

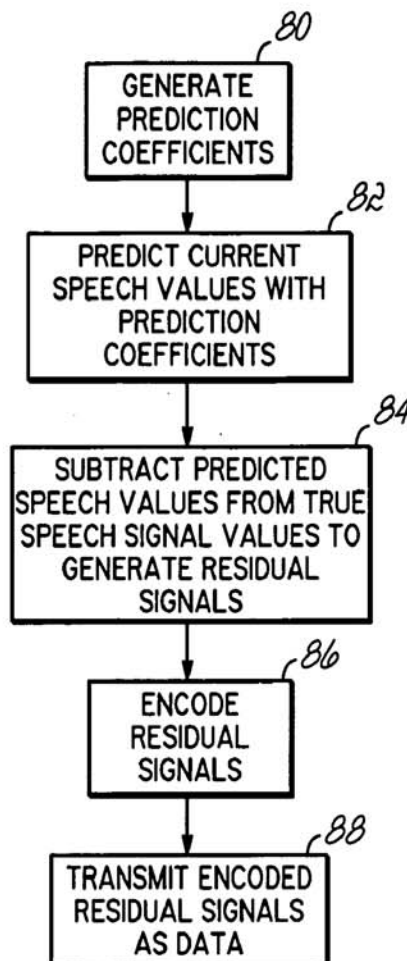


US 20070143105A1

(19) **United States**(12) **Patent Application Publication** (10) **Pub. No.: US 2007/0143105 A1****Braho et al.**(43) **Pub. Date: Jun. 21, 2007**(54) **WIRELESS HEADSET AND METHOD FOR ROBUST VOICE DATA COMMUNICATION**(52) **U.S. CL. 704/231; 704/246**(76) Inventors: **Keith Braho**, Murrysville, PA (US);
Roger Graham Byford, Apollo, PA (US); **Thomas S. Kerr**, Murrysville, PA (US); **Amro El-Jaroudi**, Pittsburgh, PA (US)(57) **ABSTRACT**

A wireless device for use with speech recognition applications comprises a frame generator for generating successive frames from digitized original audio signals, the frames representing portions of the digitized audio signals. An autocorrelation circuit generates a set of coefficients for each frame, the coefficient set being reflective of spectral characteristics of the audio signal portion represented by the frame. In one embodiment, the autocorrelation coefficients may be used to predict the original audio signal to be subtracted from the original audio signals and to generate residual signals. A Bluetooth transceiver is configured for transmitting the set of coefficients and/or residual signals as data to another device, which utilizes the coefficients for speech applications.

Correspondence Address:

WOOD, HERRON & EVANS, LLP
2700 CAREW TOWER
441 VINE STREET
CINCINNATI, OH 45202 (US)(21) Appl. No.: **11/303,271**(22) Filed: **Dec. 16, 2005****Publication Classification**(51) **Int. Cl.**
G10L 15/00 (2006.01)

