

(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2016-119600

(P2016-119600A)

(43) 公開日 平成28年6月30日 (2016.6.30)

(51) Int.Cl.			F I	テーマコード (参考)		
HO4N	5/91	(2006.01)	HO4N	5/91	N	5C053
G11B	27/031	(2006.01)	G11B	27/031		5D077
G11B	27/34	(2006.01)	G11B	27/34	P	5D110

審査請求 未請求 請求項の数 10 O L (全 16 頁)

(21) 出願番号	特願2014-258921 (P2014-258921)	(71) 出願人	000000376 オリンパス株式会社 東京都八王子市石川町2951番地
(22) 出願日	平成26年12月22日 (2014.12.22)	(74) 代理人	100076233 弁理士 伊藤 進
		(74) 代理人	100101661 弁理士 長谷川 靖
		(74) 代理人	100135932 弁理士 篠浦 治
		(72) 発明者	木村 圭之介 東京都渋谷区幡ヶ谷2丁目43番2号 オリンパスイメージング株式会社内
		(72) 発明者	川口 勝久 東京都渋谷区幡ヶ谷2丁目43番2号 オリンパス株式会社内

最終頁に続く

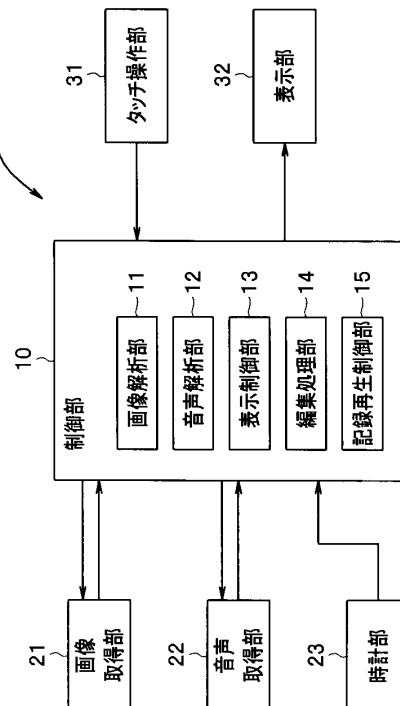
(54) 【発明の名称】 編集装置及び編集方法

(57) 【要約】

【課題】映像と音声を同一スケールの時間軸上に配置して表示することで、映像と音声の編集合成作業を容易にする。

【解決手段】 編集装置は、録画された映像を解析する画像解析部と、録音された音声を解析する音声解析部と、前記画像解析部の解析結果に基づいて前記録画された映像の再生時間に対応する映像バー表示を表示部に表示すると共に、前記音声解析部の解析結果に基づいて前記録音された音声の再生時間に対応する音声バー表示であって単位長さ当たりの時間が前記映像バー表示と同一の音声バー表示を前記表示部に表示する表示制御部とを具備する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

録画された映像を解析する画像解析部と、
録音された音声を解析する音声解析部と、
前記画像解析部の解析結果に基づいて前記録画された映像の再生時間に対応する映像バー表示を表示部に表示すると共に、前記音声解析部の解析結果に基づいて前記録音された音声の再生時間に対応する音声バー表示であって単位長さ当たりの時間が前記映像バー表示と同一の音声バー表示と前記表示部に表示する表示制御部と
を具備したことを特徴とする編集装置。

【請求項 2】

前記表示制御部は、前記画像解析部の解析結果に基づいて前記録画された映像の所定時間の映像部分に基づく映像イベント表示を前記映像バー表示の対応する再生位置近傍に表示すると共に、前記音声解析部の解析結果に基づいて前記録音された音声の所定時間の音声部分に基づく音声イベント表示を前記音声バー表示の対応する再生位置近傍に表示することを特徴とする請求項 1 に記載の編集装置。

【請求項 3】

前記表示制御部は、前記映像イベント表示を前記映像バー表示の時間軸方向に直交する方向に配置し、前記音声イベント表示を前記音声バー表示の時間軸方向に直交する方向に配置することを特徴とする請求項 1 又は 2 に記載の編集装置。

【請求項 4】

前記映像バー表示に対する操作を受け付けて前記映像バー表示に対応する時間の映像の編集ポイントを設定すると共に、前記音声バー表示に対する操作を受け付けて前記音声バー表示に対応する時間の音声の編集ポイントを設定する編集処理部
を具備したことを特徴とする請求項 1 又は 3 のいずれか 1 つに記載の編集装置。

【請求項 5】

前記映像バー表示及び前記映像イベント表示の少なくとも一方に対する操作を受け付けて前記映像バー表示に対応する時間の映像の編集ポイントを設定すると共に、前記音声バー表示及び前記音声イベント表示の少なくとも一方に対する操作を受け付けて前記音声バー表示に対応する時間の音声の編集ポイントを設定する編集処理部
を具備したことを特徴とする請求項 2 に記載の編集装置。

【請求項 6】

前記編集処理部は、前記映像バー表示に対する操作を受け付けて前記映像バー表示に対応する映像の時間軸を制御すると共に、前記音声バー表示に対する操作を受け付けて前記音声バー表示に対応する音声の時間軸を制御することを特徴とする請求項 4 に記載の編集装置。

【請求項 7】

前記編集処理部は、前記映像バー表示及び前記映像イベント表示の少なくとも一方に対する操作を受け付けて前記映像バー表示に対応する映像の時間軸を制御すると共に、前記音声バー表示及び前記音声イベント表示の少なくとも一方に対する操作を受け付けて前記音声バー表示に対応する音声の時間軸を制御することを特徴とする請求項 5 に記載の編集装置。

【請求項 8】

前記編集処理部は、前記音声イベント表示に対するタッチ操作、スライド操作又はピンチ操作を受け付けることを特徴とする請求項 4 乃至 7 のいずれか 1 つに記載の編集装置。

【請求項 9】

前記編集処理部は、前記映像の編集ポイントによって指定された映像期間の映像を前記音声の編集ポイントによって指定された音声期間に挿入するか又は、前記音声の編集ポイントによって指定された音声期間の音声を前記映像の編集ポイントによって指定された映

10

20

30

40

50

像期間に挿入する追加合成処理を行う

ことを特徴とする請求項 4 乃至 8 のいずれか 1 つに記載の編集装置。

【請求項 10】

録画された映像を解析し、

録音された音声を解析し、

前記画像解析部の解析結果に基づいて前記録画された映像の再生時間に対応する映像バー表示を表示部に表示すると共に、前記音声解析部の解析結果に基づいて前記録音された音声の再生時間に対応する音声バー表示であって単位長さ当たりの時間が前記映像バー表示と同一の音声バー表示を前記表示部に表示する

ことを特徴とする編集方法。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、映像と音声とを合成処理する編集装置及び編集方法に関する。

【背景技術】

【0002】

従来、映像コンテンツの製作では、映像と音声とを別々に記録して合成する手法が採用されることがある。先に撮影した映像に後から音声を合成するアフターレコーディングや、先に収録した音声に後から映像を合成するプリレコーディングが行われることがある。以下、アフターレコーディング及びプリレコーディングの処理を、映像又は音声の追加合成処理という。

