

- 1 -

Title Of InventionMONITORING USAGE OF MEDIA DATA WITH NON-PROGRAM DATA ELIMINATIONField Of The Invention

[0001] The invention relates to systems and methods for monitoring usage of media data.

Background Of The Invention

[0002] Existing techniques for determining the program to which an audience is exposed include pattern-matching and encoding techniques. Existing pattern matching techniques necessarily require a large amount of data to be processed to ensure reliable measurement. Encoding techniques involve adding a code to the media data. Encoding is not always practical in every instance and the necessary encoding and decoding equipment can become very complex.

[0003] Therefore, major concerns in audience measurement are the large amounts of data that must be stored and processed in order to obtain accurate measurements as well as the need to provide a simple and relatively inexpensive monitoring device.

Summary Of The Invention

[0004] For this application the following terms and definitions shall apply, both for the singular and plural forms of nouns and for all verb tenses:

- 2 -

[0005] The term “data” as used herein means any indicia, signals, marks, domains, symbols, symbol sets, representations, and any other physical form or forms representing information, whether permanent or temporary, whether visible, audible, acoustic, electric, magnetic, electromagnetic or otherwise manifested.

[0006] The term “amplitude” as used herein refers to values of energy, power, voltage, current, charge, intensity, amount, size, magnitude and/or pressure, however measured or evaluated, whether on an absolute or relative basis, on a discrete or continuous basis, on an instantaneous or accumulated basis, or otherwise.

[0007] The term “media data” as used herein means data which is widely accessible, whether over-the-air, or via cable, satellite, network, internetwork (including the Internet), distributed on storage media, or otherwise, without regard to the form or content thereof, and including but not limited to audio data and video data.

[0008] The terms “audience” and “audience member” mean a person or persons, as the case may be, who access media data in any manner, whether alone or in one or more groups, whether in the same or various places, and whether at the same time or at various different times.

[0009] The terms “survey” and “surveying” mean an activity or activities, or conducting any such activity or activities, which gather data concerning usage of media data by an audience or an audience member, including but not limited to, data concerning the media data accessed and/or data concerning the audience or audience member.

[00010] The terms, “coupled”, “coupled to” and “coupled with” shall each mean a relationship between or among two or more devices, apparatus, files, programs, media, components, networks, systems, subsystems and/or

means, constituting any one or more of (a) a connection whether direct or through one or more other devices, apparatus, files, programs, media, components, networks, systems, subsystems or means, (b) a communications relationship whether direct or through one or more other devices, apparatus, files, programs, media, components, networks, systems, subsystems, or means, or (c) a functional relationship in which the operation of any one or more thereof depends, in whole or in part, on the operation of any one or more others thereof.

[00011] The terms “communicate” and “communication” as used herein include both conveying data from a source to a destination, and delivering data to a communications medium, system or link to be conveyed to a destination.

[00012] The term “presence of media data” as used herein encompasses both the possible presence of media data, as well as the certain presence of media data.

[00013] The term, “processor” as used herein means processing devices, apparatus, programs, circuits, systems and subsystems, whether implemented in hardware, software or both.

[00014] In accordance with an aspect of the present invention, a method is provided for gathering data for producing audience survey data. The method comprises conducting an evaluation of monitored data having intervals which include media data and intervals from which media data is absent to determine whether media data potentially is present in the monitored data of each of a plurality of the intervals; producing media data presence data indicating a potential presence of media data in the monitored data of respective ones of the plurality of the intervals based on the evaluation; and separating first monitored data in which media data potentially

- 4 -

is present from second monitored data in which media data is absent based on the media data presence data.

[00015] In accordance with another aspect of the present invention, a system is provided for gathering data for monitoring usage of media data. The system comprises evaluating means for evaluating monitored data having intervals which include media data and intervals from which media data is absent, to produce media data presence data indicating whether media data potentially is present in the monitored data of respective ones of the intervals; and separation means for separating first monitored data in which media data potentially is present from second monitored data in which media data is absent based on the media data presence data.

[00016] The invention and its particular features and advantages will become more apparent from the following detailed description considered with reference to the accompanying drawings.

Brief Description Of The Drawings

[00017] Figure 1 is block diagram of an embodiment of a system for gathering data concerning media usage as well as for producing audience survey data.

[00018] Figure 2 is a flow chart for use in explaining the operation of the Figure 1 embodiment.

Detailed Description Of Certain Advantageous Embodiments

[00019] Figure 1 is a block diagram for one embodiment of an audience measurement system 10 within the scope of the invention. A monitoring device 12 is provided for receiving monitored data. The monitoring device 12 can be either a single device, stationary at the source to be

- 5 -

monitored, or multiple devices, stationary at the multiple sources to be monitored. Alternatively, the monitoring device 12 can be incorporated in a portable monitoring device that can be carried by an individual to monitor multiple sources as the individual moves about.

