

(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2006-148425
(P2006-148425A)

(43) 公開日 平成18年6月8日(2006.6.8)

(51) Int. Cl.		F I		テーマコード (参考)
HO4N 5/272 (2006.01)		HO4N	5/272	5CO23
G06T 7/20 (2006.01)		G06T	7/20	5LO96
			A	

審査請求 未請求 請求項の数 11 O L (全 13 頁)

(21) 出願番号 特願2004-334336 (P2004-334336)
(22) 出願日 平成16年11月18日(2004.11.18)

(特許庁注：以下のものは登録商標)

1. WINDOWS
2. PENTIUM

(71) 出願人 899000079
学校法人慶應義塾
東京都港区三田2丁目15番45号

(74) 代理人 100105647
弁理士 小栗 昌平

(74) 代理人 100105474
弁理士 本多 弘徳

(74) 代理人 100108589
弁理士 市川 利光

(74) 代理人 100115107
弁理士 高松 猛

(74) 代理人 100090343
弁理士 濱田 百合子

最終頁に続く

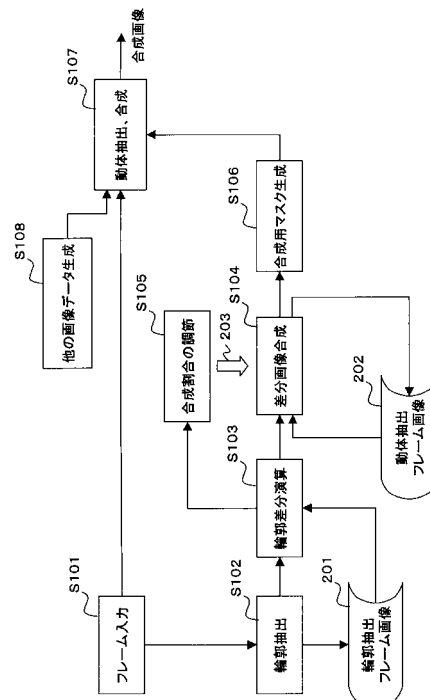
(54) 【発明の名称】 画像処理方法、画像処理装置、及びコンテンツ作成システム

(57) 【要約】

【課題】 コンピュータの処理負担を大きくすることなく鮮明な動体抽出を行って、他の画像と合成することができる画像処理方法を提供する。

【解決手段】 本発明の画像処理方法は、入力動画像の各フレーム画像に対して輪郭抽出処理を行い(ステップS102)、生成した輪郭抽出フレーム画像のフレーム間差分演算を行う(ステップS103)。そして、差分演算によって生成した動体抽出フレーム画像に基づいて、動体領域を識別するマスクデータを生成し(ステップS106)、生成したマスクデータを利用して、入力動画像における動体領域画像を他の画像と合成する(ステップS107)。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

コンピュータを用いてデジタル画像に対する処理を行う画像処理方法であって、
入力動画像の各フレーム画像に対して輪郭抽出処理を行い、輪郭抽出フレーム画像を生成する輪郭抽出ステップと、

前記輪郭抽出フレーム画像のフレーム間差分演算を行い、動体抽出フレーム画像を生成する動体抽出ステップと、

前記動体抽出フレーム画像に基づいて、前記入力動画像における動体領域を識別するマスクデータを生成するマスクデータ生成ステップと、

前記マスクデータを利用して、前記入力動画像における動体領域画像を他の画像と合成する画像合成ステップとを備える画像処理方法。 10

【請求項 2】

請求項 1 記載の画像処理方法であって、

前記動体抽出ステップは、前記フレーム間差分演算を行って生成した差分画像と、動体画像バッファに蓄積されている前フレームの動体抽出フレーム画像とを合成し、その合成画像を現フレームの動体抽出フレーム画像とするとともに、その合成画像によって前記動体画像バッファを更新する画像処理方法。

【請求項 3】

請求項 2 記載の画像処理方法であって、

前記動体抽出ステップは、前記差分画像と前フレームの動体抽出フレーム画像との合成割合を前記差分画像の平均輝度値に応じて変更する画像処理方法。 20

【請求項 4】

請求項 1 ないし 3 のいずれか 1 項記載の画像処理方法であって、

前記マスクデータ生成ステップは、前記動体抽出フレーム画像を、複数の走査直線に沿ってその走査直線の両側から走査するステップと、前記走査直線上の画素のうち、前記走査において最初に閾値以上となった画素間のすべての画素を含む領域を動体領域と認識するステップとを含む画像処理方法。

