

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5092093号
(P5092093)

(45) 発行日 平成24年12月5日(2012.12.5)

(24) 登録日 平成24年9月28日(2012.9.28)

(51) Int.Cl.		F I			
G06T	7/60	(2006.01)	G06T	7/60	150P
G06T	1/00	(2006.01)	G06T	1/00	340A
H04N	1/387	(2006.01)	H04N	1/387	
H04N	5/232	(2006.01)	H04N	5/232	Z

請求項の数 9 (全 41 頁)

(21) 出願番号	特願2008-83213 (P2008-83213)	(73) 特許権者	393031586
(22) 出願日	平成20年3月27日(2008.3.27)		株式会社国際電気通信基礎技術研究所
(65) 公開番号	特開2009-237899 (P2009-237899A)		京都府相楽郡精華町光台二丁目2番地2
(43) 公開日	平成21年10月15日(2009.10.15)	(74) 代理人	100090181
審査請求日	平成23年2月18日(2011.2.18)		弁理士 山田 義人
(出願人による申告)平成19年度独立行政法人情報通信研究機構「民間基盤技術研究促進制度／軽度脳障害者のための情報セラピーインタフェースの研究開発」、産業技術力強化法第19条の適用を受ける特許出願		(72) 発明者	米澤 朋子
			京都府相楽郡精華町光台二丁目2番地2
		(72) 発明者	株式会社国際電気通信基礎技術研究所内
			山添 大丈
		(72) 発明者	京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内
		(72) 発明者	内海 章
			京都府相楽郡精華町光台二丁目2番地2
			株式会社国際電気通信基礎技術研究所内

最終頁に続く

(54) 【発明の名称】 画像処理装置

(57) 【特許請求の範囲】

【請求項1】

複数の被写体を含む被写体像を連続的に撮影した複数の撮影画像のそれぞれに対応する複数の撮影画像データを記憶手段に記憶する画像取得手段、

前記画像取得手段によって取得された複数の撮影画像データのそれぞれを順番に読み出す読出手段、

前記読出手段によって読み出された撮影画像データに基づいて、少なくとも前記複数の被写体の各々の視線方向を含む被写体情報を検出する被写体情報検出手段、

少なくとも理想の視線方向の情報を含むテンプレートを複数種類記憶するテンプレート記憶手段、

前記テンプレート記憶手段から所望のテンプレートを選択するテンプレート選択手段、

前記被写体情報が示す視線方向と、前記テンプレート選択手段によって選択されたテンプレートが示す理想の視線方向とに基づいて、前記テンプレートに基づく理想の画像に対する前記撮影画像データの適合率を算出する適合率算出手段、

前記適合率算出手段によって算出された適合率が最も高い撮影画像データを選択する撮影画像データ選択手段、

前記撮影画像データ選択手段によって選択された撮影画像データの適合率が一定値以上であるかどうかを判断する判断手段、および

前記判断手段によって前記適合率が一定値以上であることが判断されたとき、前記撮影画像データ選択手段によって選択された撮影画像データを所望の撮影画像データとして決

定する決定手段を備える、画像処理装置。

【請求項 2】

前記撮影画像データ選択手段によって選択された撮影画像データにおいて、前記適合率を低下させている被写体についての部分画像データであって、前記適合率を向上させる部分画像データを、当該撮影画像データ選択手段によって選択されていない撮影画像データから抽出する抽出手段、および

前記抽出手段によって抽出された部分画像データを、前記撮影画像データ選択手段によって選択された撮影画像データの該当する部分に貼り付けて合成した合成画像データを生成する生成手段をさらに備える、請求項 1 記載の画像処理装置。

【請求項 3】

前記適合率を向上させるための指示を出力する指示手段をさらに備える、請求項 1 または 2 記載の画像処理装置。

【請求項 4】

前記被写体情報は、被写体の顔の向きをさらに含み、
前記テンプレートは、理想の顔の向きをさらに含み、
前記適合率は、前記被写体の顔の向きの前記理想の顔の向きに対する顔方向適合率を含む、請求項 1 ないし 3 のいずれかに記載の画像処理装置。

【請求項 5】

前記被写体情報は、被写体の体の向きをさらに含み、
前記テンプレートは、理想の体の向きをさらに含み、
前記適合率は、前記被写体の体の向きの前記理想の体の向きに対する体方向適合率を含む、請求項 1 ないし 4 のいずれかに記載の画像処理装置。

【請求項 6】

前記被写体情報は、被写体の表情をさらに含み、
前記テンプレートは、理想の表情をさらに含み、
前記適合率は、前記被写体の表情の前記理想の表情に対する表情適合率を含む、請求項 1 ないし 5 のいずれかに記載の画像処理装置。

【請求項 7】

前記被写体情報は、被写体の姿勢をさらに含み、
前記テンプレートは、理想の姿勢をさらに含み、
前記適合率は、前記被写体の姿勢の前記理想の姿勢に対する姿勢適合率を含む、請求項 1 ないし 6 のいずれかに記載の画像処理装置。

【請求項 8】

前記被写体情報は、被写体の位置をさらに含み、
前記テンプレートは、理想の位置をさらに含み、
前記適合率は、前記被写体の位置の前記理想の位置に対する姿勢適合率を含む、請求項 1 ないし 7 のいずれかに記載の画像処理装置。

【請求項 9】

複数の被写体を含む被写体像の撮影画像に対応する撮影画像データを取得する取得手段、

前記取得手段によって取得された撮影画像データに基づいて、少なくとも前記複数の被写体の各々の視線方向を含む被写体情報を検出する被写体情報検出手段、

少なくとも理想の視線方向の情報を含むテンプレートを複数種類記憶するテンプレート記憶手段、

前記テンプレート記憶手段から所望のテンプレートを選択するテンプレート選択手段、
前記被写体情報が示す視線方向と、前記テンプレート選択手段によって選択されたテンプレートが示す理想の視線方向とに基づいて、前記テンプレートに基づく理想の画像に対する前記撮影画像データの適合率を算出する適合率算出手段、

前記適合率算出手段によって算出された適合率が一定値以上であるかどうかを判断する判断手段、および

10

20

30

40

50

前記判断手段によって前記適合率が一定値以上であることが判断されたとき、前記取得手段によって取得された撮影画像データを記憶する撮影画像データ記憶手段を備える、画像処理装置。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は画像処理装置に関し、特にたとえば、複数の被写体を含む被写体像についての所望の撮影画像を得るための、画像処理装置に関する。

【背景技術】

【0002】

従来この種の画像処理装置の一例が特許文献1に開示されている。この特許文献1の画像合成装置によれば、たとえば、連続して撮影して3駒の被写体像を得て、3駒の被写体像における同一の被写体についての顔画像のうちから最も表情の良い顔画像を被写体毎に取得する。そして、たとえば、最もぶれの少ない背景画像が3駒の被写体像の中から選択され、各被写体について取得された最も表情の良い顔画像が、選択された背景画像を有する被写体像に合成される。

【特許文献1】特開2007-299297号[G06T 1/00, G06T 11/80, H04N 1/387]

【発明の開示】

【発明が解決しようとする課題】

【0003】

この背景技術では、複数駒の被写体像から、最も表情の良い顔画像や所望のポーズの人間像を被写体毎に取得し、別途選択した背景画像を有する被写体像に合成するので、記念撮影のような集合写真を撮影した場合には、たとえば、被写体の多くが笑顔であるような合成画像を得られる。しかし、被写体の視線方向や顔の向きについては何ら考慮されていないため、集合写真にも関わらず、各被写体の視線ないし顔の向きがばらばらであるような合成画像が生成されると、違和感を覚えてしまう可能性がある。また、この背景技術では、被写体毎の視線ないし顔の向きが異なるポスター写真などを生成するのは困難である。

【0004】

それゆえに、この発明の主たる目的は、新規な、画像処理装置を提供することである。

【0005】

この発明の他の目的は、所望の画像を簡単に取得することができる、画像処理装置を提供することである。

【課題を解決するための手段】

【0006】

本発明は、上記の課題を解決するために、以下の構成を採用した。なお、括弧内の参照符号および補足説明等は、本発明の理解を助けるために後述する実施の形態との対応関係を示したものであって、本発明を何ら限定するものではない。

【0007】

第1の発明は、複数の被写体を含む被写体像を連続的に撮影した複数の撮影画像のそれぞれに対応する複数の撮影画像データを記憶手段に記憶する画像取得手段、画像取得手段によって取得された複数の撮影画像データのそれぞれを順番に読み出す読出手段、読出手段によって読み出された撮影画像データに基づいて、少なくとも複数の被写体の各々の視線方向を含む被写体情報を検出する被写体情報検出手段、少なくとも理想の視線方向の情報を含むテンプレートを複数種類記憶するテンプレート記憶手段、テンプレート記憶手段から所望のテンプレートを選択するテンプレート選択手段、被写体情報が示す視線方向と、テンプレート選択手段によって選択されたテンプレートが示す理想の視線方向とに基づいて、テンプレートに基づく理想の画像に対する撮影画像データの適合率を算出する適合率算出手段、適合率算出手段によって算出された適合率が最も高い撮影画像データを選択する撮影画像データ選択手段、撮影画像データ選択手段によって選択された撮影画像デー

10

20

30

40

50

タの適合率が一定値以上であるかどうかを判断する判断手段、および判断手段によって適合率が一定値以上であることが判断されたとき、撮影画像データ選択手段によって選択された撮影画像データを所望の撮影画像データとして決定する決定手段を備える、画像処理装置である。

【0008】

第1の発明では、画像処理装置(10(または12))の画像取得手段(S306, S362)は、複数の被写体を含む被写体像を連続的に撮影した複数の撮影画像のそれぞれに対応する複数の撮影画像データを記憶手段(74)に記憶する。読出手段(S308, S326, S328, S364, S378, S380)は、画像取得手段によって取得された複数の撮影画像データのそれぞれを順番に読み出す。視線方向検出手段(S314, S370)は、読出手段によって読み出された撮影画像データに基づいて、少なくとも複数の被写体の各々の視線方向を含む被写体情報を検出(推定)する。また、テンプレート記憶手段(74(または38))は、少なくとも理想の視線方向の情報を含むテンプレートを複数種類記憶する。たとえば、ユーザの指示に従って、テンプレート選択手段(S300)は、テンプレート記憶手段から所望のテンプレートを選択する。適合率算出手段(S322, S374)は、被写体情報が示す視線方向と、テンプレート選択手段によって選択されたテンプレートが示す理想の視線方向とに基づいて、テンプレートに基づく理想の画像に対する撮影画像データの適合率を算出する。たとえば、各被写体の視線方向の適合率や複数の被写体を含む撮影画像全体としての視線方向の分散・尖度についての適合率を算出するのである。撮影画像データ選択手段(S330, S382)は、適合率算出手段によって算出された適合率が最も高い撮影画像データを選択する。判断手段(S332, S384)は、撮影画像データ選択手段によって選択された撮影画像データの適合率が一定値以上であるかどうかを判断する。決定手段(14)は、判断手段によって適合率が一定値以上であることが判断されたとき(S332で“YES”, S384で“YES”)、撮影画像データ選択手段によって選択された撮影画像データを所望の撮影画像データとして決定する。つまり、テンプレートに基づく理想の画像に少なくとも近似する撮影画像データが選択(取得)される。

10

20

【0009】

この実施例によれば、テンプレートに基づく理想の画像に近似する撮影画像データを選択するだけなので、所望の画像を簡単に取得することができる。

30

【0010】

第2の発明は第1の発明に従属し、撮影画像データ選択手段によって選択された撮影画像データにおいて、適合率を低下させている被写体についての部分画像データであって、適合率を向上させる部分画像データを、当該撮影画像データ選択手段によって選択されていない撮影画像データから抽出する抽出手段、および抽出手段によって抽出された部分画像データを、撮影画像データ選択手段によって選択された撮影画像データの該当する部分に貼り付けて合成した合成画像データを生成する生成手段をさらに備える。

【0011】

第2の発明では、抽出手段(S344)は、撮影画像データ選択手段によって選択された撮影画像データにおいて、適合率を低下させている被写体についての部分画像データであって、適合率を向上させる部分画像データを、当該撮影画像データ選択手段によって選択されていない撮影画像データから抽出する。そして、生成手段(S346)は、抽出手段によって抽出された部分画像データを、撮影画像データ選択手段によって選択された撮影画像データの該当する部分に貼り付けて合成した合成画像データを生成する。

40

【0012】

第2の発明によれば、適合率を低下させている被写体についての部分画像データについては、他の撮影画像データから抽出した適合率の高い部分画像データを貼り付けるので、より理想的な撮影画像データを取得することができる。

