

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-129971

(P2005-129971A)

(43) 公開日 平成17年5月19日(2005.5.19)

(51) Int.Cl.⁷H04N 5/91
// H04N 5/278

F I

H04N 5/91
H04N 5/278

テーマコード (参考)

5C023
5C053

審査請求 未請求 請求項の数 3 O L (全 17 頁)

(21) 出願番号 特願2002-19188 (P2002-19188)
(22) 出願日 平成14年1月28日 (2002.1.28)(71) 出願人 592256623
通信・放送機構
東京都港区芝2-31-19
(71) 出願人 591053926
財団法人エヌエイチケイエンジニアリング
サービス
東京都世田谷区砧一丁目10番11号
(71) 出願人 000004237
日本電気株式会社
東京都港区芝五丁目7番1号
(71) 出願人 000004352
日本放送協会
東京都渋谷区神南2丁目2番1号
(74) 代理人 100083806
弁理士 三好 秀和

最終頁に続く

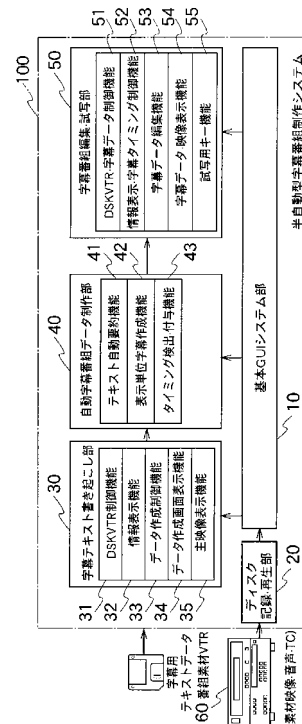
(54) 【発明の名称】 半自動型字幕番組制作システム

(57) 【要約】

【課題】字幕番組制作者による手動字幕制作を効率的に支援する実用性のより高い半自動型字幕番組制作システムを提供する。

【解決手段】テレビ素材番組の映像、音声、タイムコードに基づいてスピーチに対する字幕用テキストの書き起こしおよび背景音などの付加情報データを入力する字幕テキスト書き起こし部30と、書き起こされた字幕テキストに基づいて字幕画面作成およびタイミング付与を行う自動字幕番組データ制作部40と、作成された字幕番組データの編集および試写を行う字幕番組編集・試写部50とを備える。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

テレビ素材番組の映像、音声、タイムコードに基づいてスピーチに対する字幕用テキストの書き起こしおよび背景音などの付加情報データを入力する字幕テキスト書き起こし部と、
書き起こされた字幕テキストに基づいて字幕画面作成およびタイミング付与を行う自動字幕番組データ制作部と、
作成された字幕番組データの編集および試写を行う字幕番組編集・試写部と、
を備えたことを特徴とする半自動型字幕番組制作システム。

【請求項 2】

請求項 1 に記載の半自動型字幕制作システムにおいて、
前記字幕テキスト書き起こし部におけるマンマシンインタフェースとなるメイン画面は、
現在作業中の映像を表示する映像表示エリアと、この映像に対応する書き起こし中の字幕用テキストを表示する字幕用テキスト表示エリアと、前記作業中の映像に対応した一定間隔毎若しくはカット変わりなどの画像と、各画像に対応する書き起こされた字幕用テキストと、各字幕用テキストの書き起こしの基となる音声に関する波形とがタイムライン上に一覧表示された一覧表示エリアと、各エリアに対する制御エリアと、を少なくとも備え、
このメイン画面により作業者の字幕テキスト書き起こし作業を支援する、
ことを特徴とする半自動型字幕番組制作支援システム。

【請求項 3】

請求項 1 または 2 に記載の半自動型字幕番組制作システムにおいて、
前記字幕番組編集・試写部におけるマンマシンインタフェースとなるメイン画面は、
現在編集・試写中の映像を表示する映像表示エリアと、この映像に対応する書き起こし済の字幕用テキストを表示する字幕用テキスト表示エリアと、前記編集・試写中の映像に対応した一定間隔毎若しくはカット変わりなどの画像と、各画像に対応する書き起こされた字幕用テキストと、各字幕用テキストの書き起こしの基となった音声に関する波形とがタイムライン上に一覧表示された一覧表示エリアと、各エリアに対する制御エリアと、を少なくとも備え、このメイン画面により作業者の字幕テキスト編集・試写作業を支援する、
ことを特徴とする半自動型字幕番組制作支援システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、人手による制作機能と自動による制作機能とを効果的に組み合わせた半自動型字幕番組制作システムに関する。

【0002】

〔発明の概要〕

本発明の半自動型字幕番組制作システムは、ディスク記録再生装置に収録してあるテレビ素材番組の映像、音声、タイムコードを効果的に操作し、スピーチに対する字幕用テキストの書き起こしおよび背景音などの付加情報データを効率的に入力する機能と、これらテキストやデータを利用し、自動で字幕画面作成およびタイミング付与を行う自動字幕番組データ制作機能と、字幕番組データの編集および試写を行う編集・試写機能とを備えるものである。

【0003】

この半自動型字幕番組制作システムは、電子化原稿のない番組や背景音レベルの大きい番組など、多様な番組の字幕制作にも対応可能な字幕番組制作システムであり、今後適用番組分野、番組数などの拡大が見込まれる字幕放送の番組制作システムとして、大きな効果が期待される。

【0004】

【従来の技術】

ニュースやナレーション主体のドキュメンタリー番組を対象とし、電子化原稿が存在する

10

20

30

40

50

場合のオフラインで字幕番組を制作する技術として、本発明者らは、「自動要約」、「自動同期」、「自動画面作成技術」などをテーマとして研究を進め、これまでの研究成果を集約した「自動字幕制作システム」を構築した。これらの要点は既に特許出願済である（例えば、特開 2000-270263、特開 2000-324395 等）。

【0005】

この自動字幕制作システムを字幕制作者や聴覚障害者などによって評価した結果、一定の性能、機能があることが確認されたが、実用化のためには、未だ多くの課題が残されている。

【0006】

特に要望の高いのは、対象番組範囲の拡大であり、電子化原稿が存在しない番組、ドラマやバラエティなどの番組分野への拡大である。しかし、現状の自動字幕制作システムの機能・性能から見た字幕制作可能なテレビ番組分野は、かなり限られたものであり、自動機能として限界があるため、この限界以上の部分は、手動による字幕制作や試写・修正の範囲でカバーせざるを得ない。