【請求項 5】

請求項 4 記載の画像処理方法であって、

前記複数の走査直線は、斜め方向の直線である画像処理方法。 30

【請求項 6】

請求項 1 ないし 5 のいずれか 1 項記載の画像処理方法であって、

前記マスクデータ生成ステップは、前記動体領域の輪郭近傍の合成割合を減少させたマスクデータを生成する画像処理方法。

【請求項 7】

請求項 1 ないし 6 のいずれか 1 項記載の画像処理方法における各ステップを、コンピュータに実行させるための画像処理プログラム。

【請求項 8】

請求項 7 記載の画像処理プログラムをインストールしたコンピュータを含む画像処理装置。 40

【請求項 9】

請求項 7 記載の画像処理プログラムをインストールしたコンピュータと、

前記コンピュータによる前記画像合成ステップで得られた合成画像データに基づく表示用合成画像信号を生成するビデオ信号生成手段と、

前記表示用合成画像信号に基づくデジタル動画データを含む動画ファイルを生成する動画ファイル生成手段とを備えるコンテンツ作成システム。

【請求項 10】

請求項 9 記載のコンテンツ作成システムを利用して生成した前記動画ファイルを講義ビデオとして出力する講義ビデオ作成システム。

【請求項 11】

請求項9記載のコンテンツ作成システムを利用して生成した前記デジタル動画データを、テレビ会議参加者の端末装置に配信する手段を備えるテレビ会議システム。

【発明の詳細な説明】

【技術分野】

【0001】

コンピュータを用いてデジタル画像に対する処理を行う画像処理方法、及び画像処理装置、並びに合成画像の生成を行うコンテンツ作成システムに関する。

【背景技術】

【0002】

動画像の一部又は全部と他の画像とを合成した合成画像を表示する技術は、テレビ会議システム、テレビ電話システム、講義ビデオシステム等の各種システムに利用されている（特許文献1、2参照）。動画像の一部を合成する場合、人物画像等の動体のみを他の画像に合成することが好ましい。例えば、テレビ電話システムに利用する場合、人物画像のみを抽出して任意の背景画像と合成することにより、自分の周囲の画像を送信しないようにすることができる。また、講義ビデオに講師の動画像を合成する場合、講師の輪郭を抽出して合成することにより、他の画像、例えば講義用資料の領域を拡大することができる。

10

【0003】

動画像から人物画像等の動体画像のみを抽出して他の画像と合成する技術としては、クロマキー合成によるものが周知である。しかし、クロマキー合成は、大掛かりな設備が必要であり、上記したような簡易なシステムに利用することは、困難である。

20

【0004】

動画像から動体画像を抽出する技術としては、特許文献3、4に記載されたものがある。特許文献3には、テレビ電話装置の撮像画面における人物領域抽出技術が記載されている。この文献においては、フレーム間の差分演算を行って動体を識別し、差分演算信号を所定の閾値に基づいて2値化することにより人物領域を抽出している。また、特許文献4には、安定化させた背景画像と入力動画像との差分を求めて、動体を認識している。

【0005】

しかし、特許文献3、4に記載された動体抽出技術においては、動体全体の輪郭が抽出されない場合があり、動体領域を精度よく特定するのが簡単ではない。また、動体が静止している状態での認識が困難である。

30

【0006】

【特許文献1】特開平7-67035号公報

【特許文献2】特開2000-175166号公報

【特許文献3】特開昭63-157593号公報

【特許文献4】特開平5-159060号公報

【発明の開示】

【発明が解決しようとする課題】

【0007】

本発明は、上記事情に鑑みなされたもので、コンピュータの処理負担を大きくすることなく動画像から鮮明な動体抽出を行って、他の画像と合成することができる画像処理方法、及び画像処理装置を提供することを目的とする。

40

【課題を解決するための手段】

【0008】

本発明の画像処理方法は、コンピュータを用いてデジタル画像に対する処理を行う画像処理方法であって、入力動画像の各フレーム画像に対して輪郭抽出処理を行い、輪郭抽出フレーム画像を生成する輪郭抽出ステップと、前記輪郭抽出フレーム画像のフレーム間差分演算を行い、動体抽出フレーム画像を生成する動体抽出ステップと、前記動体抽出フレーム画像に基づいて、前記入力動画像における動体領域を識別するマスクデータを生成するマスクデータ生成ステップと、前記マスクデータを利用して、前記入力動画像における

50

動体領域画像を他の画像と合成する画像合成ステップとを備えるものである。本発明によれば、コンピュータの処理負担を大きくすることなく動画像から鮮明な動体抽出を行って、他の画像と合成することができる。

【0009】

本発明の画像処理方法は、前記動体抽出ステップが、前記フレーム間差分演算を行って生成した差分画像と、動体画像バッファに蓄積されている前フレームの動体抽出フレーム画像とを合成し、その合成画像を現フレームの動体抽出フレーム画像とするとともに、その合成画像によって前記動体画像バッファを更新するものを含む。本発明によれば、動体抽出フレーム画像の輝度値が大きく低下しないので、動体の動きが小さくなったときでも、精度良く動体領域の認識ができる。

10

【0010】

本発明の画像処理方法は、前記動体抽出ステップが、前記差分画像と前フレームの動体抽出フレーム画像との合成割合を前記差分画像の平均輝度値に応じて変更するものを含む。本発明によれば、動体抽出フレーム画像の輝度値の変化を動体の動きの大きさに拘わらず抑えることができるので、さらに精度良く動体領域の認識ができる。

