, etc via a simple database lookup query, for example but not limited to SQL query. The content database 18 is adapted to receive queries from at least one client device 20, each query including time stamp and station ID that corresponds to the broadcast station 12 which the client devices 20 is tuned to. The content database 18 organizes the information received from the recognition engine 16 according to the time received and is thus able to provide a historical records of content received from the content recognition engine 16. The historical records which the user may obtain depends on the size and capacity of the database 18.

10

The client device 20 is typically a mobile device. In this embodiment, the mobile device 20 is enabled with a FM tuner so that the user of the mobile device 20 could tune in to a selected broadcast station 12. The mobile device 20 is adapted to be in data communication with the content database 18, and may query the content database 18 using data query mechanisms such as HTTP POST request, method invocation, keyword-based SMS query etc.

15

The system 10 will now be described in the context of its operation.

As an illustration, the user of mobile device 20 tunes to a particular broadcast station 12 of the plurality of broadcast stations 12 via adjusting the FM tuner incorporated in the mobile device 20. The user wishes to know the title of a broadcasted music tune which he/she is tuned to, hence he accesses the mobile device interface to send a query to the playlist database 18 based on the protocol as mentioned earlier (via HTTP POST, SMS query etc).

20

The content database 18 performs a simple lookup based on two parameters:

25

- a. the time stamp of the query and
- b. the station ID.

The content database 18 checks the time stamp of the query to determine if the time stamp falls within a fixed known sampling interval t_k , which is the time period that the broadcast program content is certain to be playing. If the time

stamp does not fall within the known sampling interval t_k , the content database 18 returns a detection failure error to the client device 20, and optionally prompts the user of client device 20 to retry.

Typically, the value of the sampling interval t is determined by the music recognition algorithm considering the minimum sample time required to successfully identify the music sample. Any time interval shorter would effectively prevent the recognition engine 14 from identifying the audio sample.

However, in order to prevent the sampling from spanning across two different music samples, which prevents the recognition engine from identifying the sample, it is desirable to keep the sampling interval t as short as possible. It thus follows that in order to prevent sampling of overlapping samples, the sampling interval t should ideally be kept less than or at most equal to the playback length of the broadcasted content. Practically however, each broadcasted content will not have the same playback lengths, and thus the sampling interval t should be kept small, and at most equal to the playback length of the shortest broadcasted content. Even so, there will be the possibility of overlaps since the playback lengths of each broadcasted program content will vary. The approach in the embodiment is to match the sampling interval t with the minimum sampling interval required by the content recognition engine 16 employed, and resolve any overlapping samples. In the event where overlaps of sample occur (known as an overlap condition/scenario, a sample algorithm within the content recognition engine 16 will execute the necessary operations to resolve the overlap condition (described subsequently).

Regardless of whether requests are made by user(s), each receiver 14 continuously samples the music broadcasted by its broadcast station 12. The sampling is performed based on the sampling algorithm illustrated in Fig. 3 and described as follows:

The process begins by taking a sample of the broadcasted content from the broadcast source 12 (step 32). Although the sampling is continuous, but one sample is taken at a finite period is understood as the step of taking a sample.

The algorithm checks if the sample is ready (step 34). The checking step includes verifying if a sample is available. By default, the sample is a raw sample. The checking step may alternatively be performed on a transcoded sample (i.e. includes an optional transcoding step which converts the raw sample to a hexadecimal, plain text, or any unique code). In such instance the checking step verifies the availability of the transcoded sample. Once the sample is ready, it is fed to the recognition engine 16 (step 36). The recognition engine 16 performs recognition of the content based on known techniques and checks if the recognition is successful (step 38). If the content recognition is determined to be successful, the information is stored in the content database 18 (step 40). The sample is also retained in a circular queue buffer. The circular queue buffer may be a First-In-First-Out queue system, and is a means for discarding buffered content in case the content is not required anymore. If the content recognition is not successful, then a fail mark is made (step 42). A failed recognition means that the sample falls into the category of the overlap condition and hence requires splitting (step 44), the steps of which are illustrated in Fig. 5. The split sample and a first portion are then appended to the previous sample which is held in a storage buffer queue (step 46). The concatenated sample is then fed to the content recognition engine (c.f. step 36). The process repeats until the sample is recognized and stored (step 40). If the recognition fails, half of the first portion of the earlier split sample earlier will be split and then appended to the previous sample in the buffer queue.

Step 44 is further described below and illustrated in Fig. 5.

Fig. 5 shows the program content A and B being played at a particular broadcast station 12 and the samples 1, 2, 3 and 4 taken while the program content A and B are played consecutively. It is easily appreciated that samples 1, 2 and 4 are definitive, i.e. content A can be ascertained using

samples 1 and 2, while content B can be ascertained using sample 4. Sample 3, however, poses an overlap problem because part of sample 3 is taken while content A is aired and the other part of sample 3 is taken while content B is aired.