20

【0003】

このような映像又は音声の追加合成処理を行う編集装置は、映像撮影時に映像に適した音声を収録することができない場合や音声収録時に音声に適した映像を撮影することができない場合であっても、映像に適した音声、音声に適した映像を合成することができる。また、編集装置は、音声や映像とは無関係な映像や音声を合成することもでき、映像コンテンツの価値を高めることも可能である。

【0004】

例えば、特許文献 1 においては、画像に応じた擬音や画像をイメージに合成する装置が開示されている。

30

【先行技術文献】

【特許文献】

【0005】

【特許文献 1】特開 2014 - 17852 号公報

【発明の概要】

【発明が解決しようとする課題】

【0006】

このように、従来、編集装置は、映像及び音声を別々に取得し、取得した映像及び音声を編集して追加合成することができる。この場合において、編集装置は、映像を撮影時間順に時間軸上に並べて配置することで、音声を合成するタイミングを視覚的に確認できるようになっている。例えば、特許文献 1 の装置では、動画像の一部の区間に対してインデックスを付加すると共に、動画像の時間変化をバー表示するようになっている。

40

【0007】

しかしながら、従来の編集装置においては、音声については、時間軸で視覚的に確認することができず、映像と音声とを所望のタイミングで追加合成する作業には熟練を要するという問題があった。

【0008】

本発明は、映像と音声を同一スケールの時間軸上に配置して表示することで、映像と音声の追加合成処理の作業を容易にすることができる編集装置及び編集方法を提供することを目的とする。

50

【課題を解決するための手段】

【0009】

本発明に係る編集装置は、録画された映像を解析する画像解析部と、録音された音声を解析する音声解析部と、前記画像解析部の解析結果に基づいて前記録画された映像の再生時間に対応する映像バー表示を表示部に表示すると共に、前記音声解析部の解析結果に基づいて前記録音された音声の再生時間に対応する音声バー表示であって単位長さ当たりの時間が前記映像バー表示と同一の音声バー表示を前記表示部に表示する表示制御部とを具備する。

【0010】

本発明に係る編集方法は、録画された映像を解析し、録音された音声を解析し、前記画像解析部の解析結果に基づいて前記録画された映像の再生時間に対応する映像バー表示を表示部に表示すると共に、前記音声解析部の解析結果に基づいて前記録音された音声の再生時間に対応する音声バー表示であって単位長さ当たりの時間が前記映像バー表示と同一の音声バー表示を前記表示部に表示する。

【発明の効果】

【0011】

本発明によれば、映像と音声を同一スケールの時間軸上に配置して表示することで、映像と音声の編集合成作業を容易にすることができるという効果を有する。

【図面の簡単な説明】

【0012】

【図1】本発明の第1の実施の形態に係る編集装置の回路構成を示すブロック図。

【図2】追加合成処理に用いる映像及び音声の撮像及び収録の様子を説明するための説明図。

【図3】カメラ制御を示すフローチャート。

【図4】アフターレコーディングにおける追加合成処理を説明するためのフローチャート。

【図5】追加合成処理画面を説明するための説明図。

【図6】追加合成処理画面を説明するための説明図。

【図7】変形例を示すフローチャート。

【図8】変形例における追加合成処理画面を説明するための説明図。

【図9】変形例における追加合成処理画面を説明するための説明図。

【発明を実施するための形態】

【0013】

以下、図面を参照して本発明の実施の形態について詳細に説明する。

【0014】

(第1の実施の形態)

図1は本発明の第1の実施の形態に係る編集装置の回路構成を示すブロック図である。本実施の形態は、追加合成処理において、映像と音声の時間軸のスケールを一致させて、単位長さ当たりの時間が同一の映像バー表示及び音声バー表示を表示させると共に、各バー表示上のタイミングに対応する映像や音声等のイベント表示を表示可能とすることにより、映像又は音声の追加合成処理の作業を容易にするようにしたものである。

【0015】

図1において編集装置1は、制御部10を有している。制御部10は、図示しないCPU等のプロセッサによって構成することができ、図示しないメモリに記憶されたプログラムに従って動作して所定の機能を実現するものであってもよい。

【0016】

画像取得部21は動画像を取得して制御部10に与える。例えば、画像取得部21を撮像部によって構成し、撮像部が動画撮影して取得した撮像画像(動画)を制御部10に出力するようにしてもよい。また、音声取得部22は音声を取得して制御部10に与える。例えば、音声取得部22をマイク等の収録部によって構成し、収録部が取得した音声を制

10

20

30

40

50

御部 10 に出力するようにしてもよい。また、時計部 23 は時間情報を発生して、発生した時間情報を制御部 10 に出力する。こうして、制御部 10 は、入力された時間情報によって、画像取得部 21 からの撮像画像をその撮像時間に対応させて取得することができるようになっている。また、制御部 10 は、入力された時間情報によって、音声取得部 22 からの音声をその收音時間に対応させて取得することができるようになっている。制御部 10 は、画像取得部 21 及び音声取得部 22 を制御することができるようになっている。

【0017】

制御部 10 中の画像解析部 11 は、入力された撮像画像に対する画像解析処理を実行して画像解析結果を得る。例えば、画像解析部 11 は、画像判定及びシーンチェンジ判定等を行って、シーンチェンジ及びシーンチェンジタイミングにおける画像等の解析結果を時間情報に対応させて映像イベント情報として出力する。また、音声解析部 12 は、入力された音声に対する解析処理を行って音声解析結果を得る。例えば、音声解析部 12 は、音量解析、周波数解析や音声認識処理等を行って音声の解析結果を時間情報に対応させて音声イベント情報として出力する。

なお、画像解析部 11 は、主被写体の位置の変化や、登場人物そのものや表情、ポーズの変化、輝度やフォーカス情報その他の変化によってシーンチェンジを判定する他、同時に記録された音声の周波数やレベルの変化に応じて、シーンを分類し、シーンチェンジ判定をしてもよい。また、音声解析部 12 は、近年研究が進んでいる周波数変化やレベル変化による人の声と雑音の分離技術などを利用して、各タイミングの音声特徴を分類したりしてもよい。

【0018】

表示制御部 13 は表示に関する各種処理を実行する。表示制御部 13 は、収集した動画を表示部 32 に与えて表示させることができる。表示部 32 は、LCD 等の表示画面を有しており、表示制御部 13 から与えられた画像を表示する。また、表示制御部 13 は、各種メニュー表示等を表示部 32 の表示画面に表示させることもできるようになっている。更に、本実施の形態においては、表示制御部 13 は、後述する編集処理部 14 に制御されて、追加合成処理の作業のための各種表示を表示部 32 に表示させることができるようになっている。

【0019】

表示部 32 の表示画面にはタッチ操作部 31 が設けられている。タッチ操作部 31 は、ユーザが指で指し示した表示画面上の位置に応じた操作信号を発生することができる。この操作信号は、制御部 10 に供給される。これにより、制御部 10 は、ユーザが表示画面上をタッチしたりスライドさせたりした場合には、ユーザのタッチ位置、指を近接させたり離間させる操作、スライド操作やスライド操作によって到達した位置、スライド方向、タッチしている期間等の各種操作を検出することができ、ユーザ操作に対応した処理を実行することができるようになっている。