【0011】

本発明の画像処理方法は、前記マスクデータ生成ステップが、前記動体抽出フレーム画像を、複数の走査直線に沿ってその走査直線の両側から走査するステップと、前記走査直線上の画素のうち、前記走査において最初に閾値以上となった画素間のすべての画素を含む領域を動体領域と認識するステップとを含むものを含む。本発明によれば、動体の輪郭を構成する画素を簡単な処理で認識できるので、動体領域の認識処理の負荷を軽減することができる。

20

【0012】

本発明の画像処理方法は、前記複数の走査直線が、斜め方向の直線であるものを含む。

本発明によれば、ノイズの影響を減少させたマスクデータを生成することができる。

【0013】

本発明の画像処理方法は、前記マスクデータ生成ステップが、前記動体領域の輪郭近傍の合成割合を減少させたマスクデータを生成するものを含む。本発明によれば、滑らかな合成が可能となる。

【0014】

本発明の画像処理プログラムは、前記した画像処理方法における各ステップを、コンピュータに実行させるためのものである。

30

【0015】

本発明の画像処理装置は、前記した画像処理プログラムをインストールしたコンピュータを含むものである。

【0016】

本発明のコンテンツ作成システムは、前記した画像処理プログラムをインストールしたコンピュータと、前記コンピュータによる前記画像合成ステップで得られた合成画像データに基づく表示用合成画像信号を生成するビデオ信号生成手段と、前記表示用合成画像信号に基づくデジタル動画データを含む動画ファイルを生成する動画ファイル生成手段とを備えるものである。

40

【0017】

本発明の講義ビデオ作成システムは、前記したコンテンツ作成システムを利用して生成した前記動画ファイルを講義ビデオとして出力するものである。

【0018】

本発明のテレビ会議システムは、前記したコンテンツ作成システムを利用して生成した前記デジタル動画データを、テレビ会議参加者の端末装置に配信する手段を備えるものである。

【発明の効果】**【0019】**

50

以上の説明から明らかなように、本発明によれば、コンピュータの処理負担を大きくすることなく動画像から鮮明な動体抽出を行って、他の画像と合成することができる画像処理方法、及び画像処理装置を提供することができる。

【発明を実施するための最良の形態】

【0020】

以下、本発明の実施の形態について、図面を用いて説明する。なお、以下の説明では、動画像を含むコンテンツを作成するコンテンツ作成システムを適用例としている。

【0021】

図4は、コンテンツ作成システムの一例である講義ビデオ作成システムの概略構成を示す図である。図1の講義ビデオ作成システムは、教室等での講義と同時に講義ビデオを作成するものであり、講師用コンピュータ1、カメラ2、タブレット3、プロジェクタ4、スキャンコンバータ5、録画用コンピュータ6、マイクロホン7、ビデオサーバ8を含んで構成される。

10

【0022】

講師用コンピュータ1は、講師が講義に使用するコンピュータであり、例えばノート型PCである。講師用コンピュータ1には、予め、Power Point等のプレゼンテーションソフトウェアで作成された講義用素材が用意されている。また、Webサイトのコンテンツを講義に使用する場合は、Webブラウザをインストールしておくと共にインターネットに接続可能としておく。

【0023】

講師用コンピュータ1には、カメラ2とタブレット3が、例えばUSB接続により接続される。カメラ2は、講義中の講師を撮影する講師撮影用カメラであって、動画像を講師用コンピュータ1に入力するものであり、タブレット3は、講義中の板書と同様に、講師が手書きデータを入力するためのものである。講師用コンピュータ1には、カメラ2からの映像をデスクトップ上に表示させるソフトウェアと、タブレット3からの手書き情報をデスクトップ上に描画するためのソフトウェアが予めインストールされる。これらのソフトウェアは、周知の技術により簡単に作成することができる。既に作成されたソフトウェアは、例えば、「COE e-Learning Tools」、<URL: <http://coe-el.sfc.keio.ac.jp/>>でダウンロードすることができる。このサイトからダウンロードされるソフトウェアは、カメラ2からの撮影動画像及びタブレット3からの手書き画像と1又は複数の講義用素材画像とを合成した合成画像データを生成するものである。

20

30

【0024】

ここで生成される合成画像データは、複数の画像を重ね合わせたものでも、一部の画像を部分的に上書きしたものでも、それぞれの画像を所定の大きさの領域に配置したものでよい。ただし、カメラ2からの動画像については、講師撮影領域等の動体領域を認識し、認識した動体領域の画像のみが合成される。カメラ2からの動画像との合成処理については、後述する。また、合成する各アプリケーション画像（カメラ画像、手書き画像を含む。）の大きさは、任意であり、講師が変更可能である。

【0025】

講師用コンピュータ1の外部モニタ出力端子（図示せず）には、デスクトップの画面を映像として図示しない大規模スクリーンに表示するためのプロジェクタ4が接続される。

40

スキャンコンバータ5は、講師用コンピュータ1の外部モニタ出力端子（図示せず）に接続され、この出力端子から出力されるデジタル信号を表示用画像信号の1つであるアナログビデオ信号に変換するものである。