5 As a result, the recognition of sample 3 will fail based on step 38 as the music recognition algorithm would not be able to determine whether sample 3 should be associated with content A or content B. Upon detection of the failed sample 3, the sampling algorithm proceeds to mark the failed sample (step 42) and split the sample 3 into sample S3L (the first portion) and sample S3R.

10 Upon appending Sample S3L to Sample 2, the content recognition engine would be able to make a valid recognition of Sample 2+S3L. However, sample S3R would still fail based on step 38 and thus remains as an overlap problem. The failed sample S3R then triggers another splitting of the sample RL and sample RR. It will be noted that appending RR to sample 4 will result

15 in the valid identification of content B. Appending RL to either sample 2 or sample 4, however, will both result in the failure of the identification process. It is appreciated that while the illustration and description has described a scenario of 'left append', 'right append' (i.e. appending sample S3R to sample 3, for example) is equally supported by the system.

20 As illustrated in Figure 5, the interval of certainty increases as the overlap sample, sample 3, is split. Without the splitting step 44, the interval of certainty, that is, the period where the content recognition engine 16 is able to definitely identify the program content A aired, known as the First Interval of Certainty, is shorter because sample 3 falls on a overlap condition. A user

25 query at the time after the First Interval of Certainty and falling within sample 3 would thus return an error.

Applying the splitting process/step 44, however, the interval of certainty is increased. The first portion Sample S3L is added to the First Interval of Certainty, resulting to the Second Interval of Certainty which is longer. Hence,

30 the same user query falling within the Second Interval of Certainty would result in a positive identification of the program content A. For the same point

of query as illustrated, results would be better under the Second Interval of Certainty.

Upon successful recognition, the samples and its information is stored and the recognition engine 16 updates the known interval t_k to include the sampling time t plus the time taken for the appended portion. For each stored sample, the system checks whether a fail mark under step 42 has been made (step 48). If a fail mark is detected, it means that the sample recently processed is a split and concatenated/appended sample (as contrasted to a regular length sample), hence it would be time to move on to the next regular length sample (step 32). When there is no fail mark, the system checks whether there is a further split to process (step 50). If there is no other samples to process, the sampling algorithm proceeds to the next sample (step 32).

The process of splitting (step 44) and appending as described is iterative, but it has a terminating condition. Such condition may be dictated by business rules, such as fixing the number of iterations to a certain number n , or until $\frac{1}{2}t$ is of duration a minimum sampling time t_{min} . t_{min} is the minimum sampling time interval required to have a useful sample; any sample sampled for less than t_{min} will not be recognizable. n could be initially set to 2.

In accordance with a second embodiment of the invention, where like numerals reference like parts, there is a content recognition system 10.

The content recognition system 10 comprises a plurality of broadcast receivers 14 each configured to receive broadcasted program content from a broadcast source 12, a content recognition engine 16 and a content database 18. The broadcast sources 12 are typically broadcast stations. For the purpose of illustration, the broadcast sources 12 are FM frequency broadcast stations. Each broadcast station broadcasts program content in a different FM frequency bandwidth. For the purposes of illustration in this embodiment, the broadcast content from each broadcast sources is music, although it is easily appreciated that they may be other audio content including advertisements etc.

In addition to the first embodiment, there is a mobile network operator or application content manager 900 and a profiling database 950.

Instead of a regular query as described in the first embodiment, a client device 20 sends passive information on the station ID and the time stamp to the content database 18 on a regular basis without the need for the user to actively request the information. This may be done, for example, via the client device settings and will not be described in further detail. The passive information sent thus will be able to reflect if the user of client device 20 has switched to another station 12 (based on switch in the station ID). The passive information could be sent through the GSM network via SMS, MMS, IP, proprietary messaging mechanism etc. which the client device 20 is using, or through other available wireless connectivity such as Wi-fi, Bluetooth, Near Field Communication (NFC), and the like, if the client device 20 is so equipped.

The content database 18 is in data communication with the profiling database 950. Information from the content database 18 and profile database 950 may be further adapted, aggregated and consolidated (data-mined) to arrive at certain user-specific conclusion, for example, showing what specific content makes the user of client device 20 switch channels and what encourages them to stay tuned in, listening preference of the user, although other information could also be tracked, such as songs listened to, length of staying at a particular channel, content being aired when the user changed channels etc. The profiling database 950 feeds this information to the mobile network operator or the application content manager 900. The content manager 900 then is able to customize content for the mobile device 20 by way of suggestive marketing and targeted advertisements, such as for instance the availability for sale of optical media contents of the same genre as that preferred by the subscriber, or an upcoming concert of artists identified in such genre, and the like.