【0013】

第3の発明は第1または第2の発明に従属し、適合率を向上させるための指示を出力す

50

る指示手段をさらに備える。

【0014】

第3の発明では、指示手段(S334, S386)は、適合率を向上させるための指示を出力する。たとえば、適合率を向上させるためのメッセージを音声で出力したり、画面にテキスト表示したりする。

【0015】

第3の発明によれば、適合率を向上させる指示を出力するので、次回以降の撮影時にその内容を反映させることができる。

【0016】

第4の発明は第1ないし第3の発明に従属し、被写体情報は、被写体の顔の向きをさらに含み、テンプレートは、理想の顔の向きをさらに含み、適合率は、被写体の顔の向きの理想の顔の向きに対する顔方向適合率を含む。

10

【0017】

第4の発明では、被写体情報は、被写体の顔の向きをさらに含み、同様に、テンプレートは、理想の顔の向きを含む。したがって、被写体の顔の向きの理想の顔の向きに対する顔方向適合率を含む適合率が算出される。

【0018】

第4の発明によれば、視線方向のみならず、顔方向についても理想の画像に近似する画像を取得することができる。

【0019】

20

第5の発明は第1ないし第4の発明に従属し、被写体情報は、被写体の体の向きをさらに含み、テンプレートは、理想の体の向きをさらに含み、適合率は、被写体の体の向きの理想の体の向きに対する体方向適合率を含む。

【0020】

第5の発明においても、第4の発明と同様に、視線方向のみならず、体方向についても理想の画像に近似する画像を取得することができる。さらには、視線方向、顔方向および体方向のすべてが理想の画像に近似する画像を取得することができる。

【0021】

第6の発明は第1ないし第5の発明に従属し、被写体情報は、被写体の表情をさらに含み、テンプレートは、理想の表情をさらに含み、記適合率は、被写体の表情の理想の表情に対する表情適合率を含む。

30

【0022】

第6の発明においても、第4の発明と同様に、視線方向のみならず、被写体の表情についても理想の画像に近似する画像を取得することができる。また、視線方向、体方向および被写体の表情が理想の画像に近似する画像を取得することができる。さらに、視線方向、顔方向および被写体の表情が理想の画像に近似する画像を取得することもできる。さらにまた、視線方向、顔方向、体方向および被写体の表情が理想の画像に近似する画像を取得することもできる。

【0023】

第7の発明は第1ないし第6の発明のいずれかに従属し、被写体情報は、被写体の姿勢をさらに含み、テンプレートは、理想の姿勢をさらに含み、適合率は、被写体の姿勢の理想の姿勢に対する姿勢適合率を含む。

40

【0024】

第7の発明においても、第4の発明と同様に、視線方向のみならず、被写体の姿勢についても理想の画像に近似する画像を取得することができる。かかる場合、視線方向に加えて、顔方向、体方向、被写体の表情および姿勢の少なくとも1つが理想の画像に近似する画像を取得することもできる。

【0025】

第8の発明は第1ないし第7の発明のいずれかに従属し、被写体情報は、被写体の位置をさらに含み、テンプレートは、理想の位置をさらに含み、適合率は、被写体の位置の理

50

想の位置に対する姿勢適合率を含む。

【 0 0 2 6 】

第 8 の発明においても、第 4 の発明と同様に、視線方向のみならず、被写体の位置についても理想の画像に近似する画像を取得することができる。かかる場合、視線方向に加えて、顔方向、体方向、被写体の表情、姿勢および被写体の位置の少なくとも 1 つが理想の画像に近似する画像を取得することもできる。

【 0 0 2 7 】

第 9 の発明は、複数の被写体を含む被写体像の撮影画像に対応する撮影画像データを取得する取得手段、取得手段によって取得された撮影画像データに基づいて、少なくとも複数の被写体の各々の視線方向を含む被写体情報を検出する被写体情報検出手段、少なくとも理想の視線方向の情報を含むテンプレートを複数種類記憶するテンプレート記憶手段、
10 テンプレート記憶手段から所望のテンプレートを選択するテンプレート選択手段、被写体情報が示す視線方向と、テンプレート選択手段によって選択されたテンプレートが示す理想の視線方向とに基づいて、テンプレートに基づく理想の画像に対する撮影画像データの適合率を算出する適合率算出手段、適合率算出手段によって算出された適合率が一定値以上であるかどうかを判断する判断手段、および判断手段によって適合率が一定値以上であることが判断されたとき、取得手段によって取得された撮影画像データを記憶する撮影画像データ記憶手段を備える、画像処理装置である。

【 0 0 2 8 】

第 9 の発明では、画像処理装置 (1 2) の取得手段 (S 5 0 4 , S 5 4 0) は、複数の被写体を含む被写体像の撮影画像に対応する撮影画像データを取得する。視線方向検出手段 (S 5 0 8 , S 5 4 4) は、取得手段によって取得された撮影画像データに基づいて、
20 少なくとも複数の被写体の各々の視線方向を含む被写体情報を検出 (推定) する。テンプレート記憶手段 (1 2 , 3 8) は、少なくとも理想の視線方向の情報を含むテンプレートを複数種類記憶する。たとえば、ユーザの指示に従って、テンプレート選択手段 (S 5 0 0) は、テンプレート記憶手段から所望のテンプレートを選択する。適合率算出手段 (S 5 1 6 , S 5 4 8) は、被写体情報が示す視線方向と、テンプレート選択手段によって選択されたテンプレートが示す理想の視線方向とに基づいて、テンプレートに基づく理想の画像に対する撮影画像データの適合率を算出する。判断手段 (S 5 2 0 , S 5 5 0) は、
30 適合率算出手段によって算出された適合率が一定値以上であるかどうかを判断する。そして、撮影画像データ記憶手段 (S 5 2 2 , S 5 2 4 , S 5 5 2 , S 5 5 4) は、判断手段によって適合率が一定値以上であることが判断されたとき、取得手段によって取得された撮影画像データを記憶する。

【 0 0 2 9 】

第 9 の発明によれば、テンプレートに基づく理想の画像に近似する撮影画像データを取得するので、所望の画像を簡単に取得することができる。

【 発明の効果 】

【 0 0 3 0 】

この発明によれば、テンプレートに基づく理想の画像に近似する撮影画像データを選択するだけなので、所望の画像を簡単に取得することができる。
40

【 0 0 3 1 】

この発明の上述の目的、その他の目的、特徴および利点は、図面を参照して行う以下の実施例の詳細な説明から一層明らかとなろう。

【 発明を実施するための最良の形態 】

【 0 0 3 2 】

図 1 を参照して、この実施例の画像処理システム 1 0 は、デジタルカメラ (以下、単に「カメラ」という。) 1 2 およびサーバ 1 4 を含み、カメラ 1 2 とサーバ 1 4 とは互いに通信可能に接続される。また、サーバ 1 4 には、スピーカ 1 6 が接続される。

【 0 0 3 3 】

カメラ 1 2 は、CPU 2 0 を含み、CPU 2 0 には、バス 2 2 を介してカメラ処理回路
50

24、JPEGコーデック26、距離センサ28、メモリ制御回路30、LCDドライバ32、インターフェイス(I/F)34, 36およびRAM38が接続される。また、カメラ処理回路24には、イメージセンサ40が接続され、このイメージセンサ40には図示しない光学レンズなどを含む光学系42が接続される。さらに、CPU20には、モータドライバ52が接続され、このモータドライバ52は、CPU20の指示の下、光学系42に含まれるフォーカスマータ(図示せず)、ズームモータ(図示せず)または絞りモータ(図示せず)などを駆動する。

【0034】

光学系42を通して入射された光(光学像)は、CCD(Charge Coupled Device)イメージセンサやCMOS(Complementary Metal-Oxide Semiconductor)イメージセンサのようなイメージセンサ40の受光面に照射され、光電変換によって光学像に対応する電気信号が生成される。この電気信号は、カメラ処理回路24で、A/D変換、YUV変換、ガンマ補正、色補正などの各種の処理を施され、これによって、YUV形式に従ったデジタルの画像データが生成される。たとえば、イメージセンサ40からの電気信号は所定のフレームレートで読み出され、各フレームの画像データ(この実施例では、静止画像の画像データを意味する。以下、同じ。)が生成されるのである。

【0035】

撮影モードにおいて、LCD44をビューファインダとして用いてスルー画像(リアルタイム同画像)を表示する場合には、CPU20の指示の下、メモリ制御回路30の制御によって、フレーム毎の画像データは順次SDRAM44に書き込まれ、その後、メモリ制御回路30によってSDRAM44から順次読み出される。そして、SDRAM44から読み出された画像データは、CPU20の指示の下、LCDドライバ32の制御に従ってLCD46に出力される。LCD46の撮影画像表示領域46a(図5参照)には、フレーム毎に更新される画像データが出力され、したがって動画像のスルー画像が表示されるのである。

【0036】

また、撮影モードにおいて、たとえば、オートフォーカス機能がオンにされている場合には、入力装置50に含まれるリリースボタン(シャッターボタン)が半押しされると、距離センサ28によって検出された被写体18とカメラ12との距離に応じて、CPU20はモータドライバ52を制御して、フォーカスマータを駆動し、焦点(距離)を調整する。また、ユーザが、入力装置50に含まれる図示しないズームボタンないしズームスイッチを操作すると、これに応じて、CPU20はモータドライバ52を制御して、ズームモータを駆動し、ズームアウトまたはズームインする。

【0037】

シャッターボタンが半押しされた状態からさらに押し込まれる(全押しされると)、撮影指示がCPU20に与えられ、そのときの画像データが上述のようにして生成され、生成された画像データは、メモリ制御回路30の制御によって、SDRAM44に書き込まれる。次に、CPU20の指示の下、メモリ制御回路30の指示に従ってSDRAM44から画像データが読み出され、読み出された画像データはJPEGコーデック26に与えられる。JPEGコーデック26では、画像データがJPEG方式で圧縮され、圧縮された画像データ(圧縮画像データ)はCPU20の指示に従ってRAM38に記憶される。

【0038】

ただし、この実施例では、カメラ12にSDカードやMMCのようなメモリカード(外部記憶媒体)48が装着されているため、圧縮画像データを、I/F回路34を介してその外部記憶媒体48に記憶するようにしてもよい。または、ユーザ(撮影者)の操作に従って、RAM38に記憶された圧縮画像データを外部記憶媒体48に移動またはコピーしたり、逆に、外部記憶媒体48に記憶された圧縮画像データをRAM38に移動またはコピーしたりすることもできる。

【0039】

なお、図示は省略するが、そのような移動ないしコピーの指示を入力するためのボタン

10

20

30

40

50

ないしスイッチもまた、入力装置 50 に含まれている。

【0040】

さらに、RAM 38 または外部記憶媒体 48 に記憶された圧縮画像データを、I/F 36 を介してカメラ 10 と接続される視線サーバ 14 に送信することもできる。図示は省略するが、このような送信の指示を入力するためのボタンないしスイッチも、入力装置 50 に含まれている。ただし、ボタンないしスイッチの操作によらず、視線サーバ 14 からの要求に応じて、圧縮画像データを送信するようにしてもよい。また、圧縮画像データに対応する画像データを送信することもできるし、スルー画像を表示する際のフレーム毎の画像データを送信することもできる。

【0041】

また、再生モードでは、ユーザは、入力装置 50 に含まれる指示（カーソル）ボタン（図示せず）および再生ボタン（図示せず）を操作することにより、所望の画像（圧縮画像データ）を選択して、再生指示を入力することができる。再生指示が入力されると、CPU 20 は、選択された圧縮画像データを RAM 38 または外部記憶媒体 48 から読み出し、JPEG コーデック 26 に与える。JPEG コーデック 26 は、CPU 20 の指示の下、圧縮画像データを伸張し、その結果得られた画像データは、CPU 20 の指示の下、メモリ制御回路 30 によって SDRAM 44 に書き込まれる。その後、画像データは、CPU 20 の指示の下、メモリ制御回路 30 によって SDRAM 44 から読み出され、LCD ドライバ 32 によって LCD 46 に出力される。

【0042】

なお、撮影モードまたは再生モードを選択するためのボタンないしスイッチも入力装置 50 に含まれており、そのボタンないしスイッチを操作することにより、各モードが設定される。