【0026】

録画用コンピュータ6は、スキャンコンバータ5で取得したアナログビデオ信号をビデオキャプチャボードにより入力し、既存のビデオキャプチャソフトを用いて動画ファイル、例えばWindows Media形式(.WMV)にリアルタイムでエンコードする。Windows Media形式の動画ファイルは、非常に軽量である。例えば、録画解像度を640pixels×480pixels、配信ビットレートを250bpsに設定すると、1時間あたりのファイル容量は約1

50

00MBである。録画解像度を640pixels×480pixelsで、講師用コンピュータ1の画面上の資料及びタブレット描画による板書は、問題なく判読可能である。また、フレームレートは、10fps程度であり、講師の表情や板書の動き等を違和感なく閲覧することが可能である。録画用コンピュータ6の性能は、例えば、PentiumIV2.4GHzプロセッサ、メモリ1GB、ハードディスク容量180GBである。

【0027】

マイクロホン7は、講師の音声信号取得するためのものであり、録画用コンピュータ6に接続される。録画用コンピュータ6は、動画ファイルの生成時に音声データの付加を行う。なお、図4では、マイクロホン7を録画用コンピュータ6に接続したが、講師用コンピュータ1に接続し、講師用コンピュータ1で取得した音声データを録画用コンピュータ6

10

【0028】

ビデオサーバ8は、録画用コンピュータ6で作成された動画ファイルがアップされ、ストリーミング配信するものである。ビデオサーバ8は、例えば、Windows 2000Serverがインストールされたコンピュータであり、その性能は、Pentium III, 750MHzプロセッサ、メモリ512MB、ハードディスク容量240GBである。

【0029】

このような構成を有する講義ビデオを作成システムの動作について説明する。講義室には、予め、講師用コンピュータ1以外の機器が用意されている。講師は、講義用素材を記憶した自己のコンピュータ1のUSB端子にカメラ2、タブレット3を接続し、ビデオ出力端子にプロジェクタ4及びスキャンコンバータ5を接続する。そして、全ての機器を動作させ、講師用コンピュータ1に用意した講義資料表示用のアプリケーションを起動する。

20

【0030】

講師は、このようなシステムの状態で講義を開始し、講師用コンピュータ1に必要な講義用資料を表示させながら講義を進める。講師用コンピュータ1の画像表示信号は、プロジェクタ4に送られるので、図示しない大規模スクリーンにも表示される。講師用コンピュータ1には、講義用資料の一部にカメラ2からの撮影画像が表示される。図5に、表示画像の一例を示す。図5は、表示画面400のほぼ大部分の領域に、プレゼンテーションソフトウェアによる表示画像410を表示させ、さらに表示画面400の右下部に講師の撮影映像420が表示されている状態を模式的に示したものである。図5に示すように、講師の撮影映像420は、講師の撮影領域(動体領域)のみが抽出されて合成されている。

30

【0031】

講師用コンピュータ1の画像表示信号は、同時にスキャンコンバータ5に送られ、スキャンコンバータ5では、画像表示信号に基づくアナログビデオ信号が生成される。そして、生成されたアナログビデオ信号は、録画用コンピュータ6に送られ、デジタル動画ファイルに変換される。すなわち、アナログビデオ信号は、録画用コンピュータ6のビデオキャプチャボード(図示せず)を介して入力され、既存のビデオキャプチャソフトを用いてWindows Media形式(WMV)のデジタル画像データにリアルタイムでエンコードされる。その際、マイクロホン7によって入力された音声信号も同時にデジタル化され、合わせて出力される。

40

【0032】

録画用コンピュータ6で作成されたWindows Media形式(WMV)の動画ファイルは、ビデオサーバ8にアップロードされる。そして、図示しないネットワークを介して講義ビデオの配信に供せられる。アップロードされる講義ビデオは動画ファイルであるので、ストリーム配信も可能であり、したがって、実際の講義とほぼ同時のライブ配信も可能であり、遠隔講義も実現できる。

【0033】

次に、講師用コンピュータ1が行う画像合成処理について説明する。複数の画像信号の

50

合成処理自体は既述のように周知のものであるので、ここでは、カメラ2からの動画像から講師撮影領域等の動体領域を抽出する技術を主体に説明する。

【0034】

図1は、本発明の実施の形態の画像処理方法を説明する概略フロー図である。図1に示す処理は、講師用コンピュータ1が行う。

【0035】

カメラ2からの撮影動画像に基づくデジタルフレームデータを、所定のレートで入力され(ステップS101)、輪郭抽出処理が施される(ステップS102)。輪郭抽出処理自体は、周知の技術であり、例えばラプラシアン演算が利用可能であり、輪郭が強調された画像が得られる。輪郭抽出処理が施されたフレームデータは、輪郭抽出フレーム画像201として蓄積されるとともに、ステップS103の輪郭差分演算の対象となる。なお、輪郭抽出処理は、入力されるすべてのフレームデータに対して行ってもよいが、所定間隔のフレームにのみ行ってもよい。輪郭抽出フレーム画像は、適宜のバッファメモリに蓄積され、順次更新される。