【0020】

記録再生制御部 15 は、撮像画像及び收音音声の記録及び再生に関する処理を行うことができる。例えば、記録再生制御部 15 は、画像及び音声を図示しないメモリカード等の記録媒体に与えて記録させることができる。また、記録再生制御部 15 は、記録媒体に記録されている画像及び音声を読み出して再生することも可能である。

【0021】

編集処理部 14 には、記録再生制御部 15 から撮像画像及び音声時間が時間情報と共に与えられると共に、映像イベント情報及び音声イベント情報が与えられる。編集処理部 14 は、映像又は音声の追加合成処理モードが指定されると、表示制御部 13 に追加合成処理画面を表示させる。例えば、編集処理部 14 は、映像イベント情報に対応する映像イベント表示及び音声イベント情報に対応する音声イベント表示を、共通の時間軸上に対応させて配置した操作画面を追加合成処理画面として表示させることができる。例えば、編集処理部 14 は、単位長さ当たりの時間が相互に同一の 2 つのバー表示を表示させると共に、これらのバー表示上の位置（再生位置）に対応した画像再生時間及び音声再生時間の映像イ

10

20

30

40

50

ベント表示及び音声イベント表示を各バー表示の対応する時間の位置（再生位置）近傍に配置した表示を表示させることができる。なお、画像再生時間及び音声再生時間は、録画した映像及び録音した音声をリアルタイムで再生する場合には、それぞれ録画時間及び收音時間と同一である。

【 0 0 2 2 】

編集処理部 1 4 は、タッチ操作部 3 1 の操作によって映像や音声を選択されると、選択された映像を対応する時間の音声に挿入し、選択された音声を対応する時間の映像に挿入する追加合成処理を行う。また、編集処理部 1 4 は、挿入した映像又は音声の時間軸を調整する。例えば、編集処理部 1 4 は、挿入した映像又は音声の一方の時間軸を調整してスロー又は高速再生可能にする。なお、編集処理部 1 4 は、映像効果又は音声効果を向上させるために、映像又は音声の時間軸を調整するようにしてもよい。

10

【 0 0 2 3 】

なお、本実施の形態は撮像部及び收音部を備えた例えばカメラ等によって構成することが可能である。また、本実施の形態は、制御部 1 0 に撮像時間の情報を含む撮像画像と收音時間の情報を含む音声とを供給することができれば、撮像部及び收音部を備えている必要はなく、カメラの外にコンピュータやタブレット P C や携帯電話や表示装置等によって構成することも可能である。また、これらの機器において、映像及び音声を取り込む場合に、既に時間情報が映像及び音声データ中に含まれている場合には、時計部 2 3 は不要であるが、時間情報が含まれていない場合には、時計部 2 3 の時間情報を利用して画像、音声の実時間を計測しながら取り込むようにする。

20

【 0 0 2 4 】

次に、このように構成された実施の形態の動作について図 2 乃至図 6 を参照して説明する。図 2 は追加合成処理に用いる映像及び音声の撮像及び收音の様子を説明するための説明図であり、図 2 (a) は撮像の様子を示し、図 2 (b) は收音の様子を示している。図 3 はカメラ制御を示すフローチャートであり、図 4 はアフターレコーディングにおける追加合成処理を説明するためのフローチャートである。また、図 5 及び図 6 は追加合成処理画面を説明するための説明図である。

【 0 0 2 5 】

図 2 は編集装置 1 をカメラによって構成した例を示している。図 2 (a) は人物 4 1 が徒競走中の被写体 4 5 を撮影する様子を示している。人物 4 1 は図 1 の編集装置 1 が内蔵された筐体 1 a を把持して、画像取得部 2 1 を構成する撮像部によって人物 4 6 を含む被写体 4 5 を撮影する。また、図 2 (b) は人物 4 1 が人物 4 6 の会話を録音する様子を示している。人物 4 1 は、筐体 1 a を把持して、音声取得部 2 2 を構成する收音部によって人物の会話を録音する。なお、これらの録画及び録音は、編集装置 1 以外の装置によって行ってもよい。

30

【 0 0 2 6 】

映像又は音声の追加合成処理は、図 2 (a) , (b) のように、相互に異なるタイミングや場所で録画又は録音された映像及び音声を合成するものである。編集装置 1 がカメラ等によって構成されている場合には、撮影直後や再生時等においても追加合成処理が可能である。

40

【 0 0 2 7 】

図 3 は編集装置 1 がカメラによって構成されている場合のカメラ制御を示している。図 3 のステップ S 1 において、制御部 1 0 は、記録モードであるか否かを判定する。いま、記録モードが指定されているものとする。この場合には、制御部 1 0 は、ステップ S 2 において撮影モードであるか否かを判定する。いま、撮影モードが指示されているものとする。この場合には、制御部 1 0 は、ステップ S 3 において、画像取得部 2 1 を制御して撮像を開始させ、画像取得部 2 1 からの撮影画像を取り込む。制御部 1 0 は、取り込んだ撮像画像（動画）を表示部 3 2 に与えてスルー画表示させる（ステップ S 3 ）。

【 0 0 2 8 】

50

次に、制御部 10 は、ステップ S 4 において動画記録の開始が指示されたか否かを判定する。画像取得部 21 に動画記録を開始させるためのユーザ操作が行われると、制御部 10 は、ステップ S 5 において画像取得部 21 からの動画及び音声取得部 22 からの音声を取り込んで、記録再生制御部 15 によって記録を行う。

【0029】

次のステップ S 6 では、制御部 10 は、動画記録の終了操作が行われたか否かを判定する。制御部 10 は、終了操作が行われるまで、録画及び録音を継続し（ステップ S 5）、終了操作が行われると、ステップ S 7 において記録されている動画及び音声をファイル化する。なお、ステップ S 5 では、録画及び録音が同時に行われるものとして説明したが、録画のみ又は録音のみを行うようにしてもよい。

10

【0030】

次のステップ S 8 において、制御部 10 は、追加録音が指示されているか否かを判定する。制御部 10 は、追加録音が指示されていない場合には、ステップ S 9 において電源オフ操作があるか否かを判定し、電源オフ操作が合った場合には処理を終了し、電源オフ操作がない場合には処理をステップ S 1 に戻す。

【0031】

いま、ステップ S 8 において、ユーザによる追加録音の指示操作が検出されるものとする。この場合には、制御部 10 は、ステップ S 11 に処理を移行して、録音を開始する。制御部 10 は、音声取得部 22 が収録した音声を時間情報と共に取り込む（ステップ S 12）。編集処理部 14 は、この取込みに際して、ユーザによるタグ（目印）付け操作があった場合には、記録再生制御部 15 を制御して、音声にタグ付けを行った後記録させる（ステップ S 13）。