【0043】

このような画像処理システム 10 は、図 2 に示すように、複数の被写体 18 を含む静止画像（被写体像）180 についての所望の撮影画像を取得するような場合（動画像から所定フレーム毎に静止画像を取得する場合も含む。）に用いられる。たとえば、図示しない撮影者の口頭による指示に従って、カメラ 12 に対する各被写体 18 の位置および各被写体 18 の視線方向が誘導される。撮影者ないしユーザは、視線サーバ 14 に、別途各被写体 18 の理想の視線方向として、当該理想の視線方向についてのパラメータが記述されたテンプレートを選択ないし入力する。たとえば、カメラ 12 では、撮影モードが設定されると、視線サーバ 14 からの要求に応じて、複数枚（この実施例では、10 枚）の写真（撮影画像）を撮影（連写）し、各撮影画像に対応する圧縮画像データ（画像データのままでよい。）をその視線サーバ 14 に送信する。たとえば、1 秒間に 5 枚の撮影画像が得られ、2 秒間撮影が継続され、連写（連続的に撮影）することにより、10 枚の撮影画像が得られる。詳細な説明は省略するが、この実施例では、カメラ 12 は、オートフォーカス時に距離センサ 28 によって計測した距離のデータ（距離データ）が圧縮画像データに付加されて視線サーバ 14 に送信される。ただし、距離センサ 28 によって計測されるのは、カメラ 12 の正面に存在する被写体 18 との直線距離である。

【0044】

視線サーバ 14 は、圧縮画像データを伸張し、その結果得られた画像データから各被写体 18 の視線方向を検出する。これは、すべての画像データ（撮影画像）について実行される。また、視線サーバ 14 は、距離データと画像データとから、撮影時におけるカメラ 12 と視線サーバ 14 との位置関係を算出する。この位置関係の算出に方法については後述するため、ここではその説明は省略する。

【0045】

続いて、視線サーバ 14 は、選択ないし入力されたテンプレートに基づく画像に、最も近似する撮影画像を選択する。そして、選択された撮影画像の中で、被写体 18 の視線方向が、テンプレートが示す理想の視線方向と比較的大きくずれている場合には、他の撮影画像から理想の視線方向とできる限り近似するものを検索し、最も近似する撮影画像の当

10

20

30

40

50

該被写体 18 の顔画像を、選択した撮影画像についての当該被写体 18 の顔画像として合成する。これによって、テンプレートに基づく理想の画像に近い撮影画像（合成画像）を取得することができる。

【 0 0 4 6 】

また、視線サーバ 14 は、合成前の撮影画像において、被写体 18 の視線方向が、テンプレートが示す理想の視線方向と比較的大きくずれている場合には、つまりテンプレートに対する適合率が低い被写体 18 については、その視線（注意）を誘導するための指示を出力する。これは視線方向がずれている被写体 18 毎に行われる。この実施例では、視線サーバ 14 は、そのような誘導を指示するためのメッセージをスピーカ 16 から出力することにより、撮影者または被写体 18 或いはその両方に伝達（フィードバック）する。

10

【 0 0 4 7 】

以下、視線方向の検出方法、テンプレートの生成およびテンプレートとのマッチング、およびフィードバックのそれぞれについて具体的に説明することにする。

< 視線方向の検出方法 >

この実施例では、以下に説明するように、被写体 18 の視線方向（図 3，図 4）を推定または検出するために 1 つのカメラ（単眼カメラ）を用いるユニークな方法を採用している。ただし、視線方向は、図 3 に示す水平方向（地面と水平な方向）の角度 θ と、図 4 に示す垂直方向（地面に対して垂直な方向）の角度 ϕ とで規定される。ただし、点線 $h\ line$ は、水平方向の角度 θ を求めるための基準線であり、カメラ 12 の位置とユーザないし被写体 18 の頭部の中心とを通る直線である。また、点線 $v\ line$ は、垂直方向の角度 ϕ を求めるための基準線であり、点線 $h\ line$ と同様に、カメラ 12 の位置とユーザないし被写体 18 の頭部の中心とを通る直線である。以下、この実施例において、同じである。

20

【 0 0 4 8 】

図 2 に示すように、複数の被写体 18 の前方にカメラ 12 が設置され、このカメラ 12 で撮影された画像（撮影画像）についての圧縮画像データが視線サーバ 14 に取り込まれ、視線サーバ 14 が画像処理することによって、視線 18 A の角度 θ および ϕ を推定する。

【 0 0 4 9 】

なお、図 5 に示すように、カメラ 12 により取得されたスルー画像は、それに附属して設けた LCD 46 の撮影画像表示領域 46 a に動画として表示される。特に限定されないが、たとえば、撮影画像表示領域 46 a 上に、視線方向を示す指標として、眉間から視線方向に延びる線分を表示してもよい。

30

【 0 0 5 0 】

視線サーバ 14 は、一般的な汎用のコンピュータであり、特に変わったハードウェア構成ではないので、ハードウェア自体の説明は省略するが、視線方向や視線位置の推定は、以下に説明するソフトウェアにより実現される。

【 0 0 5 1 】

視線サーバ 14 では、特徴点の追跡処理の安定性を確保するため、同一特徴点に関して異なるフレーム（この実施例では、異なる画像データ）における複数の観測テクスチャを保持している。初期校正過程では、これらの特徴点と虹彩中心の関係から顔特徴点と眼球中心の相対関係を求める。視線推定過程では、校正過程で得られた関係を元に現フレームで得られている特徴点群から眼球中心位置を推定し、その位置と虹彩中心位置から視線方向を決定する。

40

【 0 0 5 2 】

視線方向の推定処理の動作の前提として、まず、たとえば 6 分割矩形フィルタを利用して、顔検出処理が実行される。

【 0 0 5 3 】

視線サーバ 14 では、特に限定されないが、たとえば、顔を連続撮影したビデオ画像（ここでは、静止画像）を処理するにあたり、横が顔幅、縦がその半分程度の大きさの矩形

50

フィルタで画面を走査する。矩形は、たとえば、 3×2 に 6 分割されていて、各分割領域の平均明るさが計算され、それらの相対的な明暗関係が或る条件を満たすとき、その矩形の中心を眉間候補とする。

【 0 0 5 4 】

連続した画素が眉間候補となるときは、それを取囲む枠の中心候補のみを眉間候補として残す。残った眉間候補を標準パターンと比較してテンプレートマッチング等を行うことで、上述した手続きで得られた眉間候補のうちから、偽の眉間候補を捨て、真の眉間を抽出する。以下、さらに詳しく説明する。

【 0 0 5 5 】

図 6 は、眉間候補領域を検出するためのフィルタを説明するための概念図であり、図 6 (a) は、上述した 3×2 に 6 分割された矩形フィルタ (以下、「6 分割矩形フィルタ」と呼ぶ) を示す。6 分割矩形フィルタは、(1) 鼻筋は両目領域よりも明るい、(2) 目領域は頬部よりも暗い、という顔の特徴を抽出し、顔の眉間位置を求めるフィルタである。たとえば、1 点 (x, y) を中心として、横 i 画素、縦 j 画素 (i, j : 自然数) の矩形の枠を設ける。そして、図 6 (a) のように、この矩形の枠を、横に 3 等分、縦に 2 等分して、6 個のブロック $S_1 \sim S_6$ に分割する。

【 0 0 5 6 】

このような 6 分割矩形フィルタを顔画像の両目領域および頬部に当てはめてみると、図 6 (b) のようになる。ただし、図 6 (a) , (b) に示す 6 分割フィルタは各矩形領域が等分されたものであるが、この 6 分割フィルタは図 7 に示すように変形されてもよい。つまり、鼻筋の部分が目の領域よりも通常は狭いことを考慮すると、ブロック S_2 および S_5 の横幅 w_2 は、ブロック S_1, S_3, S_4 および S_6 の横幅 w_1 よりも狭い方がより望ましい。好ましくは、幅 w_2 は幅 w_1 の半分とすることができる。図 7 は、このような場合の 6 分割矩形フィルタの構成を示す。また、ブロック S_1, S_2 および S_3 の縦幅 h_1 と、ブロック S_4, S_5 および S_6 の縦幅 h_2 とは、必ずしも同一である必要もない。

【 0 0 5 7 】

図 7 に示す 6 分割矩形フィルタにおいて、それぞれのブロック S_i ($1 \leq i \leq 6$) について、画素の輝度の平均値「 $\overline{S_i}$ 」(S_i に上付きの “ - ” をつける) を求める。

【 0 0 5 8 】

ブロック S_1 に 1 つの目と眉が存在し、ブロック S_3 に他の目と眉が存在するものとすると、以下の関係式 (1) および (2) が成り立つ。

【 0 0 5 9 】

【数 1】

$$\overline{S_1} < \overline{S_2} \quad \text{and} \quad \overline{S_1} < \overline{S_4}, \quad (1)$$

【 0 0 6 0 】

【数 2】

$$\overline{S_3} < \overline{S_2} \quad \text{and} \quad \overline{S_3} < \overline{S_6} \quad (2)$$

【 0 0 6 1 】

そこで、これらの関係を満たす点を眉間候補 (顔候補) として抽出する。

【 0 0 6 2 】

矩形枠内の画素の総和を求める処理には、公知の文献 (P. Viola and M. Jones, “ Rapid Object Detection using a Boosted Cascade of Simple Features ”, Proc. Of IEEE

10

20

30

40

50

Conf. CVPR, 1, pp.511-518, 2001) において開示されている、インテグラルイメージ (Integral Image) を利用した計算の高速化手法を取り入れることができる。インテグラルイメージを利用することでフィルタの大きさに依らず高速に実行することができる。多重解像度画像に本手法を適用することにより、画像上の顔の大きさが変化した場合にも顔候補の抽出が可能となる。

【 0 0 6 3 】

このようにして得られた眉間候補 (顔候補) に対しては、両目の標準パターンとのテンプレートマッチングにより、真の眉間位置 (真の顔領域) を特定することができる。

【 0 0 6 4 】

ただし、得られた顔候補に対して、サポートベクトルマシン (SVM) による顔モデルによる検証処理を適用し顔領域を決定することもできる。髪型の違いや髭の有無、表情変化による認識率の低下を避けるため、たとえば、図 8 に示すように、眉間を中心とした画像領域を利用して SVM によるモデル化を行うことができる。

【 0 0 6 5 】

なお、このような SVM による真の顔領域の決定については、文献：S. Kawato, N. Te tsutani and K. Hosaka: "Scale-adaptive face detection and tracking in real time with ssr filters and support vector machine", IEICE Trans. on Info. and Sys., E88 - D, 12, pp. 2857 - 2863 (2005) に開示されている。6 分割矩形フィルタによる高速候補抽出と SVM による処理とを組み合わせることで実時間の顔検出が可能である。

【 0 0 6 6 】

続いて、目、鼻や虹彩中心の位置を、公知の文献、たとえば『川戸、内海、安部：「4 つの参照点と 3 枚のキャリブレーション画像に基づく単眼カメラからの視線推定」画像の認識・理解シンポジウム (MIRU2005), pp. 1337 - 1342 (2005)』あるいは、『川戸慎二郎、鉄谷信二：鼻位置の検出とリアルタイム追跡：信学技報 IE2002 - 263, pp. 25 - 29 (2003)』などの手法を用いて抽出する。

【 0 0 6 7 】

両目の位置については、前節の顔領域検出で眉間のパターンを探索しているため、眉間の両側の暗い領域を再探索することにより、大まかな両目の位置を推定することができる。しかし、視線方向の推定のためには、虹彩中心をより正確に抽出する必要がある。ここでは、上で求めた目の周辺領域に対して、ラプラシアンにより虹彩のエッジ候補を抽出し、円のハフ変換を適用することにより、虹彩および虹彩の中心の投影位置を検出する。

【 0 0 6 8 】

鼻の位置は、鼻先が凸曲面であるため周囲に対し明るい点として観測されやすいことと、両目の位置から鼻の存在範囲が限定できることを利用して抽出する。また、両目、鼻の位置を用いて、大体の顔の向きも推定できる。

【 0 0 6 9 】

図 9 (a), (b) は顔検出結果の例を示す図である。検出された顔において、虹彩中心や鼻先や口なども検出されている。たとえば、特徴点としては、鼻先や、左右の目の目尻や目頭、口の両端、鼻腔中心などを用いることができる。

【 0 0 7 0 】

視線の推定においては、視線方向は眼球中心と虹彩中心を結ぶ 3 次元直線として与えられるものとする。

【 0 0 7 1 】

図 10 は視線方向を決定するためのモデルを説明する概念図である。画像上での眼球半径を r 、画像上での眼球中心と虹彩中心との距離を d とすると、視線方向とカメラ光軸とのなす角 θ は次式 (3) で表される。

【 0 0 7 2 】

10

20

30

40

【数 3】

$$\theta = \sin^{-1}\left(\frac{d}{r}\right) \quad (3)$$

【0073】

式(3)により、視線方向を推定するためには、画像上での眼球半径と眼球中心・虹彩中心の投影位置が必要となる。ここで、虹彩中心の投影位置については、上述したとおり、ハフ変換を用いた手法により求めることができる。画像上での眼球直径 r は、解剖学的なモデル(標準的な人の眼球直径)を用いてもよいし、別途キャリブレーションにより求めてもよい。