10

【0036】

ステップS103では、ステップS102の輪郭抽出処理で生成された輪郭抽出フレーム画像と、蓄積された前フレームの輪郭抽出フレーム画像との差分を演算する。得られた差分画像は、撮影動画像の動体部分が強調された画像となる。また、差分演算の対象となる画像は輪郭が強調された画像であるので、単にフレーム間の差分演算を行ったものに比較して鮮明な画像が得られる。

20

【0037】

ステップS104では、ステップS103の輪郭差分演算処理で生成された差分画像と前フレームで得られた動体抽出フレーム画像202とを合成し、現フレーム動体抽出フレーム画像を生成する。前フレームの動体抽出フレーム画像と合成する理由は、動体の動きが小さい場合でも精度よく動体抽出を行うためである。すなわち、撮影動画中の動体の動きが小さい場合、ステップS103の輪郭差分演算処理で生成された差分画像が不鮮明になるので、前フレームで生成した動体抽出フレーム画像を合成することにより、鮮明にするためである。後述するように、このステップで生成された動体抽出フレーム画像に基づいて、動体領域の合成を行うためのマスクデータを生成するので、動体の動きが小さいばあいでも、精度良く動体領域のみの抽出及び合成が可能となる。

30

【0038】

現フレームの差分画像と蓄積された動体抽出フレーム画像との合成割合は、一定としてもよいし、現フレームの差分画像に応じて変化させてもよい(ステップS105の合成割合の調節処理)。変化させる場合、ステップS105で現フレームの差分画像の平均輝度を求め、その値に応じた合成割合制御情報203を利用する。具体的には、平均輝度値が低い場合は、動体の動きが小さく動体領域が精度よく認識できないので、前フレームの動体抽出フレーム画像の合成割合を相対的に大きくする。なお、ここでの合成割合は、その合計値を必ずしも「1」とする必要はない。例えば、現フレームの差分画像の合成割合を変化させず、前フレームの合成割合を変化させるようにする。

【0039】

合成処理で得られた合成画像は、動体抽出フレーム画像202として適宜のバッファメモリに蓄積され、順次更新される。なお、ステップS104の差分画像合成処理は、省略も可能である。その場合、動体抽出フレーム画像202の蓄積及び合成割合の調節処理も省略される。

40

【0040】

ステップS106では、ステップS104の差分合成処理で生成した動体抽出フレーム画像に基づいて、入力フレームにおける動体領域を識別するためのマスクデータを生成する。動体抽出フレーム画像は、動体領域の輪郭近傍が他の領域と比較して高輝度の画像であるので、所定の閾値より高輝度を示す領域に囲まれる部分を動体領域として認識し、マスクデータを生成する。

50

【0041】

図2は、マスクデータ生成処理の一例を説明する図である。マスクデータ生成に際しては、動体抽出フレーム画像300を斜め方向の平行な走査線301a、301b、・・・、301n、・・・に沿って走査し、各画素の輝度値と閾値とを比較する。走査及び比較は、最初に最左上の走査線301aに沿って行き、次いで走査線301bに移り、最後の最右下の走査線301zに沿って行う。各走査線上の画素の輝度値の比較は、まず、走査直線の左下端（走査線301aの場合は、端部302a）から始め、閾値以上の画素が認識できた時点で、その画素にマークを付与し、走査を中止する。そして、同じ走査直線の右上端から再開し、同様に閾値以上の画素が認識できた時点で、その画素にマークを付与し、走査を中止する。なお、閾値以上の画素が認識できない場合は、左下端からの走査を最後まで行う。図2の例では、走査線301nに沿って端部302nから右上方向に、各画素の輝度値の比較を行った結果、画素303nで初めて閾値以上になってマークが付与されたことを示している。この場合、端部304nから左下方向に、画素値の比較を再開し、画素305nで初めて閾値以上になってマークが付与されている。なお、図2においては、走査方向及び非走査部を示すために、同一の走査線を実線と破線と中抜き線とで区別して示している。また、走査線の数も間引いて記載してある。

10

【0042】

すべての走査線に沿った画素の輝度値の比較処理が終了すると、同一の走査線に沿った画素で、マークを付与した画素に挟まれる画素にもマークを付与する。図2の走査線301nの例では、画素303nと画素305nに挟まれる画素にもマークを付与する。そして、図3に示すようなマークを付与した画素位置を動体領域と認識したマスクデータを生成する。なお、動体領域と非動体領域の境界近傍の所定個数の画素については、別なマークを付与し、後述する動体抽出及び合成処理における合成割合の変更に利用してもよい。また、マスクデータが点データとして得られる場合（1つの走査直線において、1つの画素のみの輝度値が閾値以上である場合）は、ノイズとして動体領域とはしない。