【0007】

実際の字幕制作現場では、高度な専門技術、知識をもった多くの専門家が携わっており、字幕制作はこのような人間の能力に負っている部分が多くある。

【0008】

一方、字幕番組の急速な拡充が要請されている状況下、字幕制作作業もこのような専門家のみに頼るのみでは不十分であり、専門家ではないワープロ作業が一応できる程度の者であっても作業の一端を分担できるシステムが望ましい。

【0009】

【発明が解決しようとする課題】

しかしながら、課題解決のためには、自動処理を前提とした字幕制作システムのみならず、手作業を含む字幕用電子化テキストの作成や字幕画面の試写・編集などの作業も含めたトータルシステムとして、番組制作効率を考えていく必要がある。

【0010】

そのため、実用的なシステム開発の目標として、1) 多様なテレビ番組の字幕制作を、時間、人手、費用などの点で効率的にできるシステム、2) そのため、研究を進めている自動字幕制作機能と、多彩な支援機能を盛り込んだ手動字幕制作機能とを効果的にタイアップした機能・構成、3) 字幕制作担当者の観点からの作業形態などに関するシステムフレキシビリティの向上などが切望されている。

【0011】

本発明は上記事情に鑑み、これまでに開発された自動字幕制作システムのシステム評価などから得られた知見をもとに、各自動化要素技術を高性能化した新しい自動字幕制作システムを中核にするとともに、字幕番組制作者による手動字幕制作を効率的に支援する実用性のより高い半自動型字幕番組制作システムを提供することを目的としている。

【0012】

【課題を解決するための手段】

上記の目的を達成するために本発明は、請求項 1 では、テレビ素材番組の映像、音声、タイムコードに基づいてスピーチに対する字幕用テキストの書き起こしおよび背景音などの付加情報データを入力する字幕テキスト書き起こし部と、書き起こされた字幕テキストに基づいて字幕画面作成およびタイミング付与を行う自動字幕番組データ制作部と、作成された字幕番組データの編集および試写を行う字幕番組編集・試写部とを備えたことを特徴としている。

【0013】

請求項 2 では、請求項 1 に記載の半自動型字幕制作システムにおいて、前記字幕テキスト書き起こし部におけるマンマシンインタフェースとなるメイン画面は、現在作業中の映像を表示する映像表示エリアと、この映像に対応する書き起こし中の字幕用テキストを表示する字幕用テキスト表示エリアと、前記作業中の映像に対応した一定間隔毎若しくはカッ

10

20

30

40

50

ト変わりなどの画像と、各画像に対応する書き起こされた字幕用テキストと、各字幕用テキストの書き起こしの基となる音声に関する波形とがタイムライン上に一覧表示された一覧表示エリアと、各エリアに対する制御エリアと、を少なくとも備え、このメイン画面により作業者の字幕テキスト書き起こし作業を支援することを特徴としている。

【0014】

請求項3では、請求項1または2に記載の半自動型字幕番組制作システムにおいて、前記字幕番組編集・試写部におけるマンマシンインタフェースとなるメイン画面は、現在編集・試写中の映像を表示する映像表示エリアと、この映像に対応する書き起こし済の字幕用テキストを表示する字幕用テキスト表示エリアと、前記編集・試写中の映像に対応した一定間隔毎若しくはカット変わりなどの画像と、各画像に対応する書き起こされた字幕用テキストと、各字幕用テキストの書き起こしの基となった音声に関する波形とがタイムライン上に一覧表示された一覧表示エリアと、各エリアに対する制御エリアと、を少なくとも備え、このメイン画面により作業者の字幕テキスト編集・試写作業を支援することを特徴としている。

10

【0015】

上記構成の本発明によれば、電子化原稿のない番組や背景音レベルの大きい番組など、多様な番組の字幕制作にも対応可能な字幕番組制作システムが構築でき、今後適用番組分野、番組数などの拡大が見込まれる字幕放送の番組制作システムとして、大きな効果が期待される。

【0016】

20

【発明の実施の形態】

図1は本発明による半自動型字幕番組制作システムの実施形態を示すブロック図である。

【0017】

同図に示されるように、この半自動型字幕番組制作システム100は、基本GUIシステム部10と、ディスク記録・再生部20と、字幕テキスト書き起こし部30と、自動字幕番組データ制作部40と、字幕番組編集・試写部50とを備えている。

【0018】

基本GUIシステム部10は、図1に示す字幕番組制作システムの一構成部品として、マンマシンインターフェース、内部各システムのインターフェースおよび制御、外部とのデータ入出力制御を行う。また、本字幕番組制作システム100の自動字幕番組データ作成機能のインターフェース・制御機能も備えており、多様な番組の効率的な字幕制作を行う字幕番組制作システム100を統括制御するものである。

30

【0019】

ディスク記録・再生部20は、番組素材VTR（例えばデジタル・ビデオ・テープ・レコーダ）60から供給される素材番組の映像・音声を必要ならば圧縮して記録するとともに、記録された映像音声を再生するために使用される。この番組素材VTR60は、番組素材が収録されている番組素材VTRテープから、映像、音声、およびそれらのタイムコード（TC）を再生出力する機能を有しており、これら素材映像、音声およびタイムコード（TC）がディスク記録・再生部20に記録されるようになっている。

【0020】

40

＜字幕テキスト書き起こし部30＞

字幕テキスト書き起こし部30は、素材番組の音声を聞き取って、字幕用テキストの書き起こしや付加データを入力するとともに、素材番組の映像・音声の活用による効果的な手動作業支援をするための機能部分であり、具体的には、ディスク記録再生制御機能31と、情報表示機能32と、データ作成制御機能33と、データ作成画面表示機能34と、主映像表示機能35とを備えている。

【0021】

ディスク記録再生（DSKVTR）制御機能31は、各機能に応じた操作キーを備えており、素材番組の映像・音声を、ディスク記録・再生部20に記録させる記録制御を実行するとともに、記録された映像音声を再生させる再生制御を実行する。特にスピーチ区間で

