US006260011B1

(54) **METHODS AND APPARATUS FOR AUTOMATICALLY SYNCHRONIZING ELECTRONIC AUDIO FILES WITH ELECTRONIC TEXT FILES**

(75) Inventors: **David E. Heckerman**, Bellevue; **Fileno A. Alleva**, Redmond; **Robert L. Rounthwaite**, Fall City; **Daniel Rosen**, Bellevue; **Mei-Yuh Hwang**, Redmond; **Yoram Yaacovi**, Redmond; **John L. Manferdelli**, Redmond, all of WA (US)

(73) Assignee: **Microsoft Corporation**, Redmond, WA (US)

( * ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/531,054**

(22) Filed: **Mar. 20, 2000**

(51) Int. Cl.$^7$ ........................... **G10L 15/08**; G10L 15/26; G10L 11/06

(52) U.S. Cl. ........................... **704/235**; 704/231; 704/251; 704/254; 704/278

(58) Field of Search .................................... 704/260, 252, 704/270.1, 270, 243, 231, 236, 278, 238

(56) **References Cited**

U.S. PATENT DOCUMENTS

| | | | | |
|---|---|---|---|---|
| 3,700,815 | * | 10/1972 | Doddington et al. ................ | 704/246 |
| 4,779,209 | * | 10/1988 | Stapleford et al. .................. | 704/278 |
| 5,008,871 | * | 4/1991 | Howells et al. ....................... | 369/28 |
| 5,333,275 | * | 7/1994 | Wheatley et al. ................... | 704/243 |
| 5,649,060 | * | 7/1997 | Ellozy et al. ........................ | 704/278 |
| 5,737,725 | * | 4/1998 | Case ..................................... | 704/270 |
| 5,758,024 | | 5/1998 | Alleva ................................... | 395/63 |
| 5,794,197 | | 8/1998 | Alleva et al. ........................ | 704/255 |
| 5,960,447 | * | 9/1999 | Holt et al. ............................ | 704/235 |
| 6,076,059 | * | 6/2000 | Glickman et al. ................... | 704/260 |

OTHER PUBLICATIONS

Hauptmann et al, "Story Segmentation & Detection of Commercials in roadcast News Video", Research & Technology Advances in Digital Libraries, Apr. 24, 1998.*

Pedro J Moreno, Chris Joerg, Jean–Manuel Van Thong, and Oren Glickman, "*A Recursive Algorithm for the Forced Alignment of Very Long Audio Segments*"; Cambridge Research Laboratory, pp. 1–4, Nov. 20, 1998.

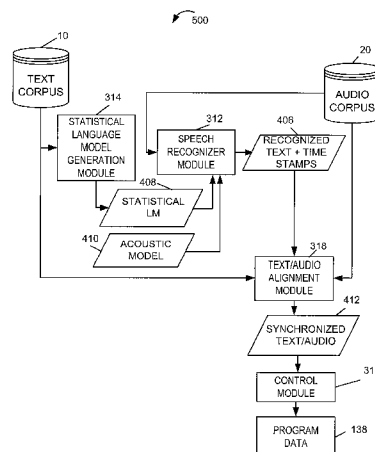Lawrence Rabiner and Biing–Hwang Juang, *Fundamentals of Speech Recognition*, pp. 434–495 (1993).

* cited by examiner

Primary Examiner—Richemond Dorvil
Assistant Examiner—Daniel A. Nolan
(74) Attorney, Agent, or Firm—Straub & Pokotylo; Michael P. Straub