[00020] Where acoustic data including media data, such as audio data, is monitored, the monitoring device 12 typically would be a microphone having an input which receives media data in the form of acoustic energy and which serves to transduce the acoustic energy to electrical data. Where media data in the form of light energy, such as video data, is monitored, the monitoring device 12 takes the form of a light-sensitive device, such as a photodiode, or a video camera. Light energy including media data could be, for example, light emitted by a video display. The device 12 can also take the form of a magnetic pickup for sensing magnetic fields associated with a speaker, a capacitive pickup for sensing electric fields or an antenna for electromagnetic energy. In still other embodiments, the device 12 takes the form of an electrical connection to a monitored device, which may be a television, a radio, a cable converter, a satellite television system, a game playing system, a VCR, a DVD player, a portable player, a computer, a web appliance, or the like. In still further embodiments, the monitoring device 12 is embodied in monitoring software running on a computer or other reproduction system to gather media data.

[00021] A processor 14, coupled to the monitoring device 12, is provided for processing the monitored data. A storage device 16, coupled to processor 14, receives data from the processor 14 for storage. A communication device 18, coupled to the processor 14, is provided for transmitting the processed data. Finally, a processor 20 located remotely from the devices 14, 16 and 18 is coupled to the communication device 18 to receive data processed by the device 14 for evaluating the processed data.

- 6 -

[00022] Figure 2 is a flow diagram for use in explaining an operation of the system 10. As indicated at 32, the monitoring device 12 receives data, as explained above, and either supplies the data as received to the processor 14 or transforms the received data to a form usable by processor 14. The processor then carries out a procedure 36 to evaluate the data received from the device 12 in order to determine in which portions media data potentially is present and in which portions it is absent.

[00023] In one embodiment, the received data is divided into time segments and each time segment is separately evaluated for this purpose. In another embodiment, the data is divided into frequency segments and each is separately evaluated. In a further embodiment, the data is divided both in time and frequency and each division is separately evaluated.

[00024] In still another embodiment the data is divided by time and/or frequency and either separate divisions are evaluated alone or multiple divisions are evaluated in one or more stages. Each of the divisions, as well as the multiple divisions, as separately evaluated in any of the above embodiments is referred to herein as an "interval" of the data.

[00025] The evaluations 36 are conducted according to one or more techniques. In certain embodiments, the data is either received by processor 14 as frequency domain data or else transformed to a frequency domain. The frequency domain data is evaluated to determine whether it possesses characteristics of media data and/or noise characteristics. In other embodiments the data is received as time domain data or transformed to time domain data, and then evaluated for characteristics of media data and/or noise. In still other embodiments both time domain data and frequency domain data are evaluated.

[00026] In certain embodiments, the absence of media data is detected by analyzing the frequency content of the data for noise-like

characteristics. Such characteristics include, but are not limited to, a relatively uniform amplitude of the data throughout a selected frequency spectrum or spectra and/or low amplitude levels. Also, the presence of noise is indicated by frequency content characteristic of machinery, power-line hum, or other types of ambient noise. Moreover, noise is often characterized by relatively little amplitude variability over time. Various embodiments of the invention utilize either one or more of these noise characteristics to determine that the data under evaluation is noise.

[00027] Acoustic or audio data constituting media data often include voice data and/or music data. Speech data typically is characterized by relatively few frequency components having amplitudes significantly higher than the noise level as well as by significant variability of data amplitude over time. Music data is often characterized by the prevalence of frequency components representing musical notes as well as amplitude variability over time characteristics of a meter of music. Video data is often characterized by well-defined periodicity at line, field and frame rates. Various embodiments of the invention utilize one or more of these media data characteristics to detect the potential presence of media data.

[00028] Still other embodiments of the invention utilize combinations of noise characteristics and media data characteristics to distinguish media data from noise.

[00029] Accordingly, in certain advantageous embodiments, the data is evaluated in a frequency domain to detect the potential presence of speech data or music data, or else video data, as appropriate. In certain techniques, data patterns are extracted from the data being evaluated, such as frequency content patterns and/or amplitude patterns over time in one or more frequency ranges, and matched with stored patterns representing known media data. In this manner, the potential presence of media data is determined. Such patterns include data sets, such as data representing amplitude variability

over time or frequency, as well as overall amplitude levels over time and/or frequency. Similarly, the presence of noise is detected in certain embodiments based on matching of extracted data patterns with stored patterns representing noise.

[00030] In still other embodiments or in addition to the techniques described above, the presence of speech is detected by utilizing a speech recognition technique simply to detect data patterns characterizing speech, as opposed to the content of the speech.

[00031] Based on one or more of the foregoing evaluations, the processor 14 produces media data presence data which indicates the potential presence of media data in the evaluated data or its absence, in each of a plurality of the data intervals.

[00032] As indicated at 42, the processor 14 proceeds to separate intervals containing potential media data 38 from non-media data intervals 40. The separation process is carried out in any of a number of ways. In certain embodiments a tag is stored with each of the evaluated intervals in storage device 16 indicating whether it contains potential media data or non-media data. As new data is received, the processor 14 overwrites the non-media data based on its tag, while retaining the potential media data.

[00033] In other embodiments, only the potential media data is stored and/or communicated to the remote processor 20, as described below.