20

【0043】

ノイズの影響でマスクデータが点データとして得られる確率は、斜め方向に走査することによって高くなるので、斜め方向の走査が好ましい。このことは、例えば図2の点Aにおいて輝度値が閾値より大きくなった場合を想定すると明らかである。すなわち、斜め方向の走査では、この点はノイズとして簡単に除去できるが、縦方向又は横方向の走査の場合、縦又は横に線状のノイズがのることになる。ただし、走査処理自体は縦方向又は横方向の方が簡単であるので、縦方向又は横方向の走査を行ってもよい。

30

【0044】

次いで、ステップS107では、ステップS101で入力されたフレーム画像から動体部分を抽出し、他の画像データと合成して合成画像を出力する。動体部分の抽出は、ステップS106で生成した図3に示すようなマスクデータを利用する。図3の例では、黒の部分の画素を動体領域の画素を認識してフレームデータから抽出し、他の画像データの該当部分の画素データを、抽出した画素データで置き換える。動体領域と非動体領域の境界近傍の画素に異なるマスクを利用する場合、境界領域の部分は、抽出した動体部分の画素データと他の画素データの画素データとを所定の比率で合成したデータとする。なお、他の画像データは、例えば、プレゼンテーションソフトウェアで生成された画像データであり、この画像データは、マスクデータ生成処理と平行して生成される（ステップS108）。

40

【0045】

以上、本発明の画像処理方法をコンテンツ作成システムの一例である講義ビデオ作成システムに適用した例について説明したが、コンテンツ作成システムの他の例であるテレビ会議に適用することも可能である。

【0046】

図6は、コンテンツ作成システムの他の例であるテレビ会議システムの概略構成を示す図である。図6のテレビ会議システムは、ネットワーク100を介して接続された会議用

50

表示サーバ20、参加者用コンピュータ30、40、及び会議用表示サーバ20に接続された主参加者用コンピュータ10を含んで構成される。

【0047】

主参加者用コンピュータ10は、会議の主参加者がテレビ会議端末として使用するコンピュータである。主参加者用コンピュータ10には、主参加者の映像を撮影するカメラ11、主参加者の音声を取得するマイクロホン12、主参加者の手書き情報を入力するタブレット13が接続されるとともに、プレゼンテーションソフトウェア等による会議資料の表示が可能とされる。そして、主参加者の撮影映像、タブレット13による手書き画像データ、会議資料データ、音声データは、直接会議用表示サーバ20に送られる。

【0048】

参加者用コンピュータ30及び40は、会議の参加者がテレビ会議端末として使用するコンピュータである。図6では2台のコンピュータを記載してあるが、台数は任意である。参加者用コンピュータ30及び40には、参加者の映像を撮影するカメラ31及び41、参加者の音声を取得するマイクロホン32及び42、参加者の手書き情報を入力するタブレット33及び43が接続される。タブレット33及び43は省略が可能である。参加者の撮影映像、タブレット33及び43による手書き画像データ、マイクロホン32及び42で取得した音声データは、ネットワークを介して会議用表示サーバ20に送られる。

10

【0049】

会議用表示サーバ20は、主参加者用コンピュータ10からの画像データと、参加者用コンピュータ30及び40からの画像データを合成し、合成した画像データに基づくアナログビデオ信号を生成し、さらに生成したアナログビデオ信号に基づくデジタル動画データを含む動画ファイルを生成する。ここで、参加者の撮影映像を合成する場合は、カメラ11、31、41からの動画像については、参加者講師撮影領域等の動体領域を認識し、認識した動体領域の画像のみが合成される。合成処理の手順は、先に説明したとおりである。

20

【0050】

そして、生成した動画ファイルを主参加者用コンピュータ10に直接送信するとともに、ネットワーク100を介して参加者用コンピュータ30及び40に送信する。また、その際、合わせて、受信した音声データを生成した動画データとともに送信する。したがって、会議の参加者は、会議資料画像に各参加者の撮影画像が合成された画像を、それぞれのコンピュータに備えられた表示器(図示せず)によって見ることができる。

30

【0051】

会議用表示サーバのアナログビデオ信号は、録画用コンピュータ50に送られ、録画用コンピュータ50では、アナログビデオ信号をビデオキャプチャボードにより入力し、既存のビデオキャプチャソフトを用いて動画ファイル、例えばWindows Media形式(.WMV)にリアルタイムでエンコードする。同時に音声データも取得し、デジタル化する。録画用コンピュータ50で生成された音声データ付きビデオデータは、ビデオサーバ60にアップロードされ、会議のストリーム配信及び記録に利用される。

【0052】

録画用コンピュータ50及びビデオサーバ60は、図4の講義ビデオ作成システムにおける録画用コンピュータ6及びビデオサーバ8と同様のものであるので、説明を省略する。

40

【0053】

なお、会議用表示サーバ20、参加者用コンピュータ30及び40相互間の画像データ及び音声データの送受信は、既存のインターネットテレビ会議システムを利用して行う。

インターネット会議システムは、例えば、<URL:http://messenger.yahoo.co.jp/>や<URL:http://www.cybernet.co.jp/webex/>に示されるものが利用可能である。