10

【0074】

図11は、図10に示した状態からユーザがカメラを注視する状態に移行した後の虹彩中心、眼球中心および投影点の関係を示す概念図である。

【0075】

眼球中心の投影位置については、一般には、画像から直接観測することはできない。しかし、被写体18がカメラ12を注視した場合について考えると、図11に示すとおり、カメラ12、虹彩中心、眼球中心の3点が1直線上に並ぶため、画像では虹彩中心と眼球中心は同一点に投影されることが分かる。

20

【0076】

そこで、この実施例での視線推定では、ユーザがカメラ12を注視しながら、顔の姿勢を変化させている画像フレーム列を撮影し、これらの画像列から虹彩位置と顔特徴点を抽出、追跡することにより、眼球中心と顔特徴点間の相対幾何関係を推定する。

【0077】

後により詳しく説明するように、この実施例の視線方向の推定では、眼球中心と顔特徴点間の相対関係の推定処理と眼球中心の投影位置推定とを行う。

【0078】

視線方向の推定のための初期設定として、視線サーバ14は、図12に示すフローチャートで表現されるキャリブレーションを実行する。

30

【0079】

まず、キャリブレーション用の画像列として、ユーザがカメラ12を注視しながら、顔の姿勢を変化させている画像フレーム列を撮影する(ステップS102)。図13は、このようにしてキャリブレーションにおいて撮影された4枚の画像フレームを示す。

【0080】

ここでは、より一般に、 $N(N \geq 2)$ 枚の画像列が得られたとする。各画像フレームを、フレーム I_1, \dots, I_N とする。

【0081】

次に、得られた各画像フレーム列に対して、上述したような方法によって顔検出処理を行い(ステップS104)、続いて、目や鼻の検出処理を行なう(ステップS106)。

40

【0082】

さらに、視線サーバ14は、特徴点の抽出、追跡を行う(ステップS108)。なお、特徴点の抽出方法としては、上述したような方法の他に、たとえば、文献: J. Shi and C. Tomasi: "Good features to track", Proc. CVPR94, pp. 593 - 600 (1994) で提案された手法を用いることもできる。

【0083】

ここで、各画像フレーム $I_i (i = 1, \dots, N)$ において $M (M \geq 4)$ 点の特徴点 $p_j (j = 1, \dots, M)$ が検出・追跡できたとする。画像フレーム I_i における特徴点 p_j の2次元観測位置を $x_j(i)$ (太字) = $[x_j(i), y_j(i)]^t (i = 1, \dots, N$

50

, $j = 1, \dots, M$) とし、両目の虹彩中心の 2 次元観測位置をそれぞれ $x_r(i)$ (太字) $= [x_r(i), y_r(i)]^t$, $x_l(i)$ (太字) $= [x_l(i), y_l(i)]^t$ ($i = 1, \dots, N$) とする。ここで、行列 W を以下のように定義する。

【0084】

【数4】

$$\mathbf{W} = \begin{bmatrix} x_1^{(1)} & \cdots & x_M^{(1)} & x_r^{(1)} & x_l^{(1)} \\ \vdots & & \vdots & & \\ x_1^{(N)} & \cdots & x_M^{(N)} & x_r^{(N)} & x_l^{(N)} \end{bmatrix} \quad (4)$$

10

【0085】

因子分解法により、特徴点の各フレームでの 2 次元観測位置を縦に並べた行列 W (計測行列) は以下のように分解できる。

【0086】

【数5】

$$\mathbf{W} = \begin{bmatrix} m_1^t \\ \vdots \\ m_N^t \end{bmatrix} \begin{bmatrix} s_1 & \cdots & s_M & s_r & s_l \end{bmatrix} \\ = \mathbf{MS} \quad (5)$$

20

【0087】

ここで、行列 M (「撮影姿勢行列」と呼ぶ) にはカメラ 1 2 の姿勢に関する情報のみが、行列 S (「相対位置関係行列」と呼ぶ) には観測対象物の形状に関する情報のみが含まれており、顔特徴点と眼球中心との 3 次元的位置の相対関係は行列 S として求まる (ステップ S110)。すなわち、正射影を仮定すると、行列 M の各要素が画像フレームでのカメラ 1 2 の姿勢を表す単位ベクトルであって、それぞれの大きさが 1 であり相互には直交するとの拘束条件のもとで、行列 W は、特異値分解により一義的に行列 M と行列 S の積に分解できることが知られている。

30

【0088】

なお、このような計測行列 W を、因子分解により、カメラ 1 2 の運動の情報を表す行列と対象物の形状情報を表す行列へ分解する点については、文献：金出，ポールマン，森田：「因子分解法による物体形状とカメラ運動の復元」，電子通信学会論文誌 D - II, J76 D - II, 8, pp.1497 - 1505 (1993) に開示がある。

40

【0089】

図 14 は、リアルタイムの視線方向の推定処理のフローチャートを示す。

【0090】

次に、以上で得られた結果を用いて、視線方向を推定する手順について説明する。

【0091】

まず、カメラ 1 2 から画像フレームを取得すると (ステップ S200)、キャリブレーション時と同様にして、顔の検出および目鼻の検出が行なわれ (ステップ S202)、取得された画像フレーム中の特徴点が抽出される (ステップ S204)。

50

【 0 0 9 2 】

画像フレーム I k が得られたとする。ここで、眼球中心以外の特徴点のうち m 点 p j (j = j 1 , … , j m) が、それぞれ、 $x_j(k)$ (太字) = [$x_j(k)$, $y_j(k)$] t に観測されたとする。このとき、観測された特徴点について、上述したように特徴点近傍のテンプレートを用いたテンプレートマッチングを実施することで、キャリブレーション時に特定された特徴点と現画像フレーム中で観測された特徴点との対応付けが行なわれて、現画像フレーム中の特徴点が特定される (ステップ S 2 0 6)。

【 0 0 9 3 】

なお、上述のとおり、特徴点を特定するためのテンプレートは、キャリブレーションの時のものに限定されず、たとえば、最近の画像フレームの所定枚数について検出された特徴点の近傍の所定の大きさの領域内の画像を所定個数だけ保持しておき、これら所定枚数のテンプレートについてマッチングをした結果、もっとも一致度の高い特徴点に特定することとしてもよい。

10

【 0 0 9 4 】

顔特徴点 p j の 2 次元観測位置 $x_j(k)$ (太字) = [$x_j(k)$, $y_j(k)$] t とキャリブレーションより求めた 3 次元位置 s_j (太字) = [X_j , Y_j , Z_j] t (j = 1 , … , M) の間には、M 個の特徴点のうち観測された m 個の特徴点について注目すると、次式の関係が得られる。

【 0 0 9 5 】

【数 6】

20

$$\begin{bmatrix} x_{j1}^{(k)} \\ \vdots \\ x_{jm}^{(k)} \end{bmatrix} = \mathbf{P}^{(k)} \begin{bmatrix} s_{j1} \\ \vdots \\ s_{jm} \end{bmatrix} \quad (6)$$

【 0 0 9 6 】

ただし、行列 P (k) は 2 × 3 の行列である。右辺の第 2 項の行列 S (k) は行列 S のうち、観測された特徴点に対応する要素のみからなる部分行列である。上述の通り、カメラ 1 2 と顔とは十分に離れているとし正射影を仮定している。ここで、4 点以上の特徴点が観測されれば、行列 P (k) は以下のように計算できる (ステップ S 2 0 8)。

30

【 0 0 9 7 】

【数 7】

$$\mathbf{P}^{(k)} = (\mathbf{X}^{(k)})^t \mathbf{S}^{(k)} \{ (\mathbf{S}^{(k)})^t \mathbf{S}^{(k)} \}^{-1}$$

$$\text{ただし、} \quad \mathbf{X}^{(k)} = \begin{bmatrix} x_{j1}^{(k)} & \cdots & x_{jm}^{(k)} \end{bmatrix}^t \quad (7)$$

40

【 0 0 9 8 】

画像フレーム I k における眼球中心の投影位置 $x_r(i)$ (太字) , $x_l(i)$ (太字) は、行列 P (k) を用いて以下のように計算できる (ステップ S 2 1 0)。

【 0 0 9 9 】

【数 8】

$$x_r^{(k)} = \mathbf{P}^{(k)} s_r \quad (8)$$

【0100】

【数 9】

$$x_l^{(k)} = \mathbf{P}^{(k)} s_l \quad (9)$$

10

【0101】

したがって、画像フレーム I k において特徴点として抽出した虹彩中心の投影位置とこの眼球中心の投影位置を用いると、視線の推定を行なうことができる（ステップ S 2 1 2）。

【0102】

なお、行列 P を QR 分解により分解することで、顔の姿勢 R が、以下のように計算できる。

【0103】

【数 10】

20

$$P = A \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \quad (10)$$

【0104】

【数 11】

$$R = \begin{bmatrix} r_1 \\ r_2 \\ r_1 \times r_2 \end{bmatrix} \quad (11)$$

30

【0105】

ただし r_1 、 r_2 はそれぞれ 1×3 のベクトルである。このような顔の姿勢 R の検出については、文献：L. Quan: "Self-calibration of an affine camera from multiple views", Int'l Journal of Computer Vision, 19, pp. 93 - 105 (1996) に開示がある。

40

【0106】

ユーザ等の指示により追跡が終了していると判断されれば（ステップ S 2 1 4）、処理は終了し、終了が指示されていなければ、処理はステップ S 2 0 2 に復帰する。

【0107】

以上説明した視線方向の推定装置の有効性を確認するため、実画像を用いた実験を行った結果について以下に説明する。

【0108】

カメラ 1 2 は Elmo 社製 PTC - 400C を用い、被写体 1 8 から約 1 5 0 [cm] の位置に設置した。

50

【0109】

まず、50フレームの画像列を用いて、眼球中心と顔特徴点のキャリブレーションを行った。キャリブレーション用の画像フレーム列と抽出した特徴点の例は、図12に示したとおりである。

【0110】

キャリブレーション用画像フレーム列の撮影に要した時間は約3秒であった。(+印は抽出された虹彩中心(眼球中心)、×印は追跡した顔特徴点)。

【0111】

次に、キャリブレーションにより求めた顔モデル(行列S)を用いて、視線推定を行った。ここで、被写体18はそれぞれ右上、上、左下の方向を注視しながら、顔の位置・向きを変化させた。

10

【0112】

図15~図17は、視線推定結果を示す。図15は、右上方注視の状態であり、図16は、上方注視の状態であり、図17は、左下方注視の状態である。ここで、視線方向は両目それぞれで計算された視線方向の平均値としている。結果より、顔の位置や向きの変化とは関係なく、視線方向が推定できた。

【0113】

以上説明したとおり、この実施例の視線方向の推定方法では、単眼カメラの観測に基づいて顔特徴点を検出し、追跡することにより視線方向を推定する。つまり、まずキャリブレーションとして視線がカメラ方向を向いたまま顔の向きのみが異なる画像列から得られる虹彩位置と顔特徴点を利用することで、眼球中心と顔特徴点の関係をモデル化し(行列Sを特定し)、その後、その関係に基づいて推定された入力画像中の眼球中心位置と虹彩位置の関係から視線方向の角度、を決定する。

20

<テンプレートの生成およびテンプレートとのマッチング>

この実施例では、テンプレートは、視線方向についてのパラメータ(角度、)である。視線方向は、上述したように、カメラ12(のレンズの中心)と被写体18の頭部の中心とを通る直線に対して、目線(眼球中心と虹彩中心とを通る線)がなす角度である。ただし、カメラ12と被写体18の頭部を通る直線を比較的太い直線と擬制して、被写体18がカメラ12を見ている場合には、カメラ12と眼球中心と虹彩中心とが一直線上に並び、このときの視線方向を示すパラメータは角度 = = 0度としてある。

30

【0114】

この実施例では、撮影者が所望の撮影画像を得るためのテンプレートを用意ないし生成する(図22参照)。たとえば、カメラ12と被写体18との位置関係および被写体18の視線方向を決定し、テンプレートは生成される。ここで、カメラ12と被写体18との位置関係とは、カメラ12に対する被写体の位置および方向であるが、この実施例では、カメラ12の位置を3次元の現実空間(ワールド座標)における原点Oとした場合における被写体18の位置(3次元座標)を意味する。

【0115】

たとえば、旅行や卒業を記念した写真のような集合写真(撮影画像)を取得する場合には、図18に示すように、各被写体18がカメラ12の方向を向くようなテンプレートが生成される。ここでは、簡単のため、4人の被写体18が一行に並んでいる場合について説明するが、2人以上の被写体18であれば、一行のみならず、二列以上に並んでいてもよい(図19, 図20も同じ)。ただし、カメラ12と被写体18との距離、ここでは、カメラ12と被写体18の列との直線距離は、カメラ12の望遠距離や被写体18の人数に応じて適宜設定される。