(57) **ABSTRACT**

Automated methods and apparatus for synchronizing audio and text data, e.g., in the form of electronic files, representing audio and text expressions of the same work or information are described. A statistical language model is generated from the text data. A speech recognition operation is then performed on the audio data using the generated language model and a speaker independent acoustic model. Silence is modeled as a word which can be recognized. The speech recognition operation produces a time indexed set of recognized words some of which may be silence. The recognized words are globally aligned with the words in the text data. Recognized periods of silence, which correspond to expected periods of silence, and are adjoined by one or more correctly recognized words are identified as points where the text and audio files should be synchronized, e.g., by the insertion of bi-directional pointers. In one embodiment, for a text location to be identified for synchronization purposes, both words which bracket, e.g., precede and follow, the recognized silence must be correctly identified. Pointers, corresponding to identified locations of silence to be used for synchronization purposes are inserted into the text and/or audio files at the identified locations. Audio time stamps obtained from the speech recognition operation may be used as the bi-directional pointers. Synchronized text and audio data may be output in a variety of file formats.
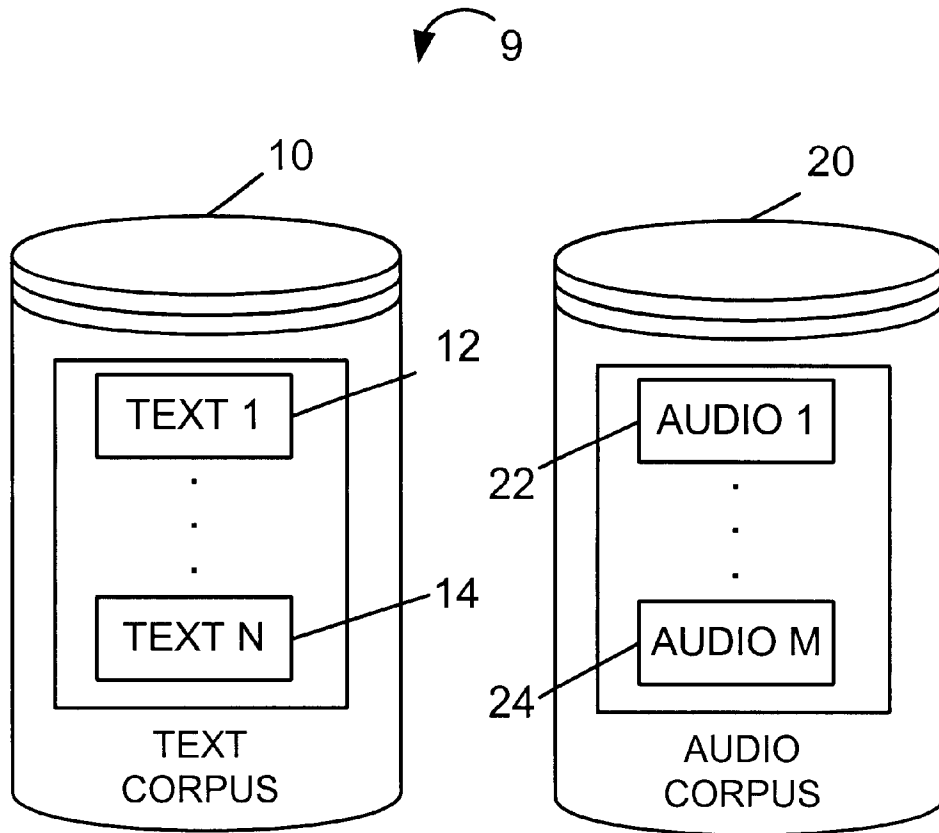
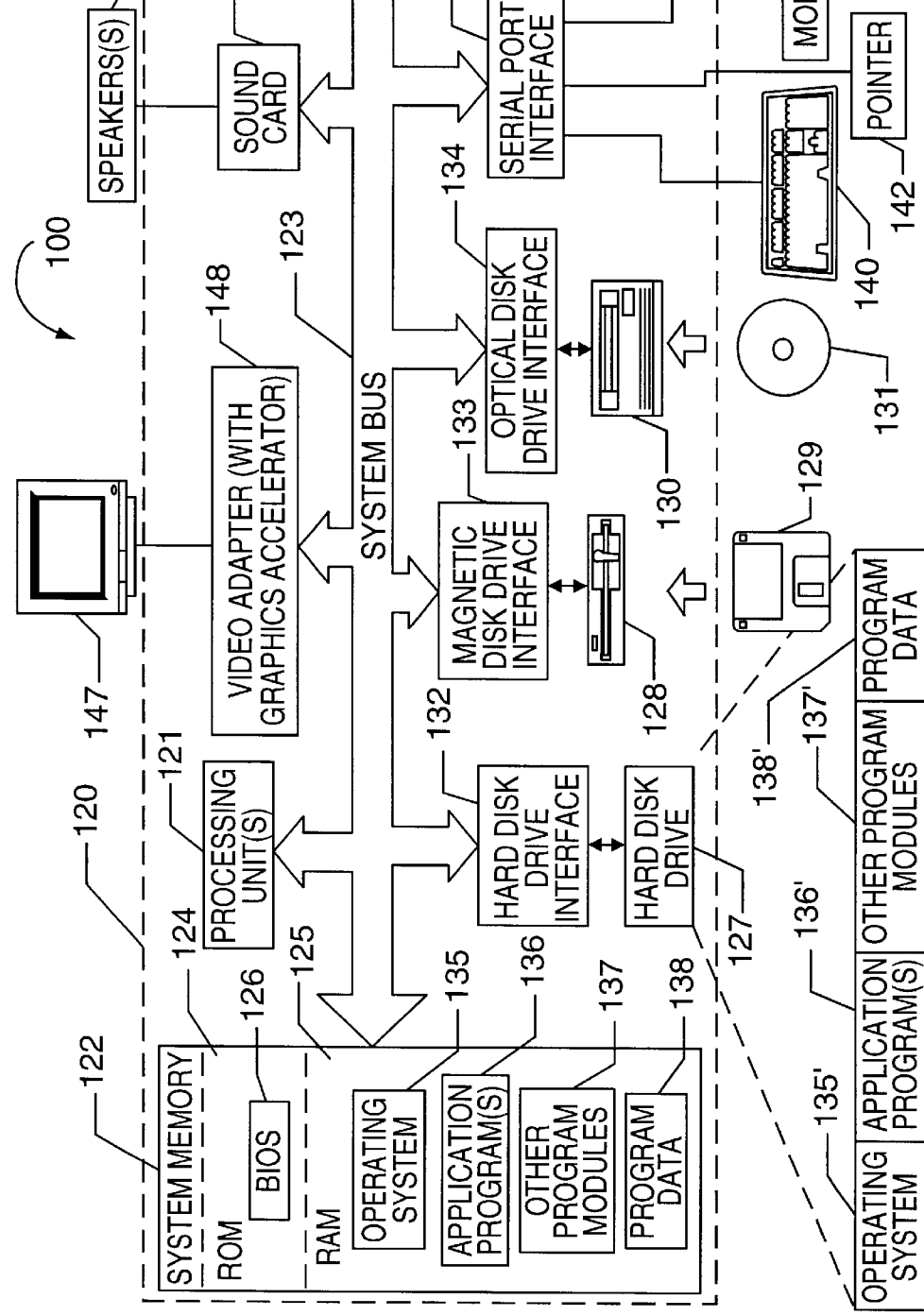**36 Claims, 9 Drawing Sheets**