20

【0032】

次のステップ S 14 では、制御部 10 は、音声記録の終了操作が行われたか否かを判定する。制御部 10 は、終了操作が行われるまで、録音及びタグ付けを継続し（ステップ S 12, S 13）、終了操作が行われると、ステップ S 15 において記録されている音声をファイル化する。なお、図 3 ではステップ S 11 ~ S 15 において、音声の追加録音を行う例を示したが、ユーザの追加撮影操作によって、動画を追加撮影するようにしてもよい。

【0033】

制御部 10 はステップ S 16 において、アフターレコーディング又はプリレコーディングの追加合成処理が指示されたか否かを判定する。制御部 10 は、追加合成処理の指示がない場合には処理をステップ S 9 に移行し、追加合成処理の指示があった場合には処理をステップ S 25 に移行する。ステップ S 25 における追加合成処理については後述する。

30

【0034】

制御部 10 は、ステップ S 1 において記録モードでないものと判定すると、処理をステップ S 21 以降の再生モードに移行する。記録再生制御部 15 は、既に記録媒体に記録されている画像の情報を取得して表示制御部 13 に与える。表示制御部 13 は、記録済み画像の一覧であるファイル一覧表示を表示部 32 の表示画面に表示させる（ステップ S 21）。制御部 10 は、次のステップ S 22 において、画像の再生を選択する操作が行われたか否かを判定する。記録再生制御部 15 は、画像の再生が選択されると、選択された画像を再生して表示制御部 13 に与え、表示制御部 13 は再生された画像を表示部 32 の表示画面に表示する（ステップ S 23）。

40

【0035】

本実施の形態においては、制御部 10 は、ステップ S 22 において、再生が指示されない場合には、ステップ S 24 において追加合成処理が指示されているか否かを判定する。追加合成処理が指示されていない場合には、編集処理部 14 は、ユーザ操作に基づいて通常の編集処理を実行する（ステップ S 26）。追加合成処理が指示されると、編集処理部 14 は追加合成処理を実行する（ステップ S 25）。

【0036】

50

次に、図4乃至図6を参照して追加合成処理について説明する。図5及び図6は筐体1aの一面に表示部32の表示画面32aが配置されている例を示している。

【0037】

追加合成処理では、選択された映像に対して選択した音声を追加合成するか、又は、選択された音声に対して選択した映像を追加合成する。図4のステップS51～S59は動画選択時の動作を示し、ステップS61～S69は音声選択時の動作を示している。上述した図3のステップS21～S23においては画像の再生処理のみを説明したが、ステップS21のファイル一覧表示において音声ファイルを選択可能にすることも可能である。

【0038】

図5(a)はステップS21におけるファイル一覧表示の一例を示している。図5(a)においては、表示画面32a上には、画像ファイルを選択するためのサムネイル表示51と音声ファイルを選択するための選択アイコン表示52が配置されている。なお、画像ファイルと音声ファイルとを区別するためのマーク等を表示させるようにしてもよい。また、表示画面32a上には、削除等の通常の編集を行うための編集ボタン表示53及び追加合成処理のために「アフレコ編集」と表示された追加合成ボタン表示54と、戻るボタン表示55とが表示されている。

10

【0039】

ユーザが編集ボタン表示53上をタッチ操作すると、図3のステップS24からステップS26に処理が移行して通常の編集処理が行われる。いま、ユーザが所定のサムネイル表示51上をタッチして画像ファイルを選択した後、追加合成ボタン表示54上をタッチするものとする。そうすると、図3のステップS25に移行して追加合成処理が開始される。

20

【0040】

追加合成処理においては、図4のステップS31において、動画と音声の選択が終了したか否かが判定される。動画及び音声の両方の選択が終了していない場合には、ステップS51において動画の画像ファイルが選択されたか否かが判定される。動画が選択されると、ステップS52において選択された動画の内容が解析される。即ち、画像解析部11は、選択された動画の画像解析を行って、解析結果を編集処理部14に与える。編集処理部14は動画の画像の変化を時系列に整理し(ステップS53)、この結果を表示制御部13に与える。表示制御部13は、これらの結果に基づいて、画面上段に映像イベント表示である代表画像、対応するバー表示及び時間表示を表示する(ステップS54～S56)。なお、ユーザが図5(a)の戻るボタン表示55上をタッチすると、ステップS58に処理が移行して動画選択が解除される(ステップS59)。

30

【0041】

更に、表示制御部13は、選択された画像ファイルに対して追加合成する音声ファイルを選択するための表示も表示させる(ステップS57)。図5(b)はこのような音声ファイルの選択画面を示している。

【0042】

本実施の形態においては、図5(b)に示すように、表示画面32aの上段にステップS55において映像バー表示61が表示されている。映像バー表示61は選択された画像ファイルの時間に対応する長さを有しており、画像ファイルの時間軸を示すものである。表示画面32a上には、選択された画像ファイルの時間が3分10秒であることを示す映像時間表示63も表示されている(ステップS56)。

40

【0043】

映像バー表示61は、画像ファイル中の各フレーム又は複数のフレームの単位毎に区切りを有しており、表示制御部13は、編集処理部14を介して画像解析部11からシーンチェンジの判定結果が与えられて、シーンチェンジのフレームに対応する映像バー表示61の部分に、図5(b)において塗り潰して示しているように、他の部分とは異なる表示64を表示する。更に、表示制御部13は、シーンチェンジ位置のフレームのサムネイル表示62を、代表画像表示として当該フレームの時間に対応する映像バー表示61の近傍

50

に表示させる（ステップS54）。なお、代表画像表示としては、シーンチェンジ位置のフレームに限らず、例えば、所定の時間間隔毎のフレームのサムネイル表示を用いてもよい。これらの表示によって、選択された映像の時間軸上における変化の様子等を把握しやすくなる。

【0044】

また、表示画面32aの下段には、音声ファイルを選択するための複数の選択アイコン表示65が配置されており（ステップS57）、各選択アイコン表示65の下方には対応する音声ファイルの時間を示す音声時間表示66も表示されている。例えば、「音声2」にて示す音声ファイルは、再生時間が2分05秒の音声を含むものであることが分かる。

【0045】

ここで、ユーザが図5（b）の選択アイコン表示65の1つをタッチして選択するものとする。そうすると、図4のステップS31、S51、S61から処理はステップS62に移行して、音声に対する解析処理が実行される（ステップS62）。即ち、音声解析部12は、選択された音声の解析を行って、解析結果を編集処理部14に与える。編集処理部14は音声の変化を時系列に整理し（ステップS63）、この結果を表示制御部13に与える。表示制御部13は、これらの結果に基づいて、画面下段に音声イベント表示である例えばキーワード表示等の代表音声表示、対応するバー表示及び時間表示を表示する（ステップS64～S66）。図5（c）はこの場合の画面表示を示している。

【0046】

本実施の形態においては、図5（c）に示すように、表示画面32aの下段にステップS65において音声バー表示71が表示されている。音声バー表示71は選択された音声ファイルの時間に対応する長さを有しており、音声ファイルの時間軸を示すものである。本実施の形態においては、映像バー表示61と音声バー表示71とは、単位時間当たりの長さが同一に設定されている。従って、映像バー表示61及び音声バー表示71によって、選択された映像と音声の時間軸上の関係が明瞭となる。