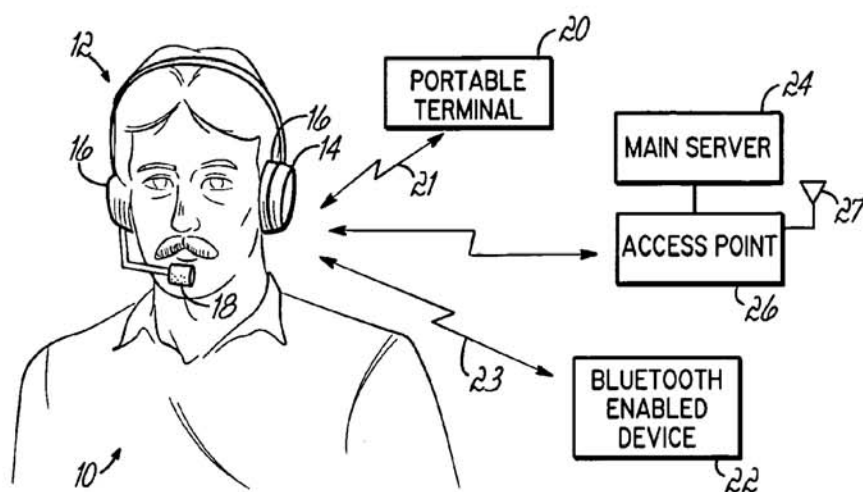


FIG. 1

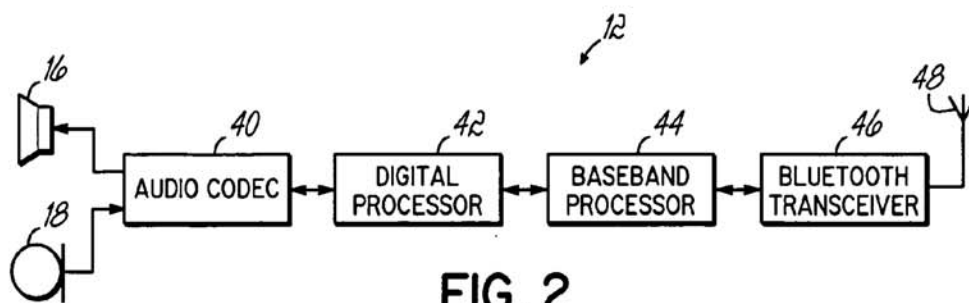


FIG. 2

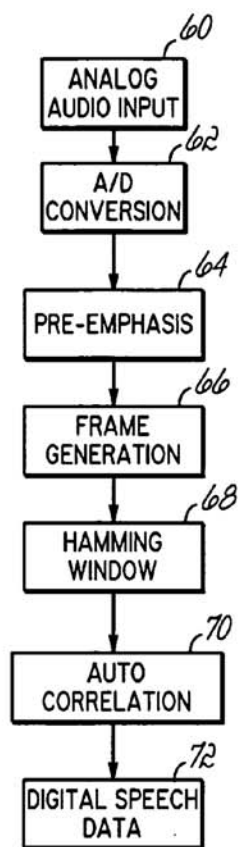
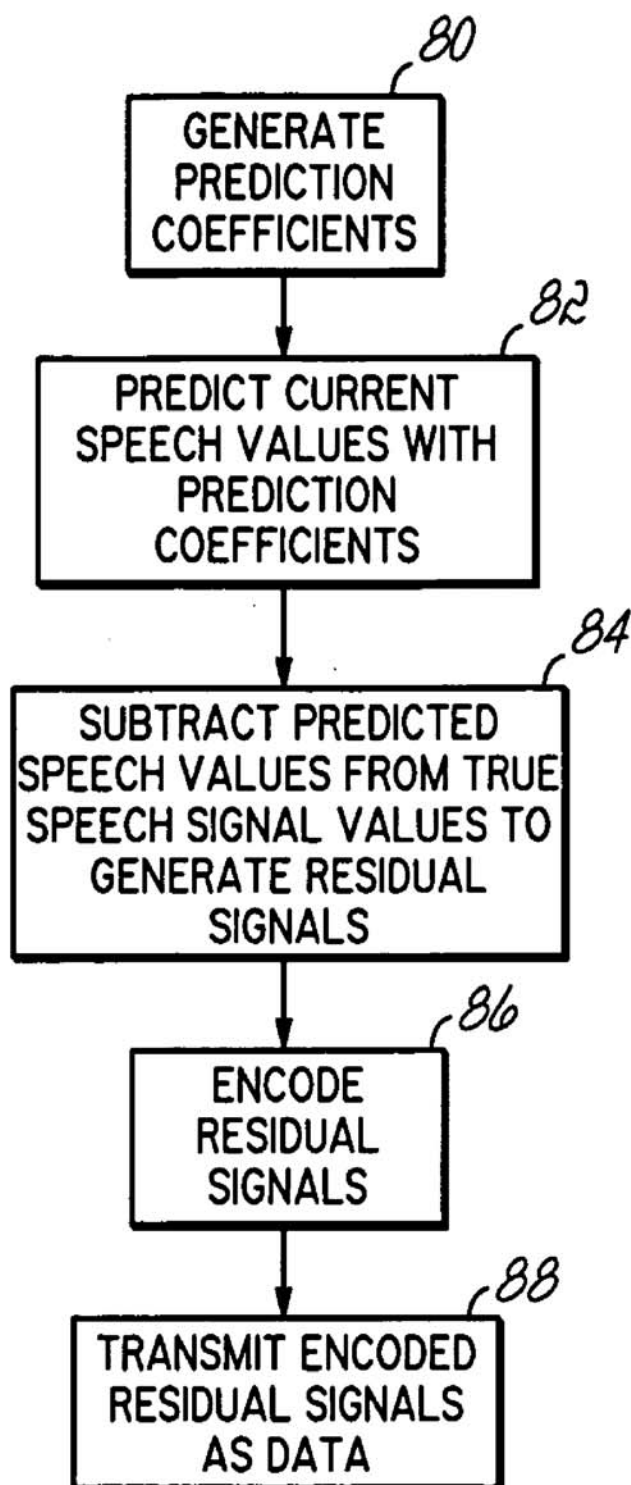


FIG. 3

BLUETOOTH LINK TYPE	BLUETOOTH PACKET TYPE	PAYLOAD SIZE (BITS)	RETRANSMISSIONS SUPPORTED	FOREWARD ERROR CORRECTION	SYMMETRIC MAX RATE (KBPS)	ASYMMETRIC MAX RATE (KBPS)	
ACL	DM1	0-136	YES	YES	108.8	108.8	108.8
ACL	DM3	0-968	YES	YES	258.1	387.2	54.4
ACL	DM5	0-1792	YES	YES	286.7	477.8	36.3
ACL	DH1	0-216	YES	NO	172.8	172.8	172.8
ACL	DH3	0-1464	YES	NO	390.4	585.6	86.4
ACL	DH5	0-2712	YES	NO	433.9	723.2	57.6
SCO	HV1	80	NO	YES	64.0	N/A	N/A
SCO	HV3	160	NO	YES	64.0	N/A	N/A
SCO	HV5	240	NO	NO	64.0	N/A	N/A

FIG. 4

**FIG. 5**

WIRELESS HEADSET AND METHOD FOR ROBUST VOICE DATA COMMUNICATION

FIELD OF THE INVENTION

[0001] This invention relates generally to wireless communication devices, and particularly to a wireless device, such as a headset, utilized for speech recognition applications, and other speech applications.