Fig. 1

**FIGURE 2**

SPEAKERS(S)

SOUND CARD

SERIAL PORT INTERFACE

MO

POINTER

100

134

148

142

140

123

OPTICAL DISK DRIVE INTERFACE

131

VIDEO ADAPTER (WITH GRAPHICS ACCELERATOR)

SYSTEM BUS

133

130

129

MAGNETIC DISK DRIVE INTERFACE

120

147

121

132

128

138'

137'

PROCESSING UNIT(S)

136'

122

124

126

125

135

136

137

138

127

135'

SYSTEM MEMORY

ROM

BIOS

RAM

OPERATING SYSTEM

APPLICATION PROGRAM(S)

OTHER PROGRAM MODULES

PROGRAM DATA

HARD DISK DRIVE INTERFACE

HARD DISK DRIVE

OPERATING SYSTEM

APPLICATION PROGRAM(S)

OTHER PROGRAM MODULES

PROGRAM DATA

136

APPLICATION PROGRAMS

302

WORD PROCESSOR

304

ELECTRONIC BOOK

306

SPREAD SHEET

308

AUDIO/TEXT SYNCHRONIZATION

310

CONTROL MODULE

312

SPEECH RECOGNIZER MODULE

314

STATISTICAL LANGUAGE MODEL GENERATION MODULE