【0054】

タブレット13、33、43からの手書き情報を合成する場合、会議用表示サーバ20は、各コンピュータ10、30、40からのタブレット使用要求に応じていずれか1つの

50

タブレットからの手書き情報をリアルタイムで合成する。

【 0 0 5 5 】

図 6 のテレビ会議システムでは、ネットワーク 1 0 0 に接続された会議用表示サーバ 2 0 が、受信した画像の合成、アナログビデオ信号の生成、デジタル画像データの生成を行うものとして記載したが、処理能力によっては、主参加者用コンピュータ 1 0 が実行しているプレゼンテーションソフトウェア等のアプリケーションプログラムの実行も行うようにしてもよい。

【 0 0 5 6 】

その場合、会議用表示サーバ 2 0 には、マルチユーザによる利用が可能となるターミナルサーバ機能が付加される。そして、ターミナルサーバのクライアントとしても動作する主参加者用コンピュータ 1 0、参加者用コンピュータ 3 0、4 0 との間でリモートデスクトッププロトコル (R D P) でデータの送受信を行い、必要なアプリケーションプログラムの実行が行われる。このような構成とすると、テレビ会議の参加者が、それぞれ必要な会議資料の提示を制御することができる。

【 図面の簡単な説明 】

【 0 0 5 7 】

【 図 1 】 本発明の実施の形態の画像処理方法を説明するための概略フロー図

【 図 2 】 本発明の実施の形態の画像処理方法における動体領域の認識処理の一例を説明する図

【 図 3 】 本発明の実施の形態の画像処理方法におけるマスクデータの一例を示す図

【 図 4 】 本発明の実施の形態の講義ビデオ作成システムの概略構成を示す図

【 図 5 】 本発明の実施の形態の講義ビデオ作成システムにおける講義用コンピュータの表示画像の一例を示す図

【 図 6 】 本発明の実施の形態のテレビ会議システムの概略構成を示す図

【 符号の説明 】

【 0 0 5 8 】

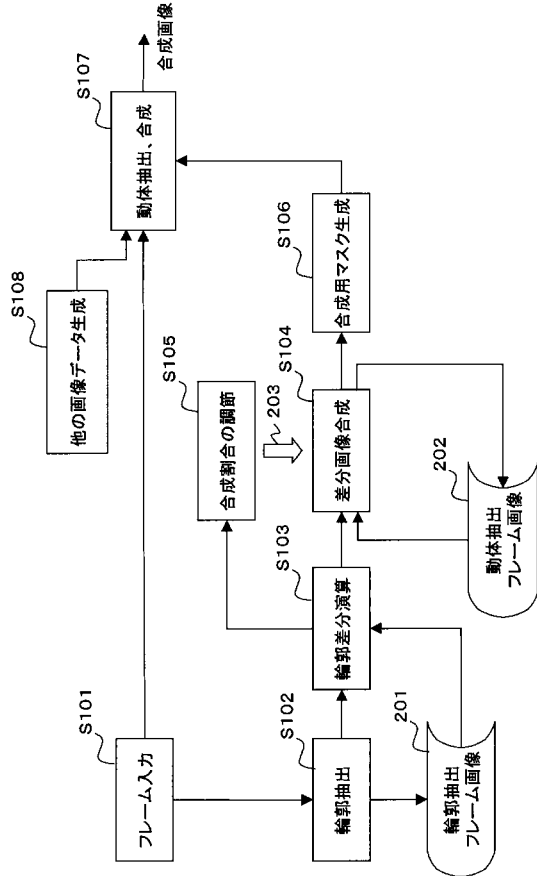
- 1・・・講師用コンピュータ
- 2、11、31、41・・・カメラ
- 3、13、33、43・・・タブレット
- 4・・・プロジェクタ
- 5・・・スキャンコンバータ
- 6、50・・・録画用コンピュータ
- 7、12、32、42・・・マイクロホン
- 8、60・・・ビデオサーバ
- 10・・・主参加者用コンピュータ
- 20・・・会議用表示サーバ
- 30、40・・・参加者用コンピュータ
- 100・・・ネットワーク