BACKGROUND OF THE INVENTION

[0002] Wireless communication devices are used for a variety of different functions and to provide a communication platform for a user. One particular wireless communication device is a headset. Generally, headsets incorporate speakers that convey audio signals to the wearer, for the wearer to hear, and also incorporate microphones to capture speech from the wearer. Such audio and speech signals are generally converted to electrical signals and processed to be wirelessly transmitted or received.

[0003] Wireless headsets have become somewhat commonplace. Wireless headsets are generally wirelessly coupled with other devices such as cell phones, computers, stereos, and other devices that process audio signals. In use, a wireless headset may be coupled with other equipment utilizing various RF communication protocols, such as the IEEE 802.11 standard for wireless communication. Other wireless communication protocols have been more recently developed, such as the Bluetooth protocol.

[0004] Bluetooth is a low-cost, low-power, short-range radio technology designed specifically as a cable replacement to connect devices, such as headsets, mobile phone handsets, and computers or other terminal equipment together. One particular use of the Bluetooth protocol is to provide a communication protocol between a mobile phone handset and an earpiece or headpiece. The Bluetooth protocol is a well-known protocol understood by a person of ordinary skill in the art, and thus all of the particulars are not set forth herein.

[0005] While wireless headsets are utilized for wireless telephone communications, their use is also desirable for other voice or audio applications. For example, wireless headsets may play a particular role in speech recognition technology. U.S. patent application Ser. No. 10/671,140, entitled "Wireless Headset for Use in a Speech Recognition Environment," and filed on Sep. 25, 2003, sets forth one possible use for a wireless headset and that application is incorporated herein by reference in its entirety. Speech recognition applications demand high quality speech or audio signal, and thus a significantly robust communication protocol. While Bluetooth provides an effective means for transmission of voice for typical telephony applications, the current Bluetooth standard has limitations that make it significantly less effective for speech recognition applications and systems.

[0006] For example, the most frequently used standard representing voice or speech data in the telephony industry utilizes 8-bit data digitized at an 8,000 Hz sample rate. This communication standard has generally evolved from the early days of analog telephony when it was generally

recent digital voice protocol standards, including the Bluetooth protocol, have built upon this legacy. In order to achieve an upper bandwidth limit of 4,000 Hz, a minimal sample rate of at least twice that, or 8,000 Hz, is required. To minimize link bandwidth, voice samples are encoded as 8 bits per sample and employ a non-linear transfer function to provide increased dynamic range on the order of 64-72 dB. The Bluetooth standard supports generally the most common telephony encoding schemes. At the physical layer, the Bluetooth protocol uses a "synchronous connection oriented" (SCO) link to transfer voice data. An SCO link sends data at fixed, periodic intervals. The data rate of an SCO link is fixed at 64,000 bits per second (64 Kbps). Voice packets transmitted over an SCO link do not employ flow control and are not retransmitted. Therefore, some packets are dropped during normal operation, thus resulting in data loss of portions of the audio signals.

[0007] For most human-to-human communication applications, such as telephony applications, the current Bluetooth voice sampling and encoding techniques using SCO links and voice packets are adequate. Generally, humans have the ability to subconsciously use reasoning, context, and other clues to mentally reconstruct the original speech over a more lossy communication medium. Furthermore, where necessary, additional mechanisms, such as the phonetic alphabet, can be employed to ensure the reliability of the information transferred (e.g., "Z" as in Zulu).

[0008] However, for human-to-machine communication, such as speech recognition systems, significantly better speech sampling and encoding performance is necessary. First, a more reliable data link is necessary, because dropped voice packets in the typical telephony Bluetooth protocol can significantly reduce the performance of a speech recognition system. For example, each dropped Bluetooth SCO packet can result in a loss of 3.75 milliseconds of speech. This can drastically increase the probability of a speech recognition error.

[0009] Additionally, the information-bearing frequency range of speech is now understood to be in the range of 250 Hz to 6,000 Hz, with additional less critical content available up to 10,000 Hz. The intelligibility of consonants has been shown to diminish when the higher frequencies are filtered out of the speech signal. Therefore, it is important to preserve this high end of the spectrum.

[0010] However, increasing the sample rate of the audio signal to 12,000 Hz, while still maintaining 8-bit encoding exceeds the capability of the Bluetooth SCO link, because such an encoding scheme would require a data rate of 96 Kbps, which is above the 64 Kbps Bluetooth SCO rate.

[0011] Speech samples digitized as 8-bit data also contain a high degree of quantization error, which has the effect of reducing the signal-to-noise ratio (SNR) of the data fed to the recognition system. Speech signals also exhibit a variable dynamic range across different phonemes and different frequencies. In the frequency ranges where dynamic range is decreased, the effect of quantization error is proportionally increased. A speech system with 8-bit resolution can have up to 20 dB additional quantization error in certain frequency ranges for the "unvoiced" components of the speech signal. Most speech systems reduce the effect of quantization error

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.