50

は任意の低速再生を行うとともに、変速率に応じたピッチコントロールを行い、書き起こしに適した速度でかつ音程がほぼ不変とする話速変換機能をも備えている。

【0022】

情報表示機能32は、書き起こしおよび付加情報データ入力の手動作業を支援するため、素材番組の映像・音声、書き起こしテキストなどに関する各種の情報を、タイムライン上にビジュアル表示する。

【0023】

データ作成制御機能33は、書き起こしたテキストやスピーチポーズの時間データなどの入力操作のための操作キーを備え、対応する動作をする。

【0024】

データ作成画面表示機能34は、作成したテキスト・入力したデータの表示・編集などのための画面を表示し、対応する動作をする。

【0025】

主映像表示機能35は、プレビュー画面として、現在作業位置の画像を表示する。

【0026】

＜自動字幕番組データ制作部40＞

自動字幕番組制作システム40は、テキスト自動要約機能41と、表示単位字幕作成機能42と、タイミング検出・付与機能43とを備えている。なお、この自動字幕番組データ制作部40の詳細については、後述する。

【0027】

＜字幕番組データ編集・試写部50＞

字幕番組データ編集・試写部50は、作成した書き起こしおよび付加情報データを基にして自動字幕番組データ制作部40で自動制作された字幕番組データを人手で編集・試写するためのものであり、素材番組の映像・音声、字幕番組データなどの高度活用による効果的な手動作業支援のための特別の機能を備えている。

【0028】

ディスク記録再生・字幕データ制御機能51は、自動制作した字幕番組データおよび記録された映像・音声に関して、指定時間、字幕ページ、字幕ブロックなど編集・試写作業支援用特殊表示操作のための専用操作キーを備え、対応する動作をする。

【0029】

情報表示・字幕タイミング制御機能52は、字幕番組データ編集の手動作業を支援するため、素材番組の映像、字幕本文、音声に関する波形等の各種情報をタイムライン上にビジュアル表示する。

【0030】

特に、字幕番組データについては、タイミング変更支援画面を表示し、この画面に対応する動作の実行が可能とされている。

【0031】

字幕データ編集機能53は、字幕データのページ単位編集のための専用操作キーを備え、字幕データをページ単位で編集する動作を実行可能に構成されている。

【0032】

字幕データ・映像表示機能54は、映像に重畳した指定字幕データ表示のための操作キーを備え、各操作キーに対応する動作が実行可能に構成されている。

【0033】

試写用キー機能55は、部分試写、通し試写など、試写形式の選択に必要な操作キーを備え、各操作キーに対応する動作が実行可能に構成されている。

【0034】

＜半自動型字幕番組制作システム100の機能構成の詳細＞

次に、上記構成を有する半自動型字幕番組制作システム100で実行される処理を字幕テキスト書き起こし部30、自動字幕番組データ制作部40および字幕番組編集・試写部50の各部毎に詳細に説明する。

10

20

30

40

50

【0035】

≪字幕テキスト書き起こし部30での処理≫

字幕テキスト書き起こし部30における主要な作業は、番組のスピーチを聞き取り、その内容を書き起こすとともに、その一区切りに対してその開始・終了タイミングとスピーチ者（話者）をデータ化するものである。

【0036】

したがって、この機能における書き起こしを容易にする支援機能として、作業者が番組のスピーチを正しく聞き取り、自分のペースで書き起こしできるようにすることが基本である。聞き取り易く書き起こしが容易であること、話者・タイミングを把握しやすいこと、書き起こしデータを簡単に修正できることなどが重要である。このように、本システム100の字幕テキスト書き起こし部30には種々の手作業支援機能が盛り込まれている。 10

【0037】

字幕テキスト書き起こし部30においては、字幕テキストファイル（XML）の編集を行う機能を備えている。

【0038】

ここで、字幕テキストとは、範囲時間（IN TIME, OUT TIME）と話者、字幕本文より構成されたデータをいう。

【0039】

編集の補助として、ディスク記録・再生部20に保存されたMPEG／AVI映像ファイルを使用して、音声の聞き取り、タイミングの調整、字幕ポジションの調整を行うことができる。 20

【0040】

図2には、字幕テキスト書き起こし部30で使用されるメイン画面の構成が示されている。同図に示すメイン画面301を使用して、字幕テキストの編集や、各補助機能の操作が行われる。画面のサイズは、例えば、1280×1024（画素数）が望ましい。

【0041】

メイン画面301は、メニュー領域302と、編集領域303と、一覧領域304と、制御領域305の4つの領域に大別される。

【0042】

メニュー領域302は、各種機能の呼び出しを行う領域であり、メニューバーとツールバーに用意された項目を実行することにより、各種機能の呼び出しを行う。 30

【0043】

編集領域303は、字幕データの入力や、入力後の字幕テキストの編集のための領域であり、図3の上段に示されるように、時間枠（IN TIME, OUT TIME）エリア303A、話者エリア303B、字幕本文エリア303Cに一对の入力データとして一覧表示されるようになっている。

【0044】

また、図3下段のデータ入力領域（入力エリア303F）には、入力する内容が表示される。この領域を利用して、例えば話者単位で各作業を行う。

【0045】

“IN TIMEとOUT TIME”の設定は、ページのIN TIME／OUT TIME設定エリア303Dから行う。前のページで設定した時間内に含める場合には、前のページの時間と同じ値のままで確定する。 40

【0046】

“話者”の設定は、メニューの“話者登録”エリア303Eで予め登録された話者の中から選択する。

【0047】

書き起こしの“本文”はページ本文の入力エリア303Fで入力される。

【0048】

追加／挿入ボタンの内、“追加”ボタン303Gは、入力設定されたエリア303D、3 50

03E, 03Fの内容を入力一覧の最後に追加する。“挿入”ボタン03Hは、入力された内容を入力一覧で選択された一对のデータの次へ挿入する。

【0049】

“切り取り”ボタン03J、“コピー”ボタン03Kは、入力一覧で選択された一単位を切り取り／コピーするものである。また、“貼り付け”ボタン03Lは、切り取り／コピー機能で記録された一単位を入力一覧で選択された箇所の次へ貼り付けるために使用される。

