(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 June 2017 (22.06.2017)

WIPO | PCT

(10) International Publication Number
WO 2017/105998 A1

(51) International Patent Classification:
G10L 21/0208 (2013.01) *G10L 21/0216* (2013.01)

(21) International Application Number:
PCT/US2016/065563

(22) International Filing Date:
8 December 2016 (08.12.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
14/973,274 17 December 2015 (17.12.2015) US

(71) Applicant: AMAZON TECHNOLOGIES, INC. [US/US]; PO Box 81226, Seattle, Washington 98108-1226 (US).

(72) Inventors: AYRAPETIAN, Robert; 410 Terry Avenue North, Seattle, Washington 98109-5210 (US). HILMES, Philip Ryan; 410 Terry Avenue North, Seattle, Washington 98109-5210 (US).

(74) Agent: BARZILAY, Ilan; 2 Seaport Lane, Suite 300, Boston, Massachusetts 02210-2028 (US).

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

(54) Title: ADAPTIVE BEAMFORMING TO CREATE REFERENCE CHANNELS

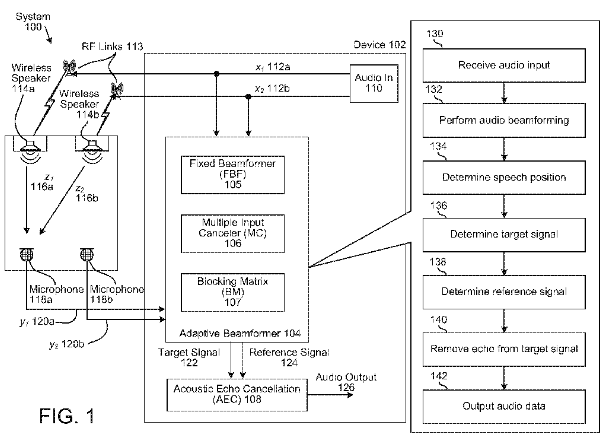

FIG. 1

(57) Abstract: An echo cancellation system that performs audio beamforming to separate audio input into multiple directions and determines a target signal and a reference signal from the multiple directions. For example, the system may detect a strong signal associated with a speaker and select the strong signal as a reference signal, selecting another direction as a target signal. The system may determine a speech position and may select the speech position as a target signal and an opposite direction as a reference signal. The system may create pairwise combinations of opposite directions, with an individual direction being selected as a target signal and a reference signal. The system may select a fixed beamformer output for the target signal and an adaptive beamformer output for the reference signal, or vice versa. The system may remove the reference signal (e.g., audio output by the loudspeaker) to isolate speech included in the target signal.

WO 2017/105998 A1

ADAPTIVE BEAMFORMING TO CREATE REFERENCE CHANNELS

CROSS-REFERENCE TO RELATED APPLICATION DATA

This application claims priority to U.S. Patent Application No. 14/973,274 filed on December 17, 2015 which is incorporated herein by reference in its entirety.

BACKGROUND

In audio systems, automatic echo cancellation (AEC) refers to techniques that are used to recognize when a system has recaptured sound via a microphone after some delay that the system previously output via a speaker. Systems that provide AEC subtract a delayed version of the original audio signal from the captured audio, producing a version of the captured audio that ideally eliminates the “echo” of the original audio signal, leaving only new audio information. For example, if someone were singing karaoke into a microphone while prerecorded music is output by a loudspeaker, AEC can be used to remove any of the recorded music from the audio captured by the microphone, allowing the singer’s voice to be amplified and output without also reproducing a delayed “echo” the original music. As another example, a media player that accepts voice commands via a microphone can use AEC to remove reproduced sounds corresponding to output media that are captured by the microphone, making it easier to process input voice commands.

BRIEF DESCRIPTION OF DRAWINGS

For a more complete understanding of the present disclosure, reference is now made to the following description taken in conjunction with the accompanying drawings.

FIG. 1 illustrates an echo cancellation system that performs adaptive beamforming according to embodiments of the present disclosure.

FIG. 2 is an illustration of beamforming according to embodiments of the present disclosure.

FIGS. 3A-3B illustrate examples of beamforming configurations according to embodiments of the present disclosure.

FIG. 4 illustrates an example of different techniques of adaptive beamforming according to embodiments of the present disclosure.

FIGS. 5A-5B illustrate examples of a first signal mapping using a first technique according to embodiments of the present disclosure.

FIGS. 6A-6C illustrate examples of signal mappings using the first technique according to embodiments of the present disclosure.

FIGS. 7A-7C illustrate examples of signal mappings using a second technique according to embodiments of the present disclosure.

FIGS. 8A-8B illustrate examples of signal mappings using a third technique according to embodiments of the present disclosure.

FIG. 9 is a flowchart conceptually illustrating an example method for determining a signal mapping according to embodiments of the present disclosure.

FIGS. 10A-10B illustrate an example of a signal mapping using a fourth technique according to embodiments of the present disclosure.

FIG. 11 is a flowchart conceptually illustrating an example method for determining a signal mapping according to embodiments of the present disclosure.

FIG. 12 is a block diagram conceptually illustrating example components of a system for echo cancellation according to embodiments of the present disclosure.