40

【0116】

この図18に示すような場合には、各被写体18の視線方向はカメラ12と当該各被写体18の頭部とを通る直線との角度差が無いため、全被写体18に対してパラメータ = = 0であるテンプレートが生成される。また、カメラ12と被写体18との位置関係は、カメラ12の位置をワールド座標の原点とした場合における被写体18の位置座標で決

50

定される。ただし、図 18 に示すように、カメラ 12 の位置を中心として、図面の横方向が X 軸方向（水平方向）であり、図面の縦方向が Z 軸方向（奥行き方向）であり、図面に対して垂直な方向が Y 軸方向（高さ方向）である。また、図面の右方向が X 軸のプラス方向であり、図面の下方向が Z 軸のプラス方向であり、図面に対して垂直上方向が Y 軸のプラス方向である。

【 0 1 1 7 】

また、図 18 においては、簡単のため、3次元の現実空間を上方から俯瞰的に見た状態を示し、水平方向（X軸方向）の角度（ θ ）のみを示し、垂直方向（Y軸方向）の角度については省略してある。さらに、図 18 においては、被写体 18 の頭部を円で示し、その視線を実線と矢印とを用いて示してある。さらにまた、図 18 においては、カメラ 12 と被写体 18 の目とを通る線を破線で示してあるが、視線と重なる部分については実線で示される。また、 x_1, x_2, x_{n-1}, x_n は、各被写体 18 の位置座標のうちの X 軸成分であり、その大きさ（絶対値）は水平方向における原点 O からの距離を示す。これらは、後述する図 19 および図 20 においても同様である。

10

【 0 1 1 8 】

また、図 19 に示すように、すべての被写体 18 が或る方向（ここでは、Z 軸方向と平行な方向）を見ている画像を取得する場合には、X 座標が x_1 の被写体 18 に対しては、 $\theta = \theta_1$ で示されるテンプレートが設定され、X 座標が x_2 の被写体 18 に対しては、 $\theta = \theta_2$ で示されるテンプレートが設定され、X 座標が x_{n-1} の被写体 18 に対しては、 $\theta = \theta_{n-1}$ で示されるテンプレートが設定され、そして、X 座標が x_n の被写体 18 に対しては、 $\theta = \theta_n$ で示されるテンプレートが生成（設定）される。

20

【 0 1 1 9 】

詳細な説明は省略するが、カメラ 12 と被写体 18 との位置関係は予め決定されているため、被写体 18 が見る方向（対象物）を決めれば、三角関数を用いた簡単な計算により、パラメータすなわち角度 θ を求めることができる。また、図 18 の説明においては省略したが、視線方向の角度 θ は、カメラ 12 と被写体 18 の頭部中心とを通る点線から右側に開く角度はプラスで表わされ、点線（*h l i n e*）から左側に開く角度はマイナスで表わされる（図 3 参照）。図示は省略してあるが、視線方向の角度 θ は、カメラ 12 と被写体 18 の頭部中心とを通る点線（*v l i n e*）から上側に開く角度がプラスで表わされ、下側に開く角度がマイナスで表わされる（図 4 参照）。以下、同じである。

30

【 0 1 2 0 】

また、図 20 は、各被写体 18 が特定の同じ方向や特定の同じ物を見るのではなく、それぞれ異なる方向であってその視線方向が放射状に分布する場合には、各被写体 18 の視線 18 A をカメラ 12 とは逆向きに延長し、たとえば、XZ 平面における Z 軸上（カメラ 12 の撮影方向の直線上）の或る点 K を通るように、視線方向のパラメータ θ が決定される。ただし、3次元の世界座標系を見た場合には、各被写体の視線 18 A をカメラ 12 とは逆方向に延長した直線が Z 軸と交わる点 K の Y 軸方向の座標は必ずしも一致しなくてもよい。

【 0 1 2 1 】

図 20 に示す例では、X 座標が x_1 の被写体 18 に対しては、 $\theta = \theta_1$ で示されるテンプレートが設定され、X 座標が x_2 の被写体 18 に対しては、 $\theta = \theta_2$ で示されるテンプレートが設定され、X 座標が x_{n-1} の被写体 18 に対しては、 $\theta = \theta_{n-1}$ で示されるテンプレートが設定され、そして、X 座標が x_n の被写体 18 に対しては、 $\theta = \theta_n$ で示されるテンプレートが生成（設定）される。

40

【 0 1 2 2 】

このように、カメラ 12 と被写体 18 との位置関係および被写体 18 の視線方向を決定した理想の撮影画像を得るためのテンプレートが生成される。この実施例では、テンプレートには、角度 θ のパラメータが記述され、当該テンプレートに対してカメラ 12 に対する被写体 18 の位置座標が付加的に記述される。ただし、被写体 18 の位置座標は、パラメータとして角度 θ とともに、テンプレートに記述するようにしてもよい。

50

【 0 1 2 3 】

図示は省略するが、各被写体 1 8 がそれぞれ別の物（方向）を見るように、視線方向を発散させた撮影画像を得ることも可能である。たとえば、カメラ 1 2 と各被写体 1 8 との位置関係を決定し、各被写者 1 8 が同じ方向や同じ物（カメラ 1 2 を含む。）を見ないように、各被写者 1 8 の視線方向が決定される。

【 0 1 2 4 】

このように、テンプレートは生成され、予め視線サーバ 1 4 に登録され、全体処理（図 2 3 - 図 2 7 参照）、撮影処理（図 2 8 - 図 3 0）および合成処理（図 3 1 , 図 3 2 参照）等において選択（設定）される。視線サーバ 1 4 は、カメラ 1 2 からの圧縮画像データを取得すると、取得した圧縮画像データに対応する画像に含まれる各被写体 1 8 について、カメラ 1 2 との位置関係を算出するとともに、各被写体 1 8 の視線方向を上述した方法によって推定（検出）する。

10

【 0 1 2 5 】

ただし、カメラ 1 2 と各被写体 1 8 との位置関係は、圧縮画像データに付加された距離データと、この圧縮画像データに対する撮影画像とに基づいて算出される。たとえば、図 1 8 - 図 2 0 に示したような例では、カメラ 1 2 と被写体 1 8 の列との距離を距離データから知ることができる。その距離データに基づく距離は、撮影画像の中心或いはその近傍に存在する現実空間の位置（説明の都合上、「第 1 位置」という。）とカメラ 1 2 の位置との距離である。撮影画像から各被写体の顔の大きさを検出し、顔の大きさ（たとえば、横幅または縦の長さ或いはその両方）と距離データに基づく距離とから、第 1 位置との X 軸方向および Z 軸方向の距離を推測する。そして、この推測した距離を用いて、ワールド座標系における位置座標が算出される。

20

【 0 1 2 6 】

なお、この実施例では、顔の大きさに応じて第 1 位置からの距離を求めるようにしてあるが、これに限定される必要はない。たとえば、左右の目の間の距離や眉間の距離を用いるなどして、第 1 位置からの距離を求めることもできる。ただし、顔の大きさ、左右の目の間の距離および眉間の距離のいずれか 2 つ以上に基づいて、第 1 位置からの距離を求めるようにしてもよい。

【 0 1 2 7 】

また、詳細な説明は省略するが、複数の被写体 1 8 を含む被写体像 1 8 0 を撮影する場合には、カメラ 1 2 と各被写体 1 8 との位置関係および各被写体 1 8 の視線方向は、撮影者ないしユーザの指示に従って予め決定される。

30

【 0 1 2 8 】

撮影画像について、カメラ 1 2 と各被写体 1 8 との位置関係および各被写体 1 8 の視線方向が算出（推測）されると、テンプレートに基づく画像におけるカメラ 1 2 と各被写体 1 8 との位置関係および視線方向と比較される。以下、分かり易く説明するために、撮影画像に基づくカメラ 1 2 に対する被写体 1 8 の位置座標を「推定位置関係」といい、同じく撮影画像に基づく被写体の視線方向を「推定視線方向」という。一方、テンプレートに付加された位置座標を「理想位置関係」といい、同じくテンプレートに基づく視線方向を「理想視線方向」という。

40

【 0 1 2 9 】

比較処理（マッチング処理）では、まず、仮想空間に、推定位置関係に従ってカメラ 1 2 と被写体 1 8 とを配置し、被写体 1 8 の推定視線方向に従う視線を描画する。一方、同じく仮想空間に、理想位置関係に従ってカメラ 1 2 と被写体 1 8 とを配置し、被写体の理想視線方向に従う視線を描画する。前者の描画結果と後者の描画結果とから理想視線方向を基準とした場合の推定視線方向のずれを検出する。このずれが小さい程、適合率が高く、逆に、ずれが大きい程、適合率が低いと言える。詳細な説明は省略するが、水平方向および垂直方向の両方も、角度差は絶対値で 0 度 - 9 0 度の間で表わすことができる。したがって、水平方向と垂直方向と角度差の合計が 1 8 0 度の場合の適合率を 0 % とし、角度差の合計が 0 度の場合の適合率を 1 0 0 % とし、角度差が 1 . 8 度増加するに従って

50

適合率を1%ずつ減少させるように、適合率を算出することができる。このような適合率は、各被写体18について算出される。

【0130】

なお、この実施例では、推定位置関係と理想位置関係とは、一致或いはほぼ一致しているものとしてある。

【0131】

このように、比較処理および適合率の算出処理を行うが、上述したように、各被験者の視線方向を発散させた撮影画像の場合には、各被写体18について適合率を算出するのではなく、撮影画像が全体としてのテンプレートに従う画像に適合しているか否かを判断するようにしてある。このように、視線方向を発散させる場合には、他の被験者18と異なる方向を見ていればよく、相対的な方向をきっちり合わせる必要性がないためである。

10

【0132】

具体的には、撮影画像における全被写体18の視線方向の分散および尖度を求め、同様に、テンプレートに基づく画像における全被写体18の視線方向の分散および尖度を求め、その適合度(画像全体としての適合率)を求める。なお、分散および尖度については、周知の公式によって簡単に求めることができ、適合率は、分散および尖度のそれぞれについて、テンプレートに基づく画像についての分散および尖度を基準として、撮影画像についての分散および尖度の百分率を求めればよい。

【0133】

適合率は各撮影画像について算出され、視線方向を発散させた撮影画像以外では、適合率が一定値よりも大きい被写体18の数が最も多い画像が基礎となる画像(基礎画像)として選択される。基礎画像が選択されると、当該基礎画像の被写体18の中で、適合率が低い被写体18については、当該被写体18についての顔画像を、他の適合率の高い撮影画像から切り取って、当該基礎画像の該当する部分に貼り付ける。つまり、合成画像が生成される。このようにすれば、各被写体18について適合率の高い画像を生成することができ、画像全体として理想に近づけることができる。

20

【0134】

詳細な説明は省略するが、合成画像を生成する方法を簡単に説明すると、適合率の高い撮影画像を検出すると、当該撮影画像において各被写体の顔領域を判別する。次に、切り取る顔画像を含む領域を方形または楕円等の簡単な平面図形で指定する。一方、切り取った顔画像を貼り付ける領域についても同じ平面図形で指定する。そして、平面図形同士が重なるように、切り取った顔画像を基礎画像の対応する位置に貼り付ける。このとき、貼り付けた顔画像のエッジ(周辺域)については、最多色を選択して、選択した最多色で埋めて、ガウスぼかしを入れる。これによって、より自然な合成画像を得るようにしてある。

30

【0135】

具体的には、貼り付けた顔画像のエッジ上に、たとえば3×3ピクセルを選択し、ポスタリゼーションして、16色に落とし込む。その中で、最多色を選択し、選択した最多色で他のピクセルを塗りつぶす。この処理をエッジ上でたとえば2ピクセル毎に実行する。その後、ガウスぼかしをかけて、エッジを隠す。

40

【0136】

また、視線方向を発散させた撮影画像では、画像全体としての適合率が最も大きい撮影画像が選択され、上述したような合成処理は行わずに、選択された撮影画像をテンプレートに従う画像に最も近似するものとして取得する。このように、視線方向を発散させた撮影画像では、個々の被写体18についての適合率を算出していない。これは、視線方向を発散させた撮影画像では、各被写体18間で異なる方向を見ていれば良いので、撮影画像全体としての適合率を高くするか否かは個々の視線方向の適合率によって左右されるのではなく、他の被写体18との関係で決まるからである。