【0050】

一覧領域04は、図4に示されるように、画像とこの画像に対応する作成中の字幕テキストを表示する領域であり、MPEG／AVI映像ファイルより一定間隔毎若しくはカット変わりなどの画像一覧を表示する画像エリア04A、入力されたテキストの内容である字幕本文を表示する字幕本文エリア04Bと、作成中の字幕本文に対応するスピーチの音声に対応する波形を表示する波形エリア04Cと、波形エリア04Cに表示される波形などこの一覧領域04の時間指標となるタイムラインが表示されるタイムラインエリア04Dとから成っている。この一覧領域04によって書き起こし作業者の入力支援が効果的に行われる。

【0051】

また、一覧領域04には、MPEG／AVI映像ファイルなどの再生位置カーソル04Eが表示されており、この再生位置04Eを変えることにより、画像の再生位置など書き起こし作業を行うタイムライン上のタイミング変更が可能になる。

【0052】

一覧領域04の具体的な機能は以下の通りである。上述したように、この一覧領域04には、映像ファイルから形成した静止画像と、編集データとしての字幕本文、波形がタイムライン上に一覧表示される。一画面で表示する幅は、“ツール(T)”の“オプション”で選択された120秒、60秒、30秒のいずれかとなる。

【0053】

“画像”は、映像ファイルより、一定間隔の画像を抜き出したもの若しくはカット変わりが表示される。画像のサイズは160×120、間隔は全体幅が120秒の時は20秒、60秒の時は10秒、30秒の時は5秒となる。

【0054】

“字幕本文”は、編集領域03によって入力された本文の内容が表示される。枠の幅は時間幅を表している。幅が狭く、本文が表示しきれない場合には省略される。

【0055】

話者エリア04Fの枠部分をドラッグで変更した場合には、対象一件の幅（時間幅）の変更を、本文の部分04Bの枠をドラッグで変更した場合には、二件の境界線の変更を行うことができる。

【0056】

枠内がダブルクリックされた際、編集領域03が編集途中でなければ、入力一覧04で指定した内容を表示できる位置に移動することができる。

【0057】

“波形”は、映像ファイルに記録されている音声（音楽、効果音を含む）信号を処理し波形化して、図4のように表示する。スピーチやポーズの位置を視覚的に判断できるようにした支援機能である。なお、波形としては、スピーチ区間とポーズ区間とを明確にするために、矩形波状に波形整形して表示するようにしても良い。

【0058】

IN TIME／OUT TIMEの編集単位は、映像ファイルからタイムコードを読み込む際の最小単位を設定する。取り込みを実行した際、設定値未満の値は切り捨てられる。この場合の“設定”は、1フレーム／5フレーム／10フレーム／15フレーム／1秒である。

【0059】

10

20

30

40

50

一覧表示の一画面の幅は、メイン画面の一覧領域 304 にどれだけの時間分を表示するかを設定する。この場合の“設定”は、30 秒／60 秒／120 秒で、それに合わせて画像も 5 秒間隔、10 秒間隔、20 秒間隔となる。

【0060】

一方、制御領域 305 は、MPEG／AVI 映像の制御のための領域であり、MPEG／AVI 映像の表示エリア 305A や、時間枠の取り出しボタン、MPEG／AVI 再生の開始、停止ボタンなどが用意されている。また、制御領域 305 には、映像ファイルの再生に関するインターフェースが用意されている。映像のサイズは、640×480 である。

【0061】

制御領域 305 の具体的な機能としては、例えば、タイムコード表示部 305B と、再生／停止ボタン 305C と、再生位置トラックバー 305D と、MARK IN ボタン 305E と、MARK OUT ボタン 305F と、繰り返し再生ボタン 305G と、話速変換ボタン 305H、<<5 秒／5 秒>> ボタンなどが用意されている。

【0062】

タイムコード表示部 305B に表示されるタイムコードは、プロジェクト新規作成の際に入力されたイニシャルタイムを基準として、映像ファイルの再生位置の時間を足した時間である。なお、映像ファイルの再生位置は、29.97 フレームを 1 秒として扱われる。

【0063】

“再生／停止”ボタン 305C は、映像の再生スタートと再生ストップを行う。なお、“停止”ボタンが実行されても、映像は消えない。

【0064】

再生位置トラックバー 305D は、現在映像ファイルのどの部分を再生しているかを示す。また、トラックバーを動かすことにより、再生位置を変更することも可能である。

【0065】

“MARK IN”ボタン 305E は、編集領域の IN TIME の欄へ、現在の映像の時間をコピーする。“MARK OUT”ボタン 305F は、編集領域の OUT TIME の欄へ、映像の現在の時間をコピーする。

【0066】

“繰り返し再生”ボタン 305G は、編集領域の IN TIME と OUT TIME の期間の画像を繰り返し作成する。

【0067】

終了は、“停止”ボタン 305C か“MARK IN”ボタン 305E をクリックすることによってなされる。

【0068】

“話速変換”ボタン 305H は、映像を再生する際の音声のスピード調節を行う際に使用される。

【0069】

“<<5 秒／5 秒>>”ボタン 305J は、映像の再生位置を前／後ろに 5 秒移動する際に使用される。

【0070】

次に、以上の構成の字幕テキスト書き起こし部 30 で実行される書き起こし処理の手順を図 5 のフローチャートを参照しつつの具体的に説明する。

【0071】

作業者が“PLAY”ボタンを押して、映像再生が開始されると、まず、発話タイミングが探索される（ステップ ST11）。

【0072】

発話の確認点で、“書起開始”ボタンが押されると、これがスピーチ区間の開始点となる（ステップ ST12）。

【0073】

10

20

30

40

50

一定時間巻き戻し、スロー再生が開始され、書き起こし作業が実行される（ステップ S T 1 3, 1 4）。スピーチ終了が認識されると、適宜巻き戻して発話終了点が探索される（ステップ S T 1 5）。

【0074】

発話終了点で“書起終了”ボタンが押されると、（終了点）通常再生モードへ移行する（ステップ S T 1 6）。

【0075】

書き起こし終了後、用字、用語チェック、要約支援が実行されるとともに、背景音情報が登録される（ステップ S T 1 7, 1 8）。

【0076】

テキスト作成処理が終了すると、自動字幕番組データ制作部 4 0 の処理へ移行する（ステップ S T 1 9）。

【0077】

上述した各ステップに示したように、一通りの書き起こしが完了した後、そのテキストの用字、用語チェック、必要ならば要約支援を実行して必要なテキストの修正を行い、その後背景音情報などをチェックし情報を登録する。登録後、テキスト作成関係の作業を終了し、自動字幕番組データ制作部 4 0 の処理へ進むのである。