【0047】

音声バー表示71は、音声ファイル中の音声の変わり目や所定時間毎やキーワードの検出毎等に区切りを有しており、表示制御部13は、編集処理部14を介して音声解析部12から区切りの判定結果が与えられて、区切りの直後等における音声バー表示71の部分に、図5（c）において塗り潰して示しているように、他の部分とは異なる表示72を表示する。更に、表示制御部13は、区切り位置の音声の認識結果をテキストで示すテキスト表示73を代表音声表示として、当該区切りの時間に対応する音声バー表示71の近傍に表示させる（ステップS64）。なお、代表音声表示としては、音声の区切り位置に限らず、例えば、所定の時間間隔毎の音声認識結果のテキスト表示を用いてもよい。これらの表示によって、選択された音声の時間軸上における変化の様子等を把握しやすくなる。なお、ユーザが図5（c）の戻るボタン表示76上をタッチすると、ステップS68に処理が移行して音声選択が解除される（ステップS69）。

【0048】

表示制御部13は、映像（動画）及び音声の選択が終了すると、図4のステップS31からステップS32に処理を移行して、編集スイッチ表示75を表示する。ここで、ユーザが図5（c）に示すように、指79によって編集スイッチ表示75をタッチ操作するものとする。図5（c）は太枠によって編集スイッチ表示75が操作されたことを示す。これにより、表示制御部13は、表示部32の表示画面32a上に図5（d）に示す編集画面を表示する。

【0049】

図5（d）に示す編集画面は、挿入ボタン表示81、選択ボタン表示82及び削除ボタン表示83が表示されていると共に、「バー上タッチで選択」という説明表示77が表示されている。ユーザは、タッチ操作部31の映像バー表示61、音声バー表示71、サムネイル表示62、テキスト表示73上の位置をタッチ操作、ピンチ操作及びスライド操作することで、位置（編集ポイント）設定、範囲設定及び調整作業を行うことができる。ま

10

20

30

40

50

た、ユーザは、挿入ボタン表示 8 1、選択ボタン表示 8 2 及び削除ボタン表示 8 3 上の位置をタッチ操作することで、選択範囲の確定、削除範囲の確定及び選択範囲の挿入を行うことができる。

【 0 0 5 0 】

編集処理部 1 4 は、図 4 のステップ S 3 4 においてタッチ操作を検出し、ステップ S 3 6 においてピンチ操作を検出し、ステップ S 3 8 3 上の位置においてスライド操作を検出する。編集処理部 1 4 は、タッチ操作を検出するとタッチ位置に対応する時間を記憶し、ピンチ操作を検出するとピンチ量に対応する時間を記憶し、スライド操作を検出するとスライド量に対応する時間を記憶する。

【 0 0 5 1 】

例えば、ユーザが図 5 (d) の音声バー表示 7 1 上の任意の 2 箇所的位置をタッチすることで、編集処理部 1 4 は、記録されている音声のうち当該 2 箇所位置 (編集ポイント) 相互間に対応する時間の音声部分を指定する操作が行われたものと判定する。図 6 (a) はユーザが指 7 9 によって音声バー表示 7 1 上の 2 箇所位置をタッチしている様子を示している。この状態でユーザが図 5 (d) に示すように、選択ボタン表示 8 2 上をタッチすると、編集処理部 1 4 は、ステップ S 4 1 からステップ S 4 2 に処理を移行して、指定された音声部分の範囲を確定する。即ち、これらの操作によって、映像に挿入する音声の部分を確定する処理が行われる。

【 0 0 5 2 】

また、ユーザが図 6 (a) の太枠で示す削除ボタン表示 8 3 上をタッチ操作すると、編集処理部 1 4 は、ステップ S 4 5 からステップ S 4 6 に処理を移行して、指定された音声部分の範囲を削除する。なお、ユーザは、範囲指定の少なくとも一方についてはテキスト表示 7 3 上をタッチすることで、映像に挿入する音声部分の範囲を指定してもよい。

【 0 0 5 3 】

このように、音声バー表示 7 1 やテキスト表示 7 3 上のタッチ操作及び選択ボタン表示 8 2 や削除ボタン表示 8 3 に対する操作によって、ユーザは極めて簡単に合成処理する音声部分を確定することができる。この場合には、音声バー表示 7 1 の時間軸と映像バー表示 6 1 とで単位長さ当たりの時間が一致していることから、また、テキスト表示 7 3 によって、範囲指定しようとする音声部分がどのような音声部分であるかを比較的認識しやすいことから、ユーザは音声合成において選択すべき音声部分を容易に決定することができる。

【 0 0 5 4 】

挿入する音声部分が確定すると、表示制御部 1 3 は、図 6 (b) に示すように、確定した部分のみの音声バー表示 8 5 を表示する。また、表示制御部 1 3 は、確定した音声部分の時間を示す確定音声時間表示 8 6 も表示させる。図 6 (b) の例では、5 6 秒間の音声部分が確定していることを示している。

【 0 0 5 5 】

ユーザは、合成処理する音声部分を確定させると、次に、確定した音声部分を映像のいずれの時間の部分に挿入するかを設定する。この場合には、ユーザは映像バー表示 6 1 上又はサムネイル表示 6 2 上をタッチする。編集処理部 1 4 は、ユーザのタッチ位置に対応する時刻が音声挿入する時間の先頭位置であるものとして、当該時刻を挿入位置 (挿入編集ポイント) として記憶する。この状態でユーザが図 6 (b) の太枠に示すように、挿入ボタン表示 8 1 上をタッチすると、編集処理部 1 4 は、ステップ S 4 3 からステップ S 4 4 に処理を移行して、既に確定されている音声部分の範囲を、画像ファイルに対応する音声の指定された挿入編集ポイントに挿入する。こうして、映像に対する音声の追加合成処理が行われる。

【 0 0 5 6 】

映像に対する音声の追加合成が行われると、表示制御部 1 3 は、図 6 (c) に示す微調整画面を表示する。図 6 (c) に示すように、微調整画面では、「微調整」という表示によって微調整画面であることが示されると共に、確定ボタン表示 8 7 が表示される。ユー

10

20

30

40

50

ザは、映像バー表示 6 1 又は音声バー表示 8 5 に対してスライド操作することにより、映像に対する音声の挿入位置（時間）を変更することができる。図 6（c）の矢印 8 9 はユーザが音声バー表示 8 5 をスライド操作することを示している。このスライド操作が行われると、編集処理部 1 4 は、ステップ S 3 8 からステップ S 3 9 に移行して、音声の映像に対する挿入位置の時間をスライド量に応じて調整する。