<フィードバック>

合成画像を生成する前の基礎画像において、推定視線方向を理想視線方向に近づけた

50

めの指示が被写体 18 または撮影者或いはその両方に与えられる。この実施例では、各被写体 18 に対して、推定視線方向を理想視線方向に近づけるように、その視線 18 A を誘導するためのメッセージをスピーカ 16 から出力するようにしてある。ただし、スピーカ 16 に代えて、ディスプレイのような他の出力装置を用いて、視線 18 A を誘導するようにしてもよい。具体的には、次のような内容のメッセージが通知される。ただし、ここでは、メッセージを通知する被写体 18 を特定していないが、たとえば、メッセージの先頭に、被写体 18 を特定するための内容が追加されるのである。たとえば、8 人の被写体 18 が前列と後列との 2 列に分かれて 4 人ずつ並んでいる場合には、たとえば、「前列の向かって右から 2 番目の人は」のような内容をメッセージの先頭に追加して、被写体 18 を特定可能にするのである。

10

【0137】

たとえば、「少しだけ左を見てください。」や「少しだけ視線を左に向けてください。」のように、方向を示す言葉を含むメッセージが被写体 18 や撮影者に通知（フィードバック）される。また、このようなメッセージの内容は時系列に従って記録され、つまり履歴（指示歴）が残され、次回以降のメッセージを通知する際に、前回通知した内容を反映したメッセージが被写体 18 に通知される。たとえば、先のメッセージによって、左方向に視線を向け過ぎてしまった場合には、「ちょっと向け過ぎましたので、ほんの少し右を見てください。」や「視線を少し右に戻してください。」というメッセージを通知することができる。また、先のメッセージを通知したが、さらに、左方向に視線を移動させたい場合には、「さらにあと少しだけ左を見てください。」や「あともう少し視線を左に向けてください。」というメッセージを通知することができる。

20

【0138】

ただし、このようなメッセージに対応する合成音声のデータ（後述するメッセージデータ）は予め用意されており、選択的にサーバ 12 からスピーカ 18 を介して出力されるのである。また、ディスプレイのような表示器にメッセージを表示する場合には、上述したようなメッセージに対応するテキストデータが記憶されるのである。

【0139】

図 21 は、視線サーバ 14 のメモリ（RAM）のメモリマップの一例を示す図解図である。図 21 に示すように、メモリには、プログラム記憶領域 72 およびデータ記憶領域 74 が設けられる。プログラム記憶領域 72 には、画像処理プログラムが記憶され、この画像処理プログラムは、画像取得プログラム 720、視線推定プログラム 722、比較プログラム 724、適合率算出プログラム 726、画像選択プログラム 728、画像合成プログラム 730 および通知プログラム 732 などによって構成される。

30

【0140】

画像取得プログラム 720 は、カメラ 12 からの画像データ（この実施例では、圧縮画像データ）を取得するためのプログラムである。視線推定プログラム 722 は、画像取得プログラム 720 に従って取得した画像データに対応する撮影画像のそれぞれについて、各被写体 18 の視線方向を推定（検出）するためのプログラムである。比較プログラム 724 は、予め設定（生成）されたテンプレートが示す理想視線方向（パラメータ、）と、視線推定プログラム 722 に従って推定された推定視線方向とを比較するためのプログラムである。

40

【0141】

適合率算出プログラム 726 は、比較プログラム 724 の比較結果に基づいて、推定視線方向が理想視線方向に適合する割合（適合率）を算出するためのプログラムである。ただし、適合率は、撮影画像がテンプレートに基づく画像に適合する割合ということもできる。画像選択プログラム 728 は、カメラ 12 から取得した複数の画像データに対応する複数の撮影画像（静止画像）からテンプレートに基づく画像に最も近似する、1 つの撮影画像すなわち基礎画像を選択するためのプログラムである。

【0142】

画像合成プログラム 730 は、基礎画像における被写体 18 のうち、適合率が一定値よ

50

りも低い被写体 18 を特定し、特定された被写体 18 について、基礎画像以外の撮影画像から適合率が最も高いものを検索し、適合率が最も高い場合被写体 18 の画像（この実施例では、顔画像）を切り切り取って、基礎画像の対応する位置に貼り付けて、合成画像を生成するためのプログラムである。通知プログラム 732 は、基礎画像と指示歴とに基づいて、視線方向などを誘導するための指示を通知（フィードバック）するためのプログラムである。

【0143】

図示は省略するが、プログラム記憶領域 72 には、カメラ 12 などの他の電子機器と通信するための通信プログラムなども記憶される。

【0144】

また、データ記憶領域 74 には、位置関係データ 740、視線データ 742、テンプレートデータ 744、撮影画像データ 746、合成画像データ 748、メッセージデータ 750 および履歴データ 752 が記憶される。

【0145】

位置関係データ 740 は、カメラ 12 の位置を 3 次元の世界座標系の原点 0 とした場合における、各被写体 18 の位置座標のデータ（位置座標データ）を、各被写体 18 に対応づけて記憶したデータである。これによって、被写体 18 の視線方向を決定するための基準となる線（ $h\ line$ 、 $v\ line$ ）を決定することができる。視線データ 742 は、撮影画像における各被写体 18 の推定視線方向を、撮影画像毎に各被写体 18 に対応づけて記憶したデータ（角度データ）である。この角度データは、上述した視線推定プログラム 722 に従って検出される。

【0146】

テンプレートデータ 744 は、複数の理想視線方向についてのパラメータ θ を定義したテンプレートのデータである。たとえば、図 22 に示すように、テンプレートデータ 744 は、第 1 パラメータ 744 a、第 2 パラメータ 744 b、第 3 パラメータ 744 c、...、および第 n パラメータ 744 n についてのデータを含む。各パラメータ 744 a - 744 n は、2 つの視線角度の組で表わされ、図 18 - 図 20 を用いて説明したように、カメラ 12 と被写体 18 との位置関係とともに決定（生成）される。ただし、図 22 において、パラメータ内における上側の視線角度は、水平方向の角度 θ_1 であり、下側の視線角度は垂直方向の角度 θ_2 である。

【0147】

図 21 に戻って、撮影画像データ 746 は、カメラ 12 から取得した撮影画像データである。上述したように、複数枚（この実施例では、10 枚）の撮影画像データがカメラ 12 から視線サーバ 14 に送信され、したがって複数枚の撮影画像についてのデータが記憶される。合成画像データ 748 は、上述した画像選択合成プログラム 730 に従って生成された合成画像についてのデータである。メッセージデータ 750 は、上述したように、通知プログラム 732 に従ってメッセージを被写体 18 等に通知する際に使用する合成音声データである。履歴データ 752 は、通知プログラム 732 に従って通知したメッセージについての履歴（指示歴）を示すデータである。ただし、指示歴は、メッセージの内容（またはメッセージの識別情報）および当該メッセージによって視線方向を指示した被写体 18 の情報（カメラ 12 側から見た並び順など）を含む。

【0148】

具体的には、図 1 に示した視線サーバ 14 が、図 23 ないし図 25 に示すフロー図に従って全体処理を実行する。図 23 に示すように、視線サーバ 14 は、全体処理を開始すると、ステップ S300 で、テンプレートおよびモードが選択されたかどうかを判断する。ここでは、すべての被写体 18 に対して、同じテンプレート（744 a - 744 n など）が選択（設定）されたり、一部または全部に異なるテンプレート（744 a - 744 n など）が選択（設定）されたりしたかどうかを判断するとともに、個別に（被写体 18 毎に）テンプレートとの適合率を判別するモード（個別モード）または総合して（撮影画像全体として）テンプレートに基づく画像との適合率を判別するモード（総合モード）が選択

10

20

30

40

50

されたかどうかを判断するのである。

【0149】

なお、個別モードにおいて、被写体18毎にテンプレートとの適合率を判別した場合であっても、それらの適合率の平均値を求めることにより、撮影画像全体としてテンプレートに基づく画像との適合率を求めることができる。

【0150】

ステップS300で“NO”であれば、つまりテンプレート(744a-744n)およびモード(個別モードまたは総合モード)の少なくとも一方が選択されていなければ、テンプレート(744a-744n)やモード(個別モードまたは総合モード)の選択中であると判断して、同じステップS300に戻る。一方、ステップS300で“YES”

10

【0151】

ステップS302で“NO”であれば、つまり総合モードが選択されれば、図26に示すステップS356に進む。一方、ステップS302で“YES”であれば、つまり個別モードが選択されれば、ステップS304で、初期化処理を実行する。ここでは、変数pおよび変数tに初期値が設定される($p = t = 0$)。ただし、変数pは、カメラ12から取得した撮影画像データに対応する撮影画像の枚数をカウントするための変数である。また、変数tは、画像処理システム10(視線サーバ14)から被写体18等に対してメッ

20

【0152】

続いて、ステップS306では、撮影画像を取得する。たとえば、視線サーバ14は、カメラ12に撮影画像の取得を要求する。すると、カメラ12では、撮影が開始され、所定の枚数(たとえば、10枚)の画像を撮影し、複数枚の撮影画像に対応する撮影画像データを視線サーバ14に送信する。ただし、各撮影画像に対して、または撮影画像データの全体に対して、撮影したときに計測されたカメラ12と被写体18(カメラ12の正面に存在する被写体18)との距離についての数値データ(距離データ)が付加されている。したがって、視線サーバ14は、撮影画像データと距離データとを取得し、データ記憶

30

【0153】

次のステップS308では、1枚目の撮影画像を読み込み、ステップS310で、変数pをインクリメントする($p = p + 1$)。続くステップS312では、全員の顔を検出する。なお、顔の検出方法は、上述した方法と同じである。そして、ステップS314では、全員の視線方向を推定する。ここでは、各被写体18について、図14を用いて説明した視線方向推定処理が実行されるのである。これによって、各被写体18の推定視線方向が得られる。さらに、ステップS316で、撮影画像におけるカメラ12と各被写体18との位置関係を算出する。ここでは、上述したように、撮影画像データと距離データとに基づいて、カメラ12の位置をワールド座標系の原点Oとした場合における各被写体18

40

【0154】

次に、ステップS318では、テンプレートを読み込む。ここでは、全体処理が開始された当初に選択されたテンプレートが読み込まれる。各被写体18に対応してテンプレートが選択されている場合には、それらのすべてが読み込まれる。そして、ステップS320では、読み出したテンプレートに基づく画像における視線方向を算出する。

【0155】

図24に示すように、続くステップS322では、各人の適合率を算出する。ここでは、推定視線方向が、テンプレートに基づく理想視線方向と一致する度合いを、被写体18毎に算出するのである。上述したように、適合率は、撮影画像における被写体18の推定

50

視線方向と、テンプレートに基づく画像の対応する被写体 18 の理想視線方向との角度差に基づいて算出される。

【0156】

続くステップ S 3 2 4 では、撮影画像と各人の適合率とを関連付けて記憶する。つまり、算出した各被写体 18 の適合率を撮影画像に関連付けて、メモリのバッファ領域に一時記憶する。そして、ステップ S 3 2 6 で、変数 p が最大値以上かどうかを判断する。つまり、すべての撮影画像について、各被写体 18 の適合率を算出したかどうかを判断する。この実施例では、変数 p の最大値は「10」である。

【0157】

ステップ S 3 2 6 で“NO”であれば、つまり変数 p が最大値未満であれば、各人の適合率を算出していない撮影画像が残っていると判断して、ステップ S 3 2 8 で、次の撮影画像を読み込み、図 2 3 に示したステップ S 3 1 0 に戻る。このようにして、撮影画像が順番に読み込まれ、撮影画像毎に、各被写体 18 の適合率が算出されるのである。一方、ステップ S 3 2 6 で“YES”であれば、つまり変数 p が最大値以上であれば、つまりすべての撮影画像について、各被写体 18 の適合率を算出すると、ステップ S 3 3 0 で、適合率の高い被写体 18 の人数が最も多い撮影画像を選択する。つまり、基礎画像が選択される。

【0158】

次のステップ S 3 3 2 では、全員の適合率が一定値（たとえば、80%）以上かどうかを判断する。ここで、“YES”なら、基礎画像がテンプレートに基づく理想の画像に少なくとも近似すると決定して、図 2 5 に示すステップ S 3 4 4 にそのまま進む。しかし、“NO”なら、基礎画像がテンプレートに基づく理想の画像に近似していないと決定して、ステップ S 3 3 4 で、適合率が一定値未満の人に、指示歴に合わせた指示を実行する。このとき、視線サーバ 1 4 は、履歴データ 7 5 2 が示す指示歴を参照して、前回の指示が有るか否かを判断し、上述したように、前回の指示が有る場合と無い場合とで、異なる内容のメッセージについての音声合成データを選択して、スピーカ 1 6 を介してメッセージの音声を出力する。つまり、視線方向の変更に関する内容（情報）をフィードバックする。