DETAILED DESCRIPTION

Typically, a conventional Acoustic Echo Cancellation (AEC) system may remove audio output by a loudspeaker from audio captured by the system's microphone(s) by subtracting a delayed version of the originally transmitted audio. However, in stereo and multi-channel audio systems that include wireless or network-connected loudspeakers and/or microphones, a major cause of problems is when there are differences between the signal sent to a loudspeaker and a signal played at the loudspeaker. As the signal sent to the loudspeaker is not the same as the signal played at the loudspeaker, the signal sent to the loudspeaker is not a true reference signal for the AEC system. For example, when the AEC system attempts to remove the audio output by the loudspeaker from audio captured by the system's microphone(s) by subtracting a delayed version of the originally transmitted audio, the audio captured by the microphone is subtly different than the audio that had been sent to the loudspeaker.

There may be a difference between the signal sent to the loudspeaker and the signal played at the loudspeaker for one or more reasons. A first cause is a difference in clock synchronization (e.g., clock offset) between loudspeakers and microphones. For example, in a wireless "surround sound" 5.1 system comprising six wireless loudspeakers that each receive an audio signal from a surround-sound receiver, the receiver and each loudspeaker has its own crystal oscillator which provides the respective component with an independent "clock" signal. Among other things that the clock signals are used for is converting analog audio signals into digital audio signals ("A/D conversion") and converting digital audio signals into analog audio signals ("D/A conversion"). Such conversions are commonplace in audio systems, such as when a surround-sound receiver performs A/D conversion prior to transmitting audio to a wireless loudspeaker, and when the loudspeaker performs D/A conversion on the received signal to recreate an analog signal. The loudspeaker produces audible sound by driving a "voice coil" with an amplified version of the analog signal.

A second cause is that the signal sent to the loudspeaker may be modified based on compression/decompression during wireless communication, resulting in a different signal being received by the loudspeaker than was sent to the loudspeaker. A third case is non-linear post-processing performed on the received signal by the loudspeaker prior to playing the received signal. A fourth cause is buffering performed by the loudspeaker, which could create unknown latency, additional samples, fewer samples or the like that subtly change the signal played by the loudspeaker.

To perform Acoustic Echo Cancellation (AEC) without knowing the signal played by the loudspeaker, devices, systems and methods may perform audio beamforming on a signal received by the microphones and may determine a reference signal and a target signal based on the audio beamforming. For example, the system may receive audio input and separate the audio input into multiple directions. The system may detect a strong signal associated with a speaker and may set the strong signal as a reference signal, selecting another direction as a target signal. In some examples, the system may determine a speech position (e.g., near end talk position) and may set the direction associated with the speech position as a target signal and an opposite direction as a reference signal. If the system cannot detect a strong signal or determine a speech position, the system may create pairwise combinations of opposite directions, with an individual direction being used as a target signal and a reference signal. The system may remove the reference signal (e.g., audio output by the loudspeaker) to isolate speech included in the target signal.

FIG. 1 illustrates a high-level conceptual block diagram of echo-cancellation aspects of an AEC system 100. As illustrated, an audio input 110 provides stereo audio "reference" signals $x_1(n)$ 112a and $x_2(n)$ 112b. The reference signal $x_1(n)$ 112a is transmitted via a radio frequency (RF) link 113 to a wireless loudspeaker 114a, and the reference signal $x_2(n)$ 112b is transmitted via an RF link 113 to a wireless loudspeaker 114b. Each speaker outputs the received audio, and portions of the output sounds are captured by a pair of microphones 118a and 118b as "echo" signals $y_1(n)$ 120a and $y_2(n)$ 120b, which contain some of the reproduced sounds from the reference signals $x_1(n)$ 112a and $x_2(n)$ 112b, in addition to any additional sounds (e.g., speech) picked up by the microphones 118.

To isolate the additional sounds from the reproduced sounds, the device 102 may include an adaptive beamformer 104 that may perform audio beamforming on the echo signals 120 to determine a target signal 122 and a reference signal 124. For example, the adaptive beamformer 104 may include a fixed beamformer (FBF) 105, a multiple input canceler (MC) 106 and/or a blocking matrix (BM) 107. The FBF 105 may be configured to form a beam in a specific direction so that a target signal is passed and all other signals are attenuated, enabling the adaptive beamformer 104 to select a particular direction. In contrast, the BM 107 may be configured to form a null in a specific direction so that the target signal is attenuated and all other signals are passed. The adaptive beamformer 104 may generate fixed beamforms (e.g., outputs of the FBF 105) or may generate adaptive beamforms using a Linearly Constrained Minimum Variance (LCMV) beamformer, a Minimum Variance Distortionless Response (MVDR) beamformer or other beamforming techniques. For example, the adaptive beamformer 104 may receive audio input, determine six beamforming directions and output six fixed beamform outputs and six adaptive beamform outputs. In some examples, the adaptive beamformer 104 may generate six fixed beamform outputs, six LCMV beamform outputs and six MVDR beamform outputs, although the disclosure is not limited thereto. Using the adaptive beamformer 104 and techniques discussed below, the device 102 may determine the target signal 122 and the reference signal 124 to pass to an acoustic echo cancellation (AEC) 108. The AEC 108 may remove the reference signal (e.g., reproduced sounds) from the target signal (e.g., reproduced sounds and additional sounds) to remove the reproduced sounds and isolate the additional sounds (e.g., speech) as audio output 126.

To illustrate, in some examples the device 102 may use outputs of the FBF 105 as the target signal 122. For example, the outputs of the FBF 105 may be shown in equation (1):

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.