【0057】

また、図 6（d）の矢印 9 0 はユーザが指 7 9 によって映像バー表示 6 1 上でピンチ操作することを示している。このピンチ操作が行われると、編集処理部 1 4 は、ステップ S 3 6 からステップ S 3 7 に移行して、音声に対する映像の時間軸をピンチ量に応じて調整する。即ち、ピンチ量に応じて映像の時間軸が伸張又は短縮する。従って、この場合には、ピンチ操作が行われた時間部分が高速又は低速再生されることになる。これにより、選択した音声の時間と挿入したい映像の期間とが異なる場合でも、選択した音声部分を映像の希望する再生位置に挿入することができる。この場合には、映像と音声の再生速度が異なることになり、例えば、映像は高速再生されている状態で音声は通常速度で再生されたりする。

10

【0058】

なお、表示制御部 1 3 は、伸張又は短縮後の時間を表示する。このピンチ操作は、映像バー表示 6 1 及び音声バー表示 8 5 のいずれに対しても行うことができ、映像及び音声の時間軸を適宜変更することが可能である。更に、これらのスライド操作及びピンチ操作は、音声を映像に挿入する前に行うことも可能である。

20

【0059】

ユーザが確定ボタン表示 8 7 上をタッチ操作すると、編集処理部 1 4 は、ステップ S 4 7 からステップ S 4 8 に処理を移行して、ユーザの操作に基づく追加合成処理を確定し、映像に音声を合成した情報をファイル化及びコンテナ化する。こうして、ユーザが選択した映像に対して、ユーザが選択した音声を追加合成するアフターレコーディング処理が可能である。制御部 1 0 は、次のステップ S 4 9 において追加合成処理の終了が指示されたか否かを判定し、終了する場合には処理を図 3 のステップ S 9 に戻す。なお、図 5 及び図 6 では、映像を先に選択して、後から音声を選択して映像に合成する例を示したが、音声を先に選択して後から映像を選択して音声に合成することもできることは明らかである。

30

【0060】

このように本実施の形態においては、追加合成処理においては、映像及び音声の時間軸を示す映像バー表示及び音声バー表示を表示させる。そして、これらの映像バー表示及び音声バー表示に対する操作によって、挿入する範囲や挿入位置等を指定するようになっており、ユーザは直感的な操作によって追加合成処理が可能である。しかも、これらの映像バー表示及び音声バー表示は、単位時間当たりの長さが同一であり、ユーザは、映像又は音声のどの位置に音声又は映像のどの部分を挿入すれば良いかを、容易に判断することができる。更に、映像バー表示及び音声バー表示の所定の時間の位置、例えばシーンチェンジや音の区切りの位置には、当該時間における映像のサムネイル表示や音声のテキスト表示を表示させており、ユーザは挿入位置及び挿入する部分を容易に判断することができる。更に、本実施の形態においては、映像及び音声の時間軸を、映像バー表示及び音声バー表示に対する簡単な操作で伸張又は短縮することができ、高速再生や低速再生等の特殊再生を容易に実現することもできる。このように、映像及び音声について、共通のスケールの時間軸を表示させることで、ユーザは音声部分の範囲や映像の挿入位置等を容易に判断でき、また、時間軸の調整も直感的に行うことができる。

40

【0061】

なお、図 4 の例では、追加合成処理において動画及び音声を選択された場合に画像及び音声の解析処理を行ったが、画像及び音声の記録又は取込み時にこれらの解析を行うようにしてもよい。

【0062】

（変形例）

50

図7は変形例を示すフローチャートである。図7において図4と同一の手順については同一符号を付して説明を省略する。なお、図7は図4のステップS51～S59及びステップS61～S69については図示を省略している。また、図8及び図9は変形例における追加合成処理画面を説明するための説明図である。図8及び図9において図5及び図6と同一物には同一符号を付して説明を省略する。

【0063】

上述した図4の例では、音声イベント表示である代表音声表示として、テキスト表示を表示する例を示した。本変形例は、代表音声表示として音声波形を表示するものである。図8(a)～(d)はそれぞれ図5(a)～(d)に対応する。図8(c)に示すように、本変形例においても、表示画面32aの下段に音声バー表示71が表示されている。音声バー表示71と映像バー表示61との時間軸は同一スケールであり、映像と音声の時間軸上の関係が明瞭である。

10

【0064】

本変形例の音声バー表示71では、図3のステップS13におけるタグ付け処理によって付加されたタグの位置を示す表示72(塗り潰し部)が表示されている。また、表示制御部13は、タグ位置に対応する音声バー表示71の近傍に、タグの番号を示すタグ番号表示91を表示させる。

【0065】

ここで、ユーザが編集スイッチ表示75をタッチ操作するものとする。図8(c)は大枠によって編集スイッチ表示75が操作されたことを示す。これにより、表示制御部13は、表示部32の表示画面32a上に図8(d)に示す編集画面を表示する。更に、図8(d)に示す編集画面において、ユーザが音声バー表示71上をタッチ操作するものとする。そうすると、表示制御部13は、ステップS34からステップS71に処理を移行して、タッチ位置が波形表示上であるか否かを判定する。この場合には、タッチ位置は音声バー表示71上であるので、表示制御部13は次のステップS73に移行して、編集処理部14から音声解析部12の解析結果である音声波形を取得して、当該波形を示す波形表示92を表示する(図7のステップS71)。なお、波形表示92は、ユーザによる音声バー表示71上のタッチ位置近傍の所定期間の波形を示すものである。また、タグ番号表示91上のタッチ操作によっても、同様の波形表示92を表示させることができるようになっている。

20

30

【0066】

この状態で、ユーザが波形表示92に対してピンチ操作を行うものとする。図9(a)の矢印93はユーザが指79によって波形表示上でピンチ操作することを示している。このピンチ操作が行われると、編集処理部14は、ステップS36からステップS81に移行してピンチ操作が波形表示上であるか否かを判定する。この場合には、ピンチ操作は波形表示上であるので、表示制御部13は、処理をステップS82に移行して、波形表示92をピンチ量に応じて拡大(又は縮小)する。図9(a)は拡大後の波形表示94を示している。即ち、波形表示94は波形表示92の期間の波形を時間軸方向(水平方向)に拡大して、波形の確認を容易にしたものである。

【0067】

40

波形表示94は拡大されているので、ユーザは編集ポイントを指定し易くなる。この状態で、ユーザが、図9(b)に示すように、波形表示94上の任意の位置をタッチ操作するものとする。そうすると、表示制御部13は、ステップS34、S71を介してステップS72に処理を移行して、タッチ位置に対応する時間を記憶する。表示制御部13は、このタッチ位置を示すポイント表示を波形表示上に表示する。

【0068】

図9(c)は同様の手法によって、もう一箇所の編集ポイントを指定し、選択ボタン表示82上のタッチ操作によって範囲を確定した状態を示している。表示画面32a上には、1点目の波形表示95aと2点目の波形表示95bとが表示されており、各波形表示95a, 95b上には、ユーザが編集ポイントとして指定した位置を示すポイント表示96

50

a, 96bが表示されている。また、音声バー表示71は、ポイント表示96a, 96b相互間に対応する音声期間が選択されていることを示す表示97が表示されている。

【0069】

こうして、この変形例では、代表音声表示として波形を表示し、この波形表示を見ながらユーザに編集ポイントを指定させることができる。音声として音楽等を追加合成する場合には、波形表示の方が編集ポイントを設定し易く、簡単に追加合成する音声部分を指定することができる。