【0159】

続くステップ S 3 3 6 では、変数 t をインクリメントし ($t = t + 1$)、ステップ S 3 3 8 では、今回の指示を指示歴に記憶する。たとえば、視線サーバ 1 4 は、今回の指示の内容（メッセージ）そのもの、または、使用した音声合成データの識別情報（ファイル名など）を、被写体 18 の情報に対応付けて記憶する。そして、ステップ S 3 4 0 では、変数 t が一定回数以上であるかどうかを判断する。これは、何度指示しても、全員の適合率が一定値を超えず、永久的に所望の撮影画像（合成画像）を取得（生成）できないような不都合を回避するためである。

【0160】

ステップ S 3 4 0 で“NO”であれば、つまり変数 t が一定回数未満であれば、そのままステップ S 3 4 4 に進む。一方、ステップ S 3 4 0 で“YES”であれば、つまり変数 t が一定回数以上であれば、ステップ S 3 4 2 で、一定値を下げてから、ステップ S 3 4 4 に進む。

【0161】

図 2 5 に示すように、ステップ S 3 4 4 では、選択した撮影画像すなわち基礎画像のうち、適合率の低い人（被写体 18）について、適合率の高い撮影画像における当該被写体 18 の顔画像を他の撮影画像から抽出する（切り取る）。次に、ステップ S 3 4 6 で、抽出した顔画像を基礎画像の対応する顔画像に貼り付ける。つまり、合成画像（合成画像データ 7 4 8）が生成される。ただし、ステップ S 3 4 4 および S 3 4 6 の処理は、複数の被写体 18 が該当する場合には、当該複数の被写体 18 について実行される。続いて、ステップ S 3 4 8 で、合成画像を出力する。ここでは、図示は省略するが、合成画像を LCD のようなディスプレイに出力してもよいし、視線サーバ 1 4 に接続される他のコンピュ

10

20

30

40

50

ータ、HDD、データベース、ディスクドライブなどの電子機器に出力するようにしてもよい。つまり、合成画像が表示されたり、保存されたりする。

【0162】

そして、ステップS350では、全体処理の終了であるかどうかを判断する。たとえば、ユーザによって、終了の指示が入力されたかどうかを判断するのである。ステップS350で“NO”であれば、つまり全体処理を終了しない場合には、ステップS352で、位置関係データ740、視線データ742、撮影画像データ746および合成画像データ748をメモリのバッファ領域から削除し、ステップS354で、変数pをリセット($p = 0$)して、図23に示したステップS306に戻る。一方、ステップS350で“YES”であれば、つまり全体処理を終了する場合には、そのまま終了する。

10

【0163】

また、上述したように、総合モードが選択され、図23のステップS302で“NO”と判断すると、図26のステップS356で、テンプレートを読み込む。なお、総合モードにおける処理(S356-S400)のうち、上述した個別モードにおける処理(S304-S354)と同様の処理については簡単に説明することにする。

【0164】

続くステップS358では、テンプレートを用いた視線方向の分散・尖度を算出し、ステップS360では、初期化処理を実行する($p = 0$, $t = 0$)。次のステップS362では、撮影画像を取得する。つまり、カメラ12から所定枚数(この実施例では、10枚)の撮影画像に対応する撮影画像データを取得する。そして、ステップS364で、1枚目の撮影画像を読み込んで、ステップS366で、変数pをインクリメントする($p = p + 1$)。次に、ステップS368では、撮影画像における被写体18全員の顔を検出し、ステップS370では、全員の視線方向を推定する。つまり、各被写体18の推定視線方向が求められる。そして、ステップS372で、撮影画像における視線方向の分散・尖度を算出する。

20

【0165】

次のステップS374では、ステップS372で算出した撮影画像における視線方向の分散・尖度が、ステップS358で算出したテンプレートに基づく画像における視線方向の分散・尖度に適合している度合(適合率)を算出し、ステップS376で、当該撮影画像と、ステップS374で算出した分散・尖度の適合率とを関連付けてメモリのバッファ領域に記憶し、図27に示すステップS378に進む。

30

【0166】

ステップS378では、変数pが最大値以上であるかどうかを判断する。ステップS378で“NO”であれば、つまり変数pが最大値未満であれば、ステップS380で、次の撮影画像を読み込んで、図26に示したステップS366に戻る。一方、ステップS378で“YES”であれば、つまり変数pが最大値以上であれば、ステップS382で、分散・尖度の適合率が最も高い撮影画像を選択する。つまり、カメラ12からの複数枚の撮影画像から、1枚の撮影画像が選択されるのである。

【0167】

1枚の撮影画像が選択されると、ステップS384で、当該撮影画像における分散・尖度の適合率が一定値(たとえば、80%)以上であるかどうかを判断する。ステップS384で“YES”であれば、つまり当該撮影画像における分散・尖度の適合率が一定値以上であれば、所望の撮影画像或いはそれに近い画像が得られたと決定して、そのままステップS394に進む。一方、ステップS384で“NO”であれば、つまり当該撮影画像における分散・尖度の適合率が一定値未満であれば、所望の撮影画像或いはそれに近い画像が得られていないと決定して、ステップS386で、視線方向をばらばらにする旨の指示を実行する。たとえば、視線サーバ14は、「もう少しばらばらの方向を見てください。」というメッセージを、スピーカ16を介して出力する。

40

【0168】

続くステップS388では、変数tをインクリメントし($t = t + 1$)、ステップS3

50

90では、変数tが一定回数以上であるかどうかを判断する。ステップS390で“NO”であれば、つまり変数tが一定回数未満であれば、そのままステップS394に進む。一方、ステップS390で“YES”であれば、つまり変数tが一定回数以上であれば、ステップS392で、一定値を下げて(たとえば、5%~10%下げて)、ステップS394に進む。

【0169】

ステップS394では、選択した撮影画像を出力する。そして、ステップS396では、全体処理を終了するかどうかを判断する。ステップS396で“NO”であれば、つまり全体処理の終了でなければ、ステップS398で、視線データおよび撮影画像データを消去し、ステップS400で、変数pをリセットして($p=0$)、図26に示したステップS362に戻る。一方、ステップS396で“YES”であれば、つまり全体処理の終了であれば、そのまま終了する。

【0170】

このように、テンプレートとモードとを選択すれば、それに従った撮影画像を取得したり、必要に応じて、合成画像を生成したりすることができる。

【0171】

上述の実施例では、カメラ12と視線サーバ14とを個別に設けて、カメラ12からの撮影画像を用いて、視線サーバ14で画像処理(全体処理)を実行するようにしたが、カメラ12のみによって撮影処理と全体処理との両方を実行することも可能である。かかる場合のカメラ12のRAM38のメモリマップは、図21に示した視線サーバ14のメモリのメモリマップと同様であるため、重複した説明は省略する。ただし、画像取得プログラム720に代えて、撮影プログラムが記憶される。また、図23-図27に示した全体処理がカメラ12のCPU20によって実行されるため、図23のステップS306と図26のステップS362とにおいて、撮影処理を実行することにより、所定枚数の撮影画像が取得される。つまり、視線サーバ14を画像処理装置として機能させることができるし、カメラ12を画像処理装置として機能させることもできるのである。

【0172】

また、上述の実施例では、複数枚の撮影画像がテンプレートに従った所望の撮影画像を取得したり、さらに、合成処理を施した画像を取得したりするようにしたが、テンプレートに従った所望の画像を撮影したり、テンプレートに従って撮影処理を実行した複数枚の撮影画像から合成画像を取得(生成)したりすることも可能である。かかる場合、たとえば、図28ないし図30に示すフロー図に従う撮影処理がカメラ12によって実行され、この撮影処理によって得られた複数の撮影画像を用いて、図31および図32に示すフロー図に従う合成処理がカメラ12によって実行される。ただし、合成処理は、視線サーバ14で実行するようにしてもよい。以下、具体的に説明するが、既に説明した処理と同じ処理については、簡単に説明することにする。

【0173】

なお、この実施例では、テンプレートに従って撮影した複数の撮影画像から合成画像を取得するようにしてあるが、テンプレートによらないで撮影した複数枚の撮影画像から所望のテンプレートに合う(たとえば、全体の適合率が80%を超える)撮影画像を検索(選択)することができ、同じ複数の被写体18についての撮影画像であれば、上述したように、さらに合成処理を施すこともできる。

【0174】

図28に示すように、カメラ12(のCPU20)が撮影処理を開始すると、ステップS500で、テンプレートおよびモードが選択されたかどうかを判断する。ステップS500で“NO”であれば、そのままステップS500に戻る。一方、ステップS500で“YES”であれば、ステップS502で、個別モードが選択されたかどうかを判断する。

【0175】

ステップS502で“NO”であれば、つまり総合モードが選択されれば、後述するよ

10

20

30

40

50

うに、図30のステップS536に進む。一方、ステップS502で“YES”であれば、つまり個別モードが選択されれば、ステップS504で、スルー画像から1フレーム分の静止画像を取得する。

【0176】

なお、図示等は省略するが、撮影処理が開始される前に、フォーカスは調整され、このときカメラ12と被写体18との距離は検出(計測)されている。

【0177】

次のステップS506では、取得した静止画像における全員の顔を検出し、ステップS508では、取得した静止画像における全員の視線方向を推定し、そして、ステップS510では、静止画像における被写体の位置関係を算出する。続いて、ステップS512で、テンプレートを読み込み、ステップS514で、テンプレートに基づく画像についての視線方向を算出し、ステップS516で、各人の適合率を算出して、図29に示すステップS518に進む。

【0178】

図29に示すように、ステップS518では、変数 t を初期化する($t = 0$)。次のステップS520では、全員の適合率が一定値以上であるかどうかを判断する。ステップS520で“YES”であれば、つまり全員の適合率が一定値以上であれば、所望の撮影画像またはそれに近い撮影画像を得ることができると判断して、ステップS522で、撮影画像の取得処理を実行し、ステップS524で、取得した撮影画像を保存して、撮影処理を終了する。詳細な説明は省略するが、ステップS522では、上述したように、シャッターボタンが全押しされたときと同じ処理が実行され、圧縮静止画像データがRAM38や外部記憶媒体48に記録される。ただし、ステップS522では、1枚の撮影画像を取得してもよいし、2枚以上の複数枚の撮影画像を取得(連写)するようにしてもよい。

【0179】

また、ステップS520で“NO”であれば、つまり適合率が一定値未満の被写体18が存在する場合には、ステップS526で、指示歴に合わせた指示を実行する。その後、ステップS528で、変数 t をインクリメントし($t = t + 1$)、ステップS530で、今回の指示を指示歴に記録して、ステップS532で、変数 t が一定回数以上かどうかを判断する。ステップS532において、“NO”であれば、そのまま図28に示したステップS504に戻るが、“YES”であれば、ステップS534で、一定値を下げてから、ステップS504に戻る。

【0180】

また、上述したように、総合モードが選択され、ステップS502で“NO”と判断すると、図30に示すステップS536で、テンプレートを読み込み、ステップS538で、テンプレートを用いた視線方向の分散・尖度を算出する。次のステップS540では、スルー画像から静止画像を取得し、ステップS542で、静止画像における全員の顔を検出し、ステップS544で、静止画像における全員の視線方向を推定する。

【0181】

そして、ステップS546で、静止画像における視線方向の分散・尖度を算出し、ステップS548で、分散・尖度の適合率を算出し、そして、ステップS550で、分散・尖度の適合率が一定値以上であるかどうかを判断する。ステップS550で“YES”であれば、つまり分散・尖度の適合率が一定値以上であれば、ステップS554で、撮影画像を取得し、ステップS556で、取得した撮影画像を保存して、撮影処理を終了する。

【0182】

また、ステップS550で“NO”であれば、つまり分散・尖度の適合率が一定値未満であれば、ステップS556で、視線方向をばらばらにする旨の指示を実行し、ステップS558で、変数 t をインクリメントして($t = t + 1$)、ステップS560で、変数 t が一定回数以上であるかどうかを判断する。ステップS560において、“NO”であれば、そのままステップS540に戻るが、“YES”であれば、ステップS562で、一定値を下げてから、ステップS540に戻る。

【0183】

このようにすることにより、通常の写真撮影のように、ビューファインダを覗きながら、被写体18に視線(顔)の向きの指示を出して、所望の撮影画像を取得すると同様に、テンプレートを用いて、自動的に被写体18に視線(顔)の向きの指示を出して、所望の撮影画像を取得することができる。

【0184】

また、図31に示すように、カメラ12(または視線サーバ14)が合成処理を開始すると、ステップS600で、テンプレートが選択されたかどうかを判断する。なお、上述したように、総合モードにおける撮影画像には合成処理は施されないため、ここでは、上述の撮影処理の個別モードにおいて取得された撮影画像を用いた合成処理が実行される。このため、通常、合成処理では、撮影処理において選択されたテンプレートが選択される。したがって、撮影画像に個別モードまたは総合モードのいずれのモードで撮影されたかの識別情報(ラベル)を付与しておくことにより、そのラベルに従って自動的にテンプレートを選択するようにすることもできる。