316

SPEECH RECOGNITION TRAINING MODULE

318

TEXT/AUDIO ALIGNMENT MODULE

Fig. 3

138

PROGRAM DATA

402

UNALIGNED TEXT FILES

404

UNALIGNED AUDIO FILES

406

RECOGNIZED TEXT + TIME STAMPS

408

STATISTICAL LM

410

ACOUSTIC MODEL

412

SYNCHRONIZED TEXT/AUDIO

Fig. 4

500

10

TEXT
CORPUS

20

AUDIO
CORPUS

314

STATISTICAL
LANGUAGE
MODEL
GENERATION
MODULE

312

SPEECH
RECOGNIZER
MODULE

406

RECOGNIZED
TEXT + TIME
STAMPS

408

STATISTICAL
LM

410

ACOUSTIC
MODEL

318

TEXT/AUDIO
ALIGNMENT
MODULE

412

SYNCHRONIZED
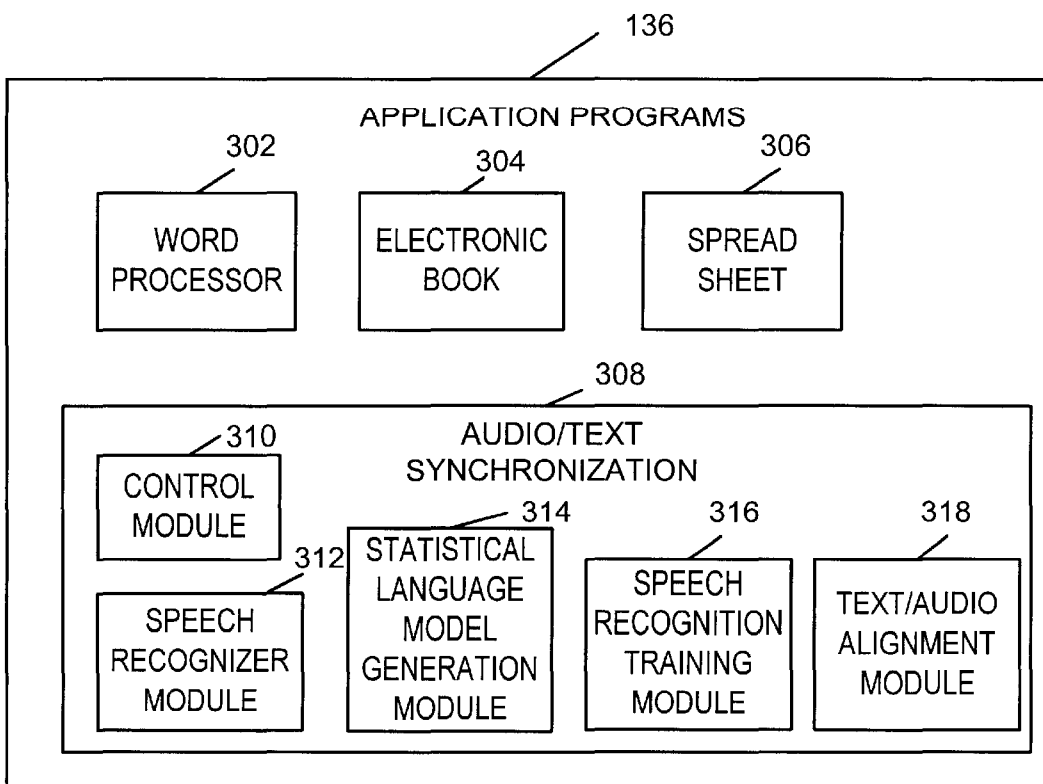TEXT/AUDIO
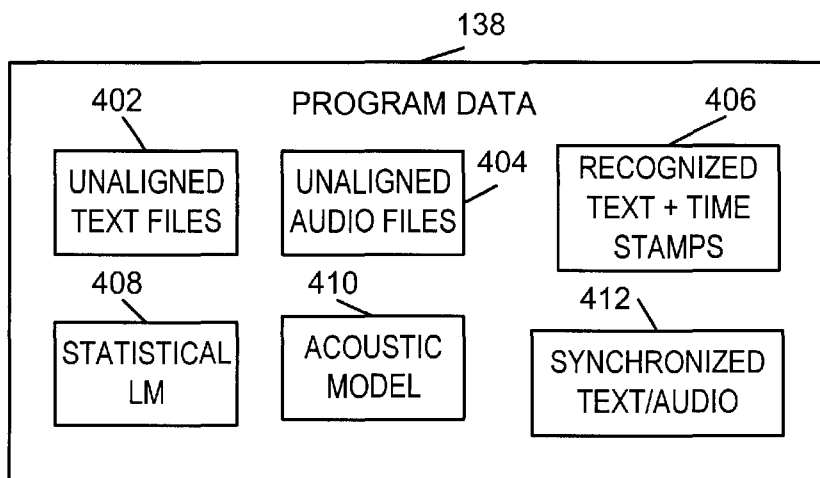
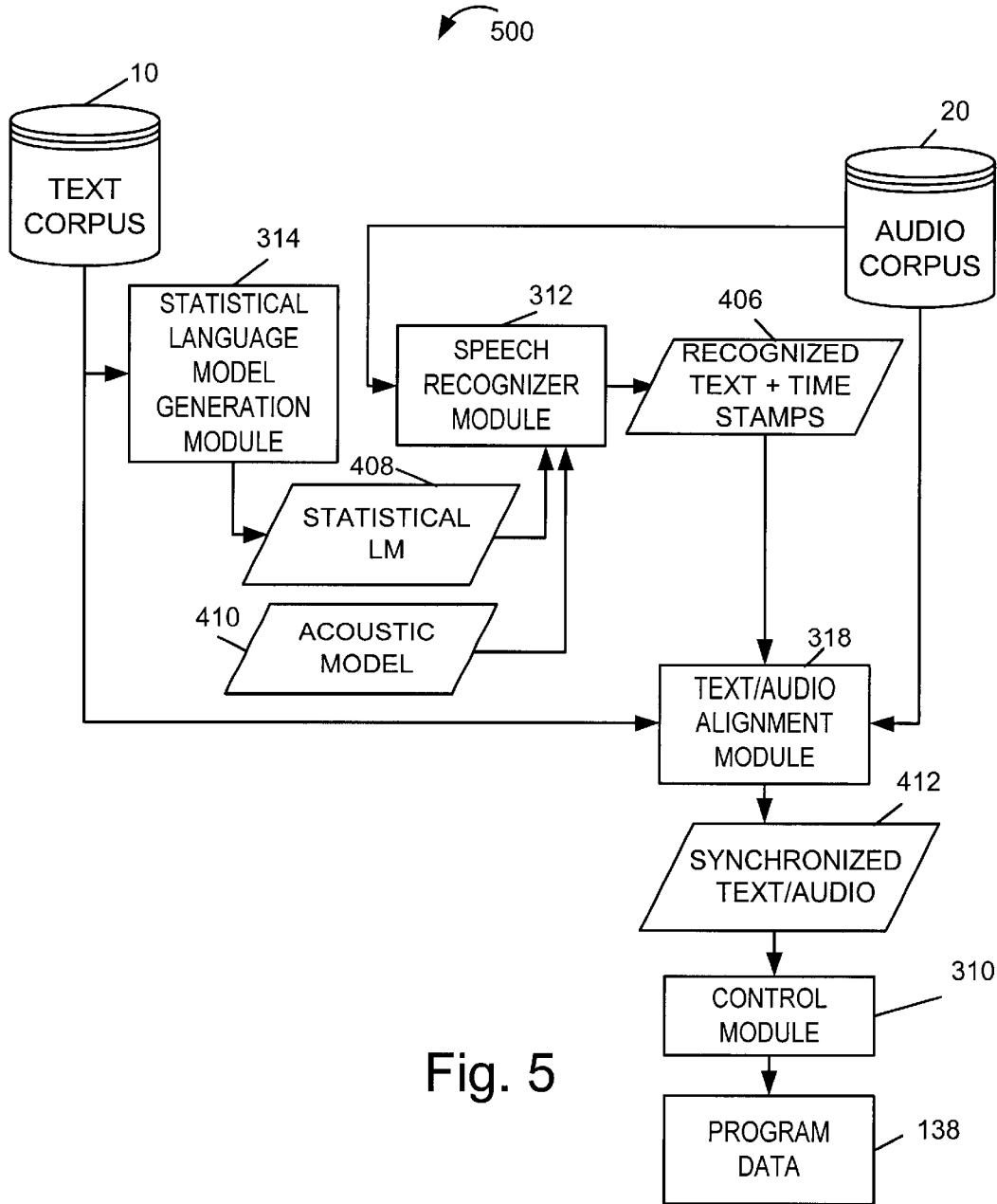310

CONTROL
MODULE

138

PROGRAM
DATA

Fig. 5

# 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.

fastcase
Smarter legal research.