10

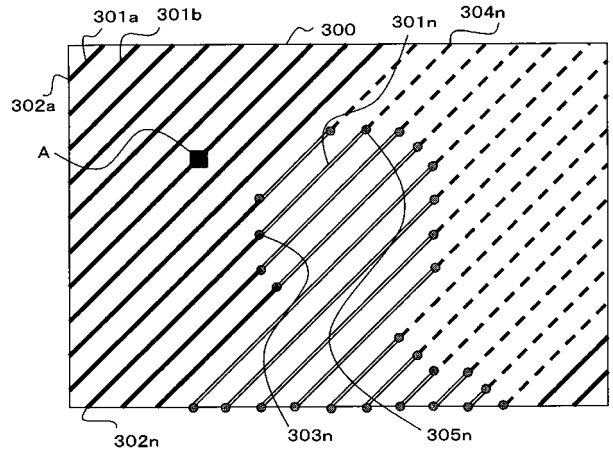
20

30

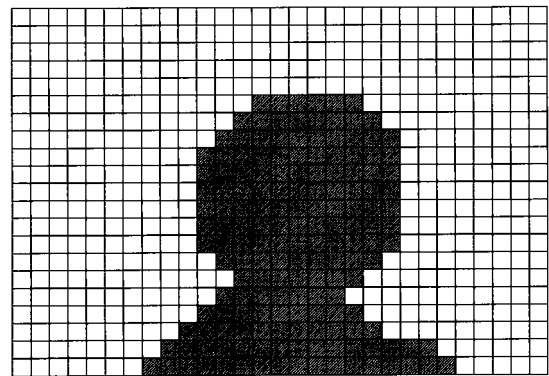
【図1】



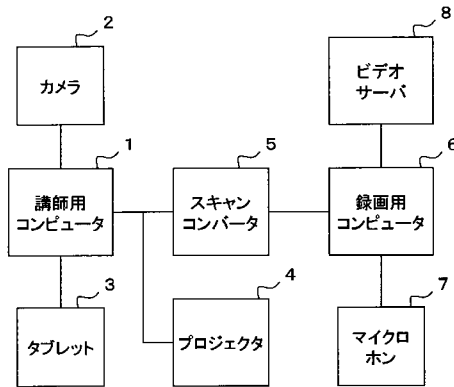
【図2】



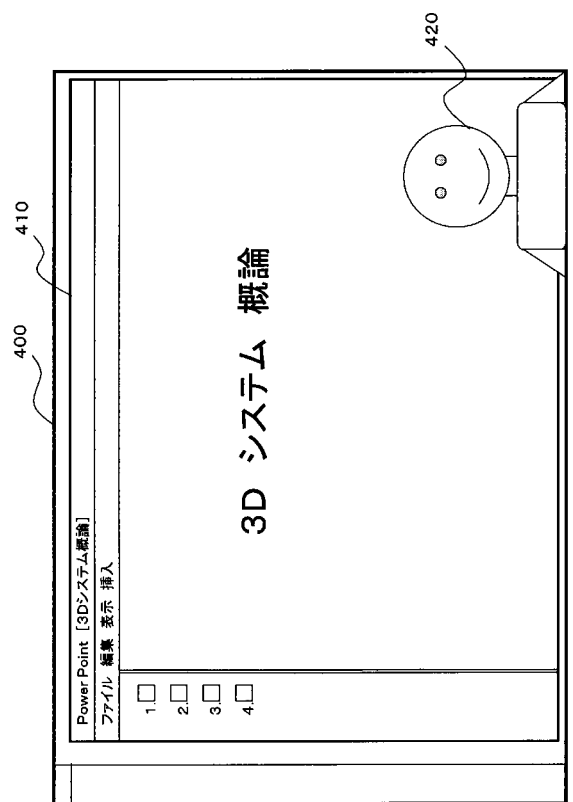
【図3】



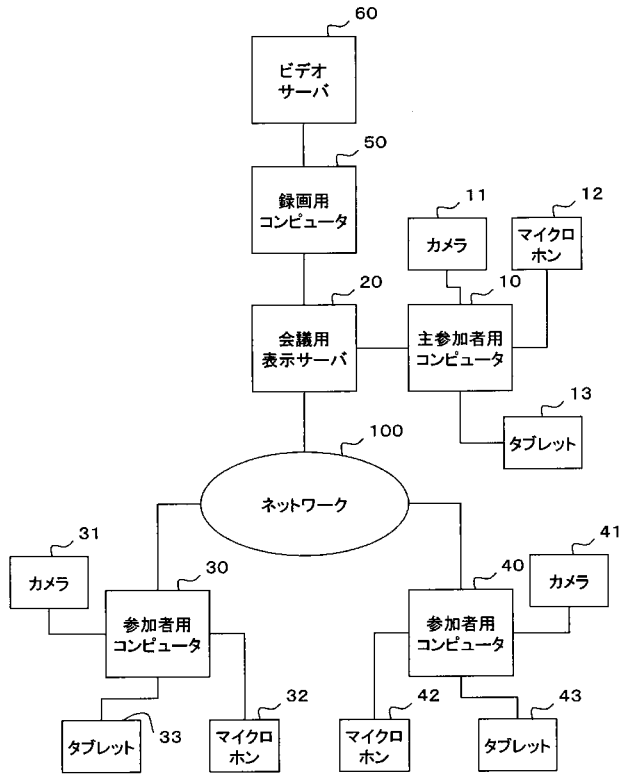
【図4】



【図5】



【 図 6 】



フロントページの続き

(72)発明者 千代倉 弘明

神奈川県藤沢市遠藤5 3 2 2 慶應義塾大学湘南藤沢キャンパス内

(72)発明者 林 佑樹

神奈川県藤沢市遠藤5 3 2 2 慶應義塾大学湘南藤沢キャンパス内

Fターム(参考) 5C023 AA06 AA16 BA04 BA13 CA01 DA04 EA03 EA05

5L096 AA06 CA04 FA06 GA08 GA51 HA03