【0078】

このように、この実施の形態は、新しい半自動型字幕番組制作システムの書き起こし支援を提示するものであり、以下のような利点を有する。

【0079】

すなわち、テレビ素材番組の映像、音声、タイムコードをノンリニア操作が容易なディスク記録再生装置に収録するとともに効果的に操作・処理することによって、スピーチに対する字幕用テキストの書き起こしおよび背景音などの付加情報データを効率的に入力することが可能となる。

【0080】

また、書き起こしそのものは、人の知識・能力に依存する手作業によらざるを得ないが、テレビ素材番組の映像、音声などを効果的に操作・処理して、この書き起こしなどの手作業を支援することが可能となる。

【0081】

書き起こしなどの作業を分析すると、通常テレビ素材番組を収録した V T R テープを適宜操作して映像、音声を再生し、音声の中のスピーチをテキストとして書き起こすとともに、その開始・終了時間も V T R のタイムコードを利用して記録しているが、その過程で該当スピーチ部分の頭出し、スピーチ部分の書き起こしに適した速度での再生や繰り返し再生、スピーチ部分終了点の確認などの複雑な V T R 操作が行われる。この点に関しては、話速変換技術を適用して書き起こし作業者の聞き取り能力に応じて話速を遅くしたり早くしたりできるので、書き起こし作業が容易にでき、また、タイミングを合わせも容易にすることができる。

【0082】

≪自動字幕番組データ制作部 4 0 の処理≫

図 6 は自動字幕番組データ制作部 4 0 において実行される表示字幕文へのタイミング付与の処理手順の一例として、音声認識手法に代わるタイミング検出手法の適用で高速化したフローチャートである。

【0083】

まず、音声データからポーズ開始タイミング、継続時間を求める。その適否を検証し必要な修正をする（図 6 のステップ S T 2 1）。なお、ポーズ時間の検出機能は、音声認識処理をすることなく、素材 V T R から供給される音声のレベルや継続時間、およびそのタイムコードから、例えばそのレベルが指定レベル以下で連続する開始、終了タイムコードを検出するものであり、ブロック・ケプストラム・フラックス法などの方法によって実行される。

10

20

30

40

50

【0084】

次に、比較的長いポーズ（例えば2秒以上）で字幕用テキストをブロックに分割し、ブロックテキスト文としてその開始、終了タイミングを付与する（ステップST22, 23）。

【0085】

次いで、ブロックテキスト文の継続時間をその総読み数（計算推定）で割り、当該範囲の平均読み速度を求める（ステップST24）。

【0086】

次いで、各ブロックテキスト文を、中の長さのポーズ箇所などを改行点とする表示単位字幕文に分割する（ステップST25）。この場合、予め定められた分割ルールを適用する 10

【0087】

次いで、各分割字幕文の文頭、文末に対応するタイミングを、ブロックテキスト文の開始、終了タイミングやポーズのタイミング、平均読み速度を基に計算し、付与する。この場合、文字数、文字種法、または発音数法を適用する（ステップST26）。

【0088】

そして、各表示単位字幕の表示時間をチェックし、必要ならば終了タイミングを修正する（ステップST27）。

【0089】

このようにして、自動処理により表示単位字幕文へのタイミング付与が実行され、自動字幕番組データが作成される。 20

【0090】

《字幕番組データ編集・試写部50の処理》

上述のようにして作成された字幕番組データは字幕番組データ編集・試写部50による編集・試写処理によって最終的な字幕番組として完成される。この機能における作業は、一応出来上がった字幕番組データを専門知識を有する作業者が試写し、必要ならば修正するものであり、作業者がチェックし易い試写形態、字幕内容、改行・改頁、タイミングなどに関する修正編集がし易いよう支援する機能が盛り込まれている。

【0091】

ここで、以下の説明では、送出／消去時間、字幕本文、および管理情報より構成された字幕番組データを字幕素材と仮称する。 30

【0092】

また、編集の補助として、MPEG画像ファイルを使用して、タイミングの調整や、字幕ポジションの調整を行うことができる。なお、字幕素材編集の入出力は、文字放送のNAB形式のファイルのみとなる。このファイルは、他の『ページ一覧』『ページ利用』『デジタル形式変換』の入出力形式になる。

【0093】

図7は、字幕番組編集・試写部50で使用されるメイン画面の構成を示している。このメイン画面は、字幕素材編集のために使用される。このメイン画面を使用して、字幕素材の編集や各補助機能の操作を行う。なお、画面のサイズは、1280×1024である。 40

【0094】

メイン画面501は大きく分けて、メニュー領域502と、編集領域503と、一覧領域504の3つの領域に分かれている。

【0095】

メニュー領域502は、各種機能の呼び出しを行う部分であり、メニューバーとツールバーに用意された項目を実行することにより、各種機能の呼び出しを行うことができる。

【0096】

編集領域503は、図8に示されるように、映像エリア503Aを備えている。この映像エリア503Aは、システム内に保存されているMPEG/AVIファイルを再生した映像を表示するエリアと、文字入力ソフトを使用して入力された字幕本文が表示されるエリ 50

アとから成っており、本文と画像の位置調整にも使用できる。通常の試写では、この編集領域503を使用して確認を行う。また、再生の停止や、IN TIME、OUT TIMEの取り込み、リピート再生などの操作を行うことが可能である。映像エリア503Aの画像のサイズは、640×480である。

【0097】

図7に示すタイムコード表示エリア503Bは、タイムコードを表示する部分であり、再生している映像の位置は29.97フレーム1秒として再生される。

【0098】

“再生”ボタン503Dは、映像の再生スタートに使用され、“停止”ボタン503Cは、映像の再生ストップに使用される。なお、“停止”ボタン503Cが押されても、映像がストップされるだけで映像そのものは画面上からは消えない。 10

【0099】

再生位置トラックバー503Eは、現在映像ファイルのどの部分を再生しているかを示すものである。このトラックバーを動かすことにより、再生位置を変更することが可能である。