[00034] As indicated at 44, the processor 14 in certain embodiments stores the potential media data for subsequent communication to the remote processor by means of the communication device 16 in step 46. In other embodiments the processor 14 communicates the potential media data to the remote processor 20 once it is separated.

[00035] The remote processor 20 processes the received potential media data to produce audience survey data. In certain embodiments, the potential media data is processed by processor 20 to detect an audience survey code therein, which can serve to identify media data or its source or sources, or otherwise characterizes the media data or its relationship to other data. In other embodiments, the processor 20 processes the potential media data by matching it or a signature extracted therefrom with a known stored pattern or patterns, to identify the media data or its sources or sources, or else to derive some characteristic thereof or relationship to other data.

[00036] Although the invention has been described with reference to a particular arrangement of parts, features and the like, these are not intended to exhaust all possible arrangements or features, and indeed many other modification and variation will be ascertainable to those of skill in the art.

What is claimed is:

1. A method for gathering data for producing audience survey data, comprising:

conducting an evaluation of monitored data having intervals which include media data and intervals from which media data is absent to determine whether media data potentially is present in the monitored data of each of a plurality of the intervals;

producing media data presence data indicating a potential presence of media data in the monitored data of respective ones of the plurality of the intervals based on the evaluation; and

separating first monitored data in which media data potentially is present from second monitored data in which media data is absent based on the media data presence data.

2. The method of claim 1, wherein conducting an evaluation comprises evaluating the monitored data of each of the plurality of the intervals in a time domain to determine whether media data potentially is present therein.

3. The method of claim 1, wherein conducting an evaluation of the monitored data and producing media data presence data comprise extracting respective data patterns from the monitored data of each of the plurality of the intervals and matching the extracted data patterns with stored data patterns representing known media data to produce the media data presence data.

4. The method of claim 1, wherein conducting an evaluation comprises evaluating the monitored data of each of the plurality of the intervals in a frequency domain to determine whether media data potentially is present therein.

5. The method of claim 4, wherein conducting an evaluation comprises evaluating the monitored data of each of the plurality of the intervals in the frequency domain to detect the presence of music data therein.

- 11 -

6. The method of claim 4, wherein conducting an evaluation comprises evaluating the monitored data of each of the plurality of the intervals in the frequency domain to detect the presence of noise therein.
7. The method of claim 4, wherein conducting an evaluation comprises evaluating a variability of data amplitudes of the monitored data of each of the plurality of the intervals within the frequency domain to detect the presence of noise therein.
8. The method of claim 4, wherein conducting an evaluation comprises evaluating the monitored data of each of the plurality of the intervals in the frequency domain to detect the presence of speech data therein.
9. The method of claim 1, further comprising the step of producing audience survey data based on the first monitored data.
10. The method of claim 9, wherein producing audience survey data comprises decoding a media data code in the first monitored data.
11. The method of claim 9, wherein producing audience survey data comprises pattern matching.
12. A system for gathering data for monitoring usage of media data comprising:

evaluating means for evaluating monitored data having intervals which include media data and intervals from which media data is absent, to produce media data presence data indicating whether media data potentially is present in the monitored data of respective ones of the intervals; and

separation means for separating first monitored data in which media data potentially is present from second monitored data in which media data is absent based on the media data presence data.

13. The system of claim 12, wherein the evaluating means is operative to evaluate the monitored data of a respective one of the intervals in a frequency domain to determine whether media data potentially is present therein.

14. The system of claim 13, wherein the evaluating means is operative to determine whether media data potentially is present in the monitored data of a respective one of the intervals by detecting the presence of speech data therein.

15. The system of claim 14, wherein the evaluating means is operative to detect the presence of speech data using a speech recognition technique.

16. The system of claim 13, wherein the evaluating means is operative to determine whether media data potentially is present in the monitored data of a respective one of the intervals by detecting the presence of music data therein.

17. The system of claim 13, wherein the evaluating means is operative to detect an absence of media data in the monitored data of a respective one of the intervals by detecting the presence of noise therein.

18. The system of claim 17, wherein the evaluating means is operative to detect the presence of noise based on a variability of data amplitudes in the monitored data of a respective one of the intervals within the frequency domain.

19. The system of claim 12, wherein the evaluating means is operative to evaluate the monitored data of a respective one of the intervals in a time domain to determine whether media data potentially is present therein.

20. The system of claim 12, further comprising storage means for storing the first monitored data.

21. The system of claim 12, further comprising communications means for communicating the first monitored data to a processor located remotely from

the separation means for producing audience survey data based on the first monitored data.

22. The system of claim 12, further comprising a sensor having an input to receive a signal to be monitored to produce the monitored data, the sensor being coupled to the evaluating means to provide the monitored data thereto.

23. The system of claim 22, wherein the sensor is mounted on a personal portable monitoring device.

24. The system of claim 12, further comprising means for producing audience survey data based on the first monitored data.

25. The system of claim 24, wherein the means for producing audience survey data comprises means for decoding a media data code in the first monitored data.

26. The system of claim 24, wherein the means for producing audience survey data comprises pattern matching means.