Information from content database 18 and profile database 950 may further be used in a variety of ways. Information on user behaviour can be aggregated

and consolidated to show what specific broadcast program content makes the users of client devices 20 shift to another channel and what encourages them to stay tuned in. Information on the number of listeners tuning in the broadcast stations 12 at any given time would be available, making available information to the application content manager 900 on the best time/period to broadcast a particular program content. The information can then be used by content providers and broadcast stations 12 to determine relevant programming which would engage listeners more. In addition, this embodiment would also be relevant in rating the broadcast stations 12. It is appreciated that traditionally, the information gathering and rating is performed using means ranging from manual random surveys to automated data gathering using devices randomly deployed to households and individuals. With this embodiment, all mobile devices 20 with the built-in tuner may be accompanied with an integrated reporting system that allows the real-time determination of how many client devices 20 (and therefore end users) are listening to a particular broadcast channel 12 at any time. The current embodiment allows advertisers to be better advised on which broadcast channel 12 to use depending on the target audience.

The integrated reporting system mentioned above is implemented in the background and is similar to the passive information sent by the client device 20. The passive information could be a simple notification when a user of client device 20 tunes in to a particular channel, and such notifications are collated in a database report from which could be another service, accessible for free, on subscription basis, pay-per-view or other business models as may be determined subsequently.

The embodiment may further be used to rank the most played music over a defined territory. Without the need for any end user participation, the combined elements 14, 16, and 18 may be used to monitor and rank the music, song or album that is popular at any given time period. Listener density may likewise be determined based on the location and number of mobile devices 20. Further, location information may be provided by the

mobile network operator 900 via the profiling database 950. With the information, geographic profiling can be made to deliver more relevant content and programming.

5 In further embodiments of the present invention the receivers 14 and music recognition engine 16 may be replaced by hybrid broadcast stations.

An associated advantage of the content recognition system 10 is the ability to track historical data. In this regard, when a user queries "what was that song that played recently?", the content database 18 retrieves the information and provides the song information.

10 In addition, the described embodiments is downward compatible with relatively older generation of user device as long as the device is sms enabled. In such a situation, the user may still send a simple text request (including the station id), and the system 10 will reply with a text message on the title of the song that was playing from the radio station, for example.

15

Variants

20 ▪ The profile database 950 and the location-based service arising from the combined elements 14, 16, and 18 can be replaced with any relevant function equivalent, such as a record of previous transactions or events from the same user or any pattern which could be mined from available records.

▪ The continuous sampling by the receivers 14 may be performed 24/7.

25 It is to be understood that the above embodiments have been provided only by way of exemplification of this invention, such as those detailed below, and that further modifications and improvements thereto, as would be apparent to persons skilled in the relevant art, are deemed to fall within the broad scope and ambit of the present invention described. Furthermore although individual

embodiments of the invention may have been described it is intended that the invention also covers combinations of the embodiments discussed.

The Claims Defining the Invention are as Follows

1. A broadcast program content recognition system comprising
at least one receiver adapted to sample broadcast program content from a
broadcast source;
5 a content recognition engine for recognizing and storing the sampled
broadcast program content; and
a content database in data communication with the content recognition
means; the content database adapted to return information relating to the
broadcast program content upon receipt of a query from a client device.
10
2. A system according to claim 1 wherein the client device is a mobile device
adapted to receive the broadcast program content.
3. A system according to claim 1 wherein the query is a SMS query or a HTTP
15 post query.
4. A system according to claim 1 wherein the query comprises the time stamp
of the broadcast program content and an identification associated with the
broadcast source.
20
5. A system according to claim 4 wherein the client device is configured to
automatically send passive information on the station ID and the time stamp to
the content database at regular time interval.
- 25 6. A system according to claim 5 wherein the content database is further in
data communication with an application content manager adapted to process
the passive information received to customize broadcast program content to
the client device.

30

7. A system according to claim 5 wherein passive information could be sent via SMS, MMS, IP, proprietary messaging, or other available wireless connectivity such as Wi-fi, Bluetooth or Near Field Communication (NFC).

- 5 8. A system according to claim 1, the system further comprises a profiling database in data communication with the content database, wherein information from the content database and profile database is adapted, aggregated and consolidated to arrive at certain user-specific conclusion.
- 10 9. A broadcast program content recognition system comprising at least one receiver adapted to sample broadcast program content and a content recognition engine for recognizing and storing the sampled broadcast program content;
wherein in event where the content recognition engine is unable to recognize
15 the sample broadcast program content, the content recognition engine splits the unrecognized sample into at least a first and a second sequential portions and appends the first portion or the second portion to a previously recognized sample.
- 20 10. A system according to claim 9, wherein the system is adapted to mark the unrecognized sample as a failed sample.
11. A system according to claim 9, wherein the system is adapted to iteratively split and append the unrecognized sample until either a terminating condition
25 is reached or the appended first or second portion is recognizable.
12. A method of recognizing broadcasted program content comprising the steps of:
a. receiving a sample of broadcasted program content;
30 b. determining if the received sample is recognizable;
c. splitting the received sample into a first and a second sequential portions if the sample is determined not to be recognizable; and

d. appending the first portion or second portion to a previously recognizable sample.

13. A method according to claim 12 wherein including the step of repeating
5 steps (b.) to (d.) until the appended sample is recognizable.

10

15

20

25

30