【0100】

“MARK IN”ボタン503Fは、編集領域のIN TIMEの欄へ、現在の映像の時間をコピーするボタンである。“MARK OUT”ボタン503Gは、編集領域のOUT TIMEの欄へ、映像の現在の時間をコピーするボタンである。

【0101】

“繰り返し再生”ボタン503Hは、編集領域のIN TIMEとOUT TIMEの期間の画像を繰り返し作成するために使用される。 20

【0102】

なお、終了は、“停止”ボタン503Bか“MARK IN”ボタンを操作することにより成される。

【0103】

“<<5秒／5秒>>”ボタン503Jは、映像ファイルの再生位置を5秒分だけ進め／戻すためのボタンである。繰り返し再生中に実行された場合には、通常再生に切り替わる。

【0104】

前述したように、編集領域503は、字幕データの入力や、入力後の編集を行う他に映像セクションを兼ねていて、映像エリア503A上に重ねて試写を実行する際にも利用される。“前ページ”と“後ページ”は、現在中央に表示されている字幕データページの前と後のページの字幕を表示する領域である。 30

【0105】

この場合、編集領域503では、映像エリア503A上に表示されている枠カーソルを使用して、新たな文字の入力を行う。入力を行った文字の色サイズは、右上のリストボックスで設定した内容になる。

【0106】

また、一旦入力を行った文字を範囲指定し、削除／複写／移動や、属性（色、サイズ、囲み等）の変更を行うことができる。 40

【0107】

“IN TIME／OUT TIME”は、ページのIN TIME／OUT TIMEの設定を行う。“MARK IN”ボタン503F、“MARK OUT”ボタン503Gにより、映像ファイルの再生タイミングを“IN TIME／OUT TIME”として取り込むことができる。

【0108】

前景色／背景色／文字サイズエリア503Lは、編集領域503に、これから新規に入力する文字の前景色／背景色／文字サイズを設定する。

【0109】

メッシュエリア 503M は、画像の上に、文字の入力位置の目安とする基準線を表示するかを設定するものである。

【0110】

“ ページ追加 ” ボタン 503N は、編集中の素材の最後に、新たなページを追加し、そのページを編集対象として表示するために使用される。“ ページ挿入 ” ボタン 503P は、現在編集しているページの前か後に、新たにページを挿入し、そのページを編集対象として表示するために使用される。“ ページ削除 ” ボタン 503Q は、現在編集中のページを削除するために使用される。なお、ページ削除を実行する際には、確認操作が入る。

【0111】

“ 通し試写 ” ボタン 503R は、現在編集対象としているページの少し前より試写を開始し、“ 試写終了 ” ボタン 503S が実行されるまで、試写を行う際に使用される。 10

【0112】

“ 部分試写 ” ボタン 503T は現在対象としているページの前のページより試写を開始し、次のページの OUT TIME 後に試写を終了する。途中で“ 試写終了 ” ボタンにて終了することも可能である。

【0113】

“ 全画面試写 ” ボタンは、モニタ全てを使用して『通し試写』と同じ動作を行う。“ 試写終了 ” ボタンは、現在行っている試写を中止する。

【0114】

図 9 は、一覧領域 504 の構成を示している。 20

【0115】

一覧領域 504 は、画像とこの画像に対応する作成済みの字幕テキストを表示する領域であり、MPEG/AVI 映像ファイルより一定間隔毎若しくはカット変わりの画像一覧を表示する画像エリア 504A、作成されたテキストの内容である字幕本文を表示する字幕本文エリア 504B と、作成された字幕本文に対応するスピーチの音声に関連した波形を表示する波形エリア 504C と、波形エリア 504C に表示される波形の時間指標となるタイムラインが表示されるタイムラインエリア 504D とから成っている。この一覧領域 504 によって編集・試写作業者の作業支援が効果的に行われる。

【0116】

また、この一覧領域 504 における一画面で表示する幅は、“ ツール(T)” の“ オプション ” で選択された 120 秒、60 秒、30 秒のいずれかとなる。 30

【0117】

“ 画像 ” は、映像ファイルより、一定間隔毎若しくはカット変わりなどの画像を抜き出し、表示する。画像のサイズは 160×120、間隔は全体幅が 120 秒の時は 20 秒、60 秒の時は 10 秒、30 秒の時は 5 秒となる。

【0118】

“ 字幕本文 ” は、編集領域 503 によって入力された本文の内容が表示される。枠の幅は時間幅を表している。なお、幅が狭く、本文が表示しきれない場合には省略される。

【0119】

ページ番号エリア 504F の枠をドラッグで変更した場合には、対象一件の幅（時間幅）の変更を、本文の部分の枠をドラッグで変更した場合には、二件の境界線の変更を行う。 40

【0120】

枠内がダブルクリックされた際、編集領域 503 が編集途中でなければ、入力一覧で指定したページを編集対象のページ変更する。

【0121】

“ 波形 ” は、映像ファイルに記録されている音声（音楽、効果音を含む）信号を処理し波形化したものが表示される。なお、波形としては、スピーチ区間とポーズ区間とを明確にするために、矩形波状に波形整形して表示するようにしても良い。

【0122】

IN TIME/OUT TIME の編集単位は、映像フィルムからタイムコードを読み込 50

む際の最小単位を設定する。取り込みを実行した際、設定値未満の値は切り捨てされる。この場合の“設定”は1フレーム／5フレーム／10フレーム／15フレーム／1秒である。

【0123】

一覧表示の一画面の幅は、メイン画面の一覧領域にどれだけの時間分を表示するかを設定する。この場合の“設定”は、30秒／60秒／120秒で、それに合わせて画像も5秒間隔、10秒間隔、20秒間隔となる。

【0124】

このようにこの実施の形態によれば、研究を進めている自動字幕制作機能と、多彩な支援機能を盛り込んだ手動字幕制作機能とを効果的にタイアップした機能・構成を構築でき、多様なテレビ番組の字幕制作を、時間、人手、費用などの点で効率的にできるシステムを提供できる。そのため、字幕制作作業者の観点からの作業形態などに関するシステムフレキシビリティが向上する。