【0070】

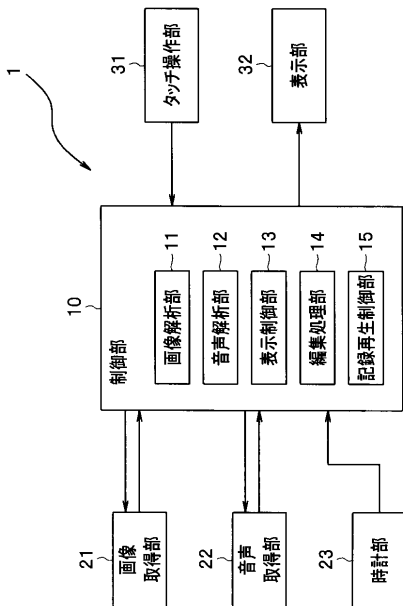
このように本変形例においては、波形表示を見ながら編集ポイントを設定することができ、操作性に優れている。

【符号の説明】

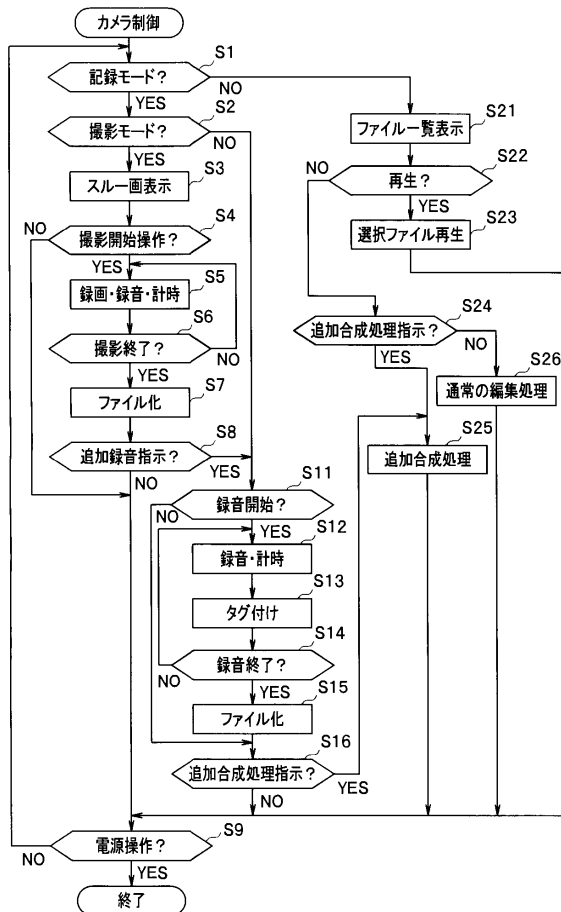
【0071】

1...編集装置、10...制御部、11...画像解析部、12...音声解析部、13...表示制御部、14...編集処理部、15...記録再生制御部、21...画像取得部、22...音声取得部、23...時計部、31...タッチ操作部、32...表示部。

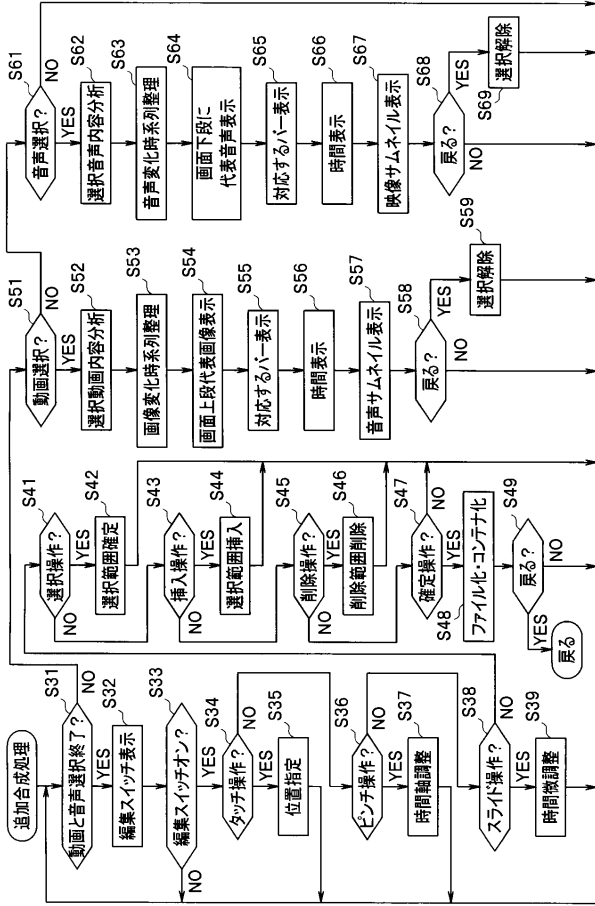
【図1】



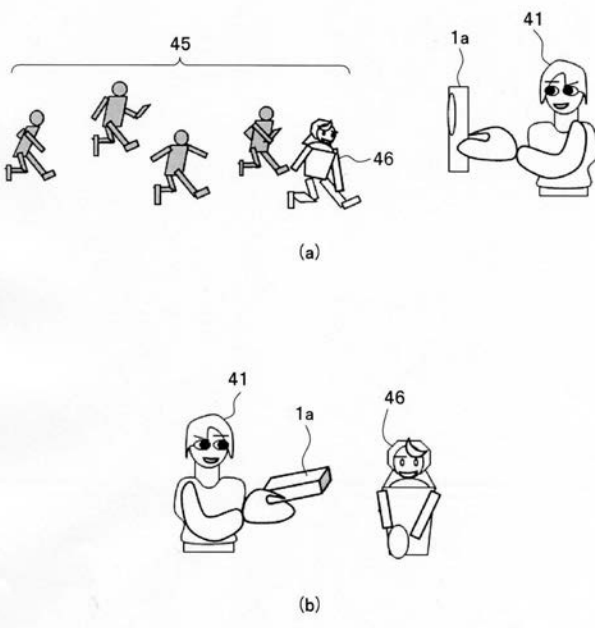
【図3】



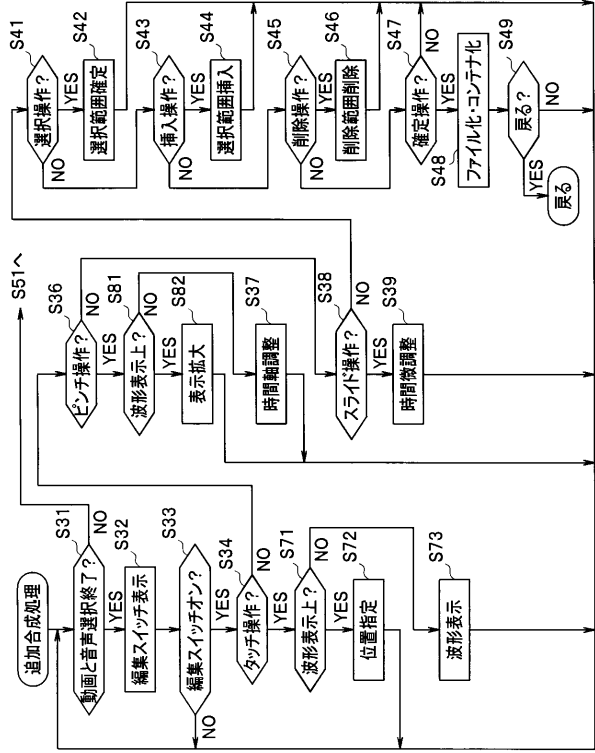
【 図 4 】



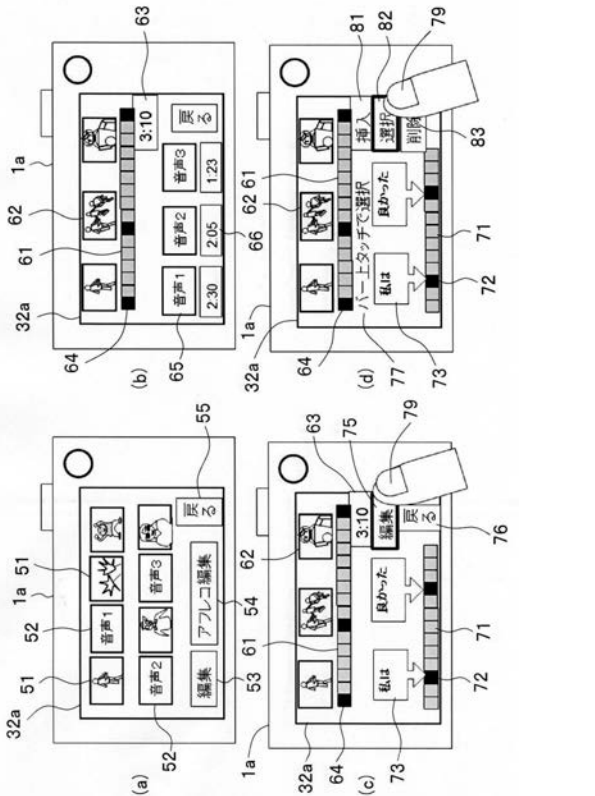
【 図 2 】



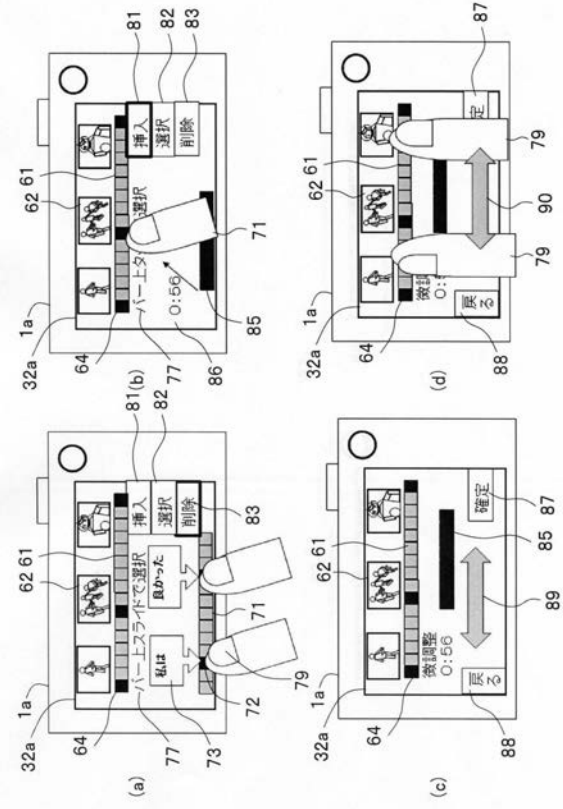
【 図 7 】



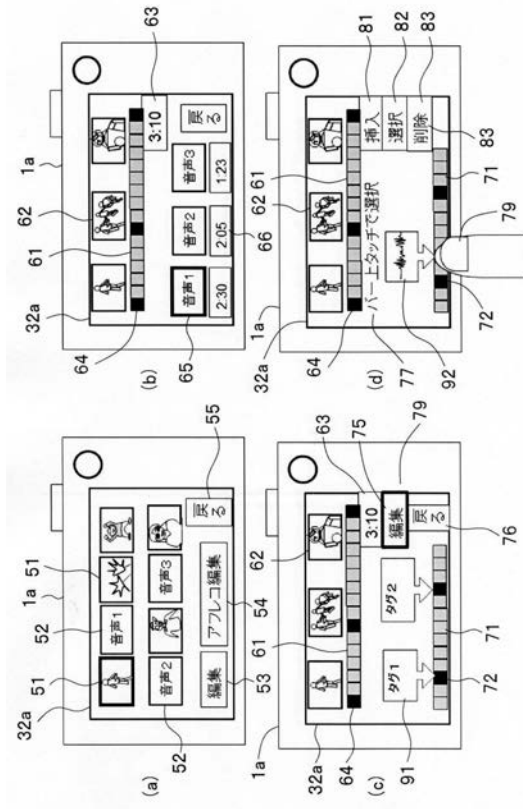
【 図 5 】



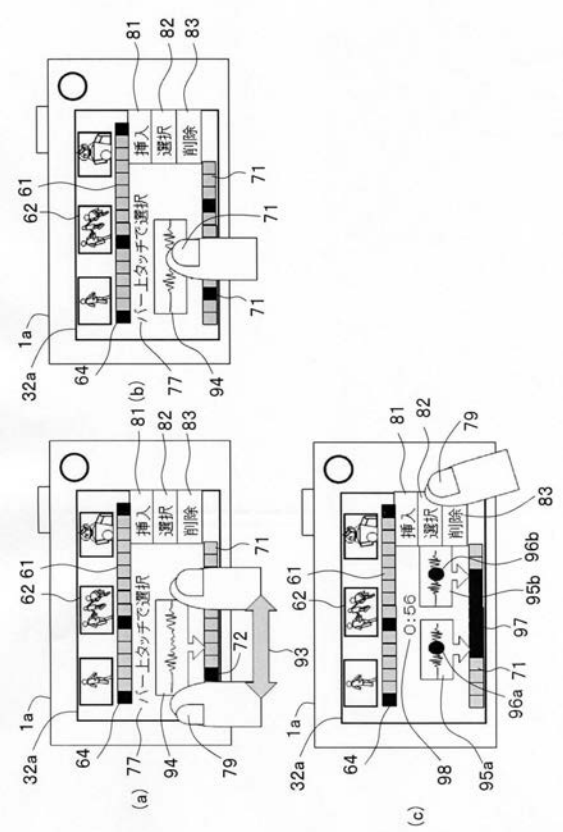
【図 6】



【図 8】



【図 9】



フロントページの続き

(72)発明者 野中 修

東京都渋谷区幡ヶ谷 2 丁目 4 3 番 2 号 オリンパス株式会社内

Fターム(参考) 5C053 FA14 FA27 LA01

5D077 HC50

5D110 CA01