【0125】

【発明の効果】

以上説明したように本発明によれば、これまでに開発された自動字幕制作システムのシステム評価などから得られた知見をもとに、各自動化要素技術を高性能化した新しい自動字幕制作システムを中核にするとともに、字幕番組制作者による手動字幕制作を効率的に支援する実用性のより高い半自動型字幕番組制作システムを提供することが可能となる。

【0126】

また、本発明の半自動型字幕番組制作システムは、電子化原稿のない番組や背景音レベルの大きい番組など、多様な番組の字幕制作にも対応可能であり、今後適用番組分野、番組数などの拡大が見込まれる字幕放送の番組制作システムとして、大きな効果が期待できる。

【図面の簡単な説明】

【図1】本発明による半自動型字幕番組制作システムの実施形態を示すブロック図である。

【図2】本発明による半自動型字幕番組制作システムを構成する字幕テキスト書き起こし部で使用されるメイン画面の構成を示す説明図である。

【図3】図2に示すメイン画面の編集領域の構成を示す説明図である。

【図4】図2に示すメイン画面の一覧領域の構成を示す説明図である。

【図5】本発明による半自動型字幕番組制作システムを構成する字幕テキスト書き起こし部の処理手順を示すフローチャートである。

【図6】自動字幕番組データ制作部において実行される表示字幕文へのタイミング付与の処理手順の一例を示すフローチャートである。

【図7】本発明による半自動型字幕番組制作システムを構成する字幕番組編集・試写部で使用されるメイン画面の構成を示す説明図である。

【図8】図7に示すメイン画面の編集領域の構成を示す説明図である。

【図9】図7に示すメイン画面の一覧領域の構成を示す説明図である。

【符号の説明】

- 10 基本GUIシステム部
- 20 ディスク記録・再生部
- 30 字幕テキスト書き起こし部
- 31 ディスク記録再生制御機能
- 32 情報表示機能
- 33 データ作成制御機能
- 34 データ作成画面表示機能
- 35 主映像表示機能
- 40 自動字幕番組データ制作部
- 41 テキスト自動要約機能

10

20

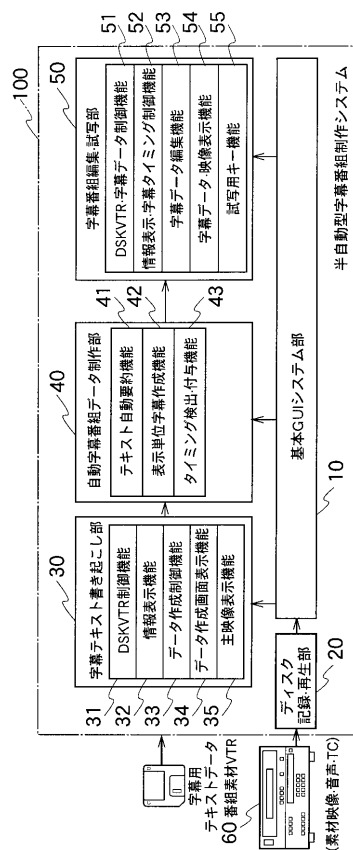
30

40

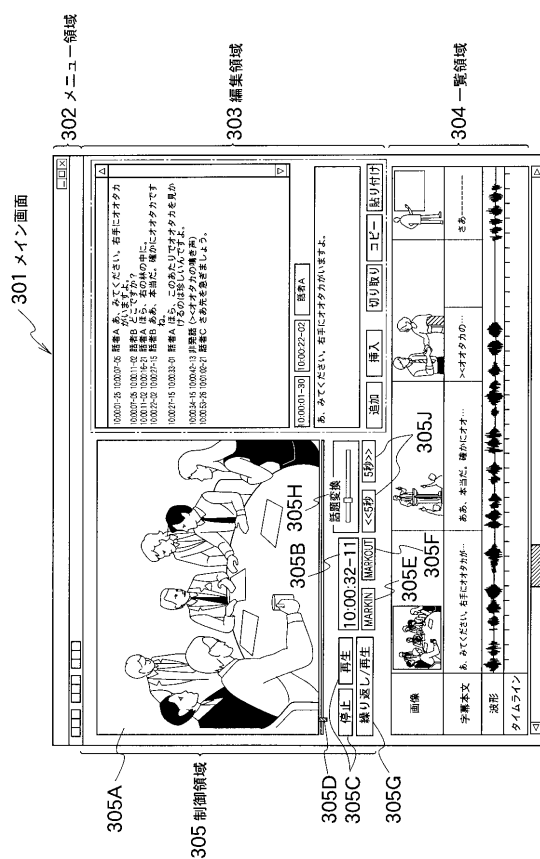
50

- | | |
|-------|--------------------|
| 4 2 | 表示単位字幕作成機能 |
| 4 3 | タイミング検出・付与機能 |
| 5 0 | 字幕番組編集・試写部 |
| 5 1 | ディスク記録再生・字幕データ制御機能 |
| 5 2 | 情報表示・字幕タイミング制御機能 |
| 5 3 | 字幕データ編集機能 |
| 5 4 | 字幕データ・映像表示機能 |
| 5 5 | 試写用キー機能 |
| 6 0 | 番組素材V T R |
| 1 0 0 | 半自動型字幕番組制作システム |

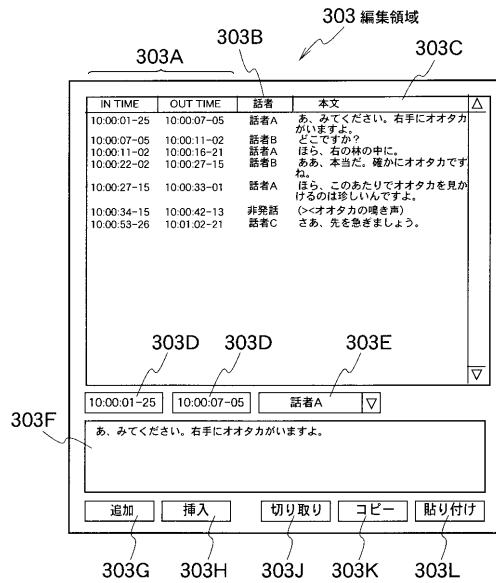
【图 1】



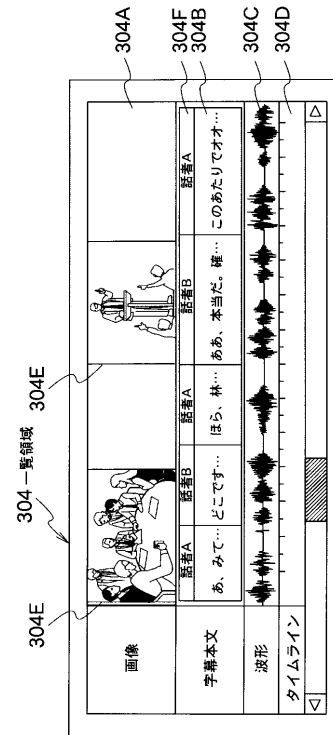
【图 2】



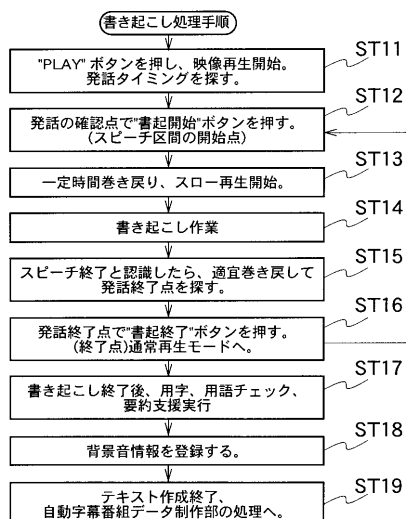
【図 3】



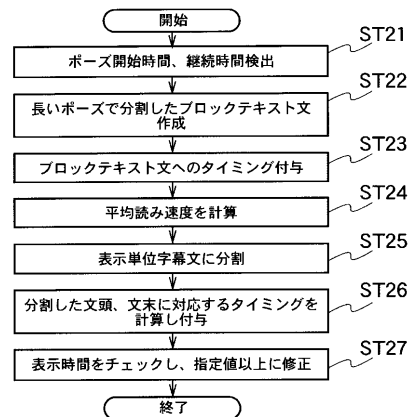
【図 4】



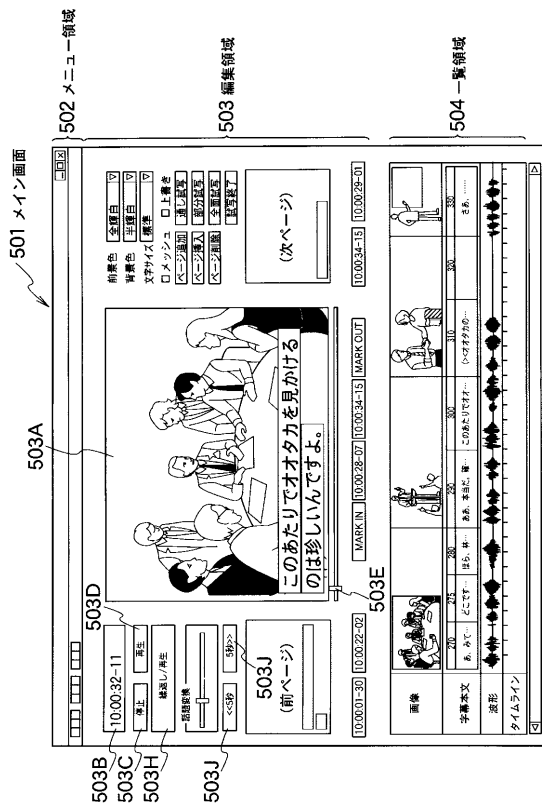
【図 5】



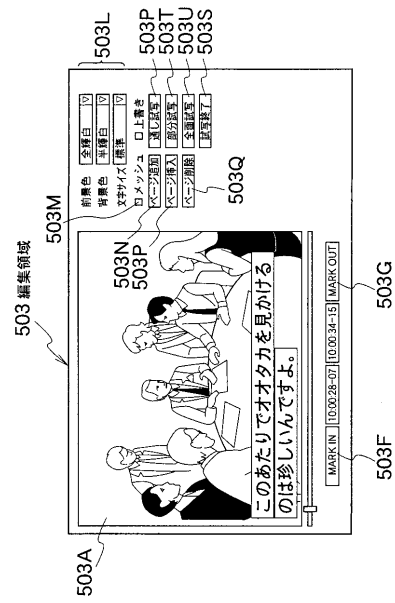
【図 6】



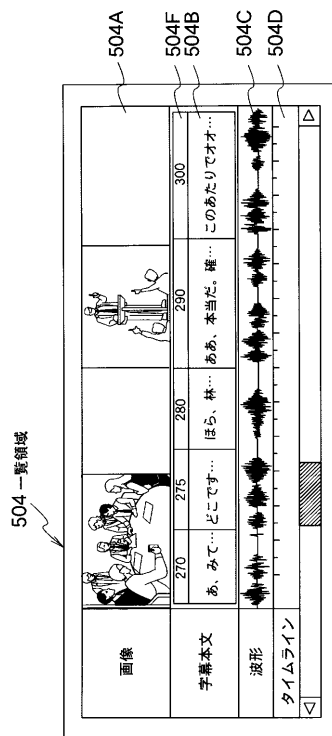
【図 7】



【図 8】



【図 9】



フロントページの続き

- (74)代理人 100100712
弁理士 岩▲崎▼ 幸邦
- (74)代理人 100101247
弁理士 高橋 俊一
- (74)代理人 100108914
弁理士 鈴木 壯兵衛
- (72)発明者 沢村 英治
東京都港区芝 2-3-1-19 通信・放送機構内
- (72)発明者 門馬 隆雄
東京都港区芝 2-3-1-19 通信・放送機構内
- (72)発明者 浦谷 則好
東京都港区芝 2-3-1-19 通信・放送機構内
- (72)発明者 大関 健二
東京都渋谷区宇田川町 37 番 18 号 財団法人 エヌエイチケイエンジニアリングサービス内
- (72)発明者 江原 暉将
東京都世田谷区砧一丁目 10 番 11 号 日本放送協会放送技術研究所内
- (72)発明者 白井 克彦
東京都港区芝 2-3-1-19 通信・放送機構内
- F ターム(参考) 5C023 AA18 AA34 AA38 BA11 CA01 CA05 EA03
5C053 FA14 FA23 GB11 GB12 JA01 JA16 KA22 LA04