10

【0185】

なお、視線サーバ14で合成処理を実行する場合には、この合成処理に先だって、個別モードで取得された撮影画像が当該視線サーバ14のメモリに記憶されるのである。

【0186】

ステップS600で“NO”であれば、つまりテンプレートが選択されていない場合は、同じステップS600に戻る。一方、ステップS600で“YES”であれば、つまりテンプレートが選択されれば、ステップS602で、初期化処理を実行する。ここでは、変数pに初期値を設定する($p = 0$)。続く、ステップS604では、1枚目の撮影画像を読み込み、ステップS606では、変数pをインクリメントする($p = p + 1$)。

20

【0187】

続いて、ステップS608では、撮影画像における全員の顔を検出し、ステップS610では、撮影画像における全員の視線方向を推定し、そして、ステップS612では、撮影画像における被写体18の位置関係を算出する。また、次のステップS614では、テンプレートを読み込み、ステップS616では、テンプレートに基づく画像における視線方向を検出する。そして、図32に示すステップS618で、各人の適合率を算出し、次のステップS620で、撮影画像と各人の適合率とを関連付けて記憶する。

30

【0188】

続くステップS622では、変数pが最大値以上であるかどうかを判断する。ステップS622で“NO”であれば、ステップS624で、次の撮影画像を読み込み、図31に示したステップS606に戻る。一方、ステップS622で“YES”であれば、ステップS626で、適合率の高い人数が最も多い撮影画像を基礎画像として選択する。

【0189】

そして、ステップS628では、全員の適合率が一定値以上であるかどうかを判断する。ステップS628で“YES”であれば、そのままステップS632に進む。一方、ステップS628で“NO”であれば、ステップS630で、適合率に応じた指示を実行して、ステップS632に進む。したがって、同じテンプレートを用いて、後で、撮影処理をやり直す場合には、撮影者は、今回の指示に従って被写体18の視線の向きを修正(誘導)することができる。また、今回取得された撮影画像がテンプレートに基づく画像と異なる点を知ることができる。

40

【0190】

なお、合成処理においては、指示歴に基づく指示や指示の回数による一定値の変更を行わないのは、撮影処理と合成処理とが連続的に行われないためであり、リアルタイムに被写体18の視線18Aの方向を誘導しないからである。

【0191】

ステップS632では、選択した撮影画像すなわち基礎画像のうち、適合率の低い人についての顔画像を、適合率の高い他の撮影画像から抽出し、ステップS634では、抽出

50

した顔画像を基礎画像の該当する位置に貼り付けて合成画像を生成し、そして、ステップ S 6 3 6 で、生成した合成画像を出力して、合成処理を終了する。

【 0 1 9 2 】

この実施例によれば、所望の撮影画像を得るためのテンプレートを設定するだけなので、簡単に所望の撮影画像を取得することができる。また、必要に応じて、適合率の高い顔画像を合成するので、より自然に、理想に近い画像を取得することができる。

【 0 1 9 3 】

なお、この実施例では、サーバ 1 4 は、オートフォーカス時に距離センサ 2 8 によって計測した距離データと画像データとから、撮影時におけるカメラ 1 2 と視線サーバ 1 4 との位置関係を算出するようにしたが、他の方法により、距離データを取得することもできる。たとえば、コントラスト検出方式により、オートフォーカスする場合には、CPU 2 0 は、スルー画像から或るフレーム画像を取得し、このフレーム画像のうち、合焦状態を評価するための予め設定された 1 つまたは複数の合焦評価領域に基づいて S D R A M 4 4 から合焦領域の画像成分を抽出する。そして、CPU 2 0 は、合焦評価領域の画像成分に基づいてフレーム画像のコントラスト値を求め、コントラスト値に基づいて合焦状態を判定し、フォーカスレンズの位置（合焦位置）を特定する。さらに、CPU 2 0 は、モータドライバ 5 2 を指示して、フォーカスレンズの位置を制御する。かかる場合には、フォーカスレンズの位置とズームレンズの位置とに対応した距離（カメラ 1 2 と被写体 1 8 との距離）のデータを予め記憶しておくことにより、ズームおよびフォーカスが調整されたときの各レンズの位置に応じて距離データが決定され、これを撮影画像データとともに、サーバ 1 4 に送信するのである。つまり、オートフォーカスの方式は、アクティブ方式とパッシブ方式とのいずれであってもよい。

【 0 1 9 4 】

他の実施例の画像処理システム 1 0 では、視線方向のみならず、他のパラメータも含むテンプレートを設定し、当該テンプレートに従う撮影画像を取得するようにした以下は、上述の実施例と同じであるため、重複した説明は省略する。

【 0 1 9 5 】

この他の実施例では、図 3 3 に示すように、テンプレート 7 4 4 a - 7 4 4 n では、パラメータとして、視線方向（視線角度 θ_1 , ϕ_1 ）に加えて、顔方向（顔の角度 θ_2 , ϕ_2 ）、体方向（体の角度 θ_3 , ϕ_3 ）および表情（顔表情 A_1 , B_1 , C_1 ）が記憶される。したがって、テンプレートの従うことにより、視線方向のみならず、他の要素も考慮された撮影画像を取得することができる。

【 0 1 9 6 】

ただし、顔の角度 θ_2 と体の角度 θ_3 とは、視線角度 θ_1 と同様に、水平方向の角度であり、顔の角度 ϕ_2 と体の角度 ϕ_3 とは、視線角度 ϕ_1 と同様に、垂直方向の角度である。また、顔方向は、頭部の中心と鼻とを通る線上であり、前方に向かう方向である。さらに、体方向は、腹部（体幹）の中心と臍とを通る直線上であり、前方に向かう方向である。

【 0 1 9 7 】

ただし、顔方向については、画像処理により、顔の縦幅に対する横幅の長さ、眉間の長さ、黒目領域の縦幅に対する横幅の長さのいずれか 1 つまたはいずれか 2 つ以上に基いて求めることができる。

【 0 1 9 8 】

また、体方向については、「複数人物によるジェスチャーの単一画像からのスポットィング認識」（<http://ci.nii.ac.jp/naid/110003274571/>）に紹介されている技術に基づいて求めることができる。

【 0 1 9 9 】

さらに、表情については、「動的表情認識による感情推定手法に関する研究」（http://hydro.energy.kyoto-u.ac.jp/Lab/mirai/face/face_OHP/index.htm）に紹介されている技術を用いて求めることができる。

10

20

30

40

50

【0200】

顔方向と体方向については、視線方向と同様にして、撮影画像から得られた顔方向と体方向と、テンプレートに基づく画像から得られた顔方向と体方向とのそれぞれの角度差に基づいて適合率を算出することができる。顔表情については、撮影画像から得られた顔の表情のそれぞれを、対応するテンプレートについての顔の表情と比較し、一致する度合いが高い程、適合率が高く、一致する度合いが低い程、適合率も低くされる。具体的には、数12に従って算出される。

【0201】

[数12]

$$\begin{aligned} \text{適合率}(\%) = & [\{ (100 - | \text{理想表情 A} - \text{推定表情 A} |) \% \\ & + (100 - | \text{理想表情 B} - \text{推定表情 B} |) \% \\ & + \dots \\ & + (100 - | \text{理想表情 N} - \text{推定表情 N} |) \} / n \text{個}] \end{aligned}$$

10

ただし、 $| \cdot |$ は絶対値を意味し、変数 n はテンプレートに設定されている顔の表情についてのパラメータの総数である。また、テンプレートが示すパラメータに基づく顔の表情を「理想表情」と表記し、撮影画像から求められた顔の表情を「推定表情」と表記してある。

【0202】

たとえば、テンプレートの顔表情(理想顔表情A)が「笑顔」でその割合が90%であり、顔表情(理想顔表情B)が「驚き」でその割合が50%であると仮定する。この場合に、撮影画像の或る被写体18について得られた顔表情のうち、「笑顔」の割合が40%で、「驚き」の割合が80%であれば、数12に従うと、適合率は60%と算出される。また、同様の仮定の下で、他の被写体18について得られた顔表情のうち、「笑顔」の割合が80%で「驚き」の割合が30%であれば、数12に従うと、適合率は85%と算出される。

20

【0203】

このように、視線方向のみならず、顔方向、体方向および表情のパラメータを含む場合には、上述した全体処理のステップS312とステップS318との間において、撮影画像に基づいて、顔方向、体方向および表情を算出し、ステップS320において、テンプレートに基づく画像における視線方向、顔方向、体方向および表情を算出し、そして、ステップS322において、視線方向、顔方向、体方向および表情のすべてを含む適合率を被写体18毎に算出するのである。

30

【0204】

同様に、撮影処理では、ステップS506とステップS512との間において、撮影画像に基づいて、顔方向、体方向および表情を算出し、ステップS514において、テンプレートに基づく画像における視線方向、顔方向、体方向および表情を算出し、そして、ステップS516において、視線方向、顔方向、体方向および表情のすべてを含む適合率を被写体18毎に算出するのである。

【0205】

さらに、合成処理では、ステップS608とステップS614との間において、撮影画像に基づいて、顔方向、体方向および表情を算出し、ステップS616において、テンプレートに基づく画像における視線方向、顔方向、体方向および表情を算出し、そして、ステップS618において、視線方向、顔方向、体方向および表情のすべてを含む適合率を被写体18毎に算出するのである。

40

【0206】

なお、他の実施例では、視線方向以外のパラメータとして、顔方向、体方向および表情の3つのパラメータを設けるようにしたが、いずれか1つ以上であってもよい。

【0207】

また、詳細な説明は省略したが、他の実施例では、各パラメータについて算出した適合率の平均値を、被写体18についての適合率としてもよいし、各パラメータに重みを付与

50

して、或るパラメータの適合率を重視するようにしてもよい。

【0208】

さらに、パラメータが体方向を含む場合には、合成画像を生成するとき、被写体の顔画像のみならず、適合率の高い撮影画像から身体全体についての画像を切り取り、その身体全体についての画像を、基礎画像の対応する位置に貼り付ける必要がある。

【0209】

さらにまた、顔方向、体方向、表情のパラメータをさらに含む場合には、フィードバックする内容として、視線方向のみならず、顔方向、体方向、表情についてテンプレートに近づけるようなメッセージを出力することも可能である。

【0210】

このように、顔方向、体方向、表情のパラメータを含むようにすることもできるが、体方向に代えて、または、体方向とともに姿勢（ジェスチャー）のパラメータを含むようにすることもできる。かかる場合には、前傾姿勢、仰け反っている姿勢などの姿勢または予め決めた姿勢についても適合率が算出される。ただし、姿勢は、全身で認識するのみならず、上半身および下半身に分けて認識したり、両手、両足および胴体に分けて認識したりすることもできる。このような姿勢（ジェスチャー）もまた、体方向と同様に、「複数人物によるジェスチャーの単一画像からのスポッティング認識」に紹介されている技術に基づいて求めることができる。

【0211】

また、上述の実施例では、被写体の位置は、撮影者の指示により、テンプレートを作成した際の位置と同じまたはほぼ同じになることを前提としたが、位置のパラメータもテンプレートに含むようにしてもよい。かかる場合には、距離データと画像データとから、カメラ12に対する被写体18の位置（位置座標）が検出され、検出された位置の適合率が、テンプレートに記述された位置のパラメータを用いて算出される。したがって、位置についてもテンプレートに近づけるようなメッセージを出力することができる。

【図面の簡単な説明】

【0212】

【図1】図1はこの発明の画像処理システムの構成の一例を示すブロック図である。

【図2】図2は図1に示す画像処理システムを用いて複数の被写体の撮影画像を取得する状況を説明するための図解図である。

【図3】図3は図1に示すカメラと被写体との位置関係およびその被写体の水平方向の視線角度を示す図解図である。

【図4】図4は図1に示すカメラと被写体との位置関係およびその被写体の垂直方向の視線角度を示す図解図である。

【図5】図5は図1に示すカメラのLCDに表示されている被写体の顔の画像の一例を示す図解図である。

【図6】図6は眉間候補領域を検出するためのフィルタを説明するための概念図である。

【図7】図7は6分割矩形フィルタの他の構成を示す概念図である。

【図8】図8は眉間を中心とした画像領域を利用してSVMによるモデル化を説明する図解図である。

【図9】図9は顔検出結果の例を示す図解図である。

【図10】図10は視線方向を決定するためのモデルを説明する概念図である。

【図11】図11は被写体がカメラを中止する状態に移行した後の虹彩の中心、眼球の中心および投影点の関係を示す概念図である。

【図12】図12は視線サーバによる初期設定処理を示すフロー図である。

【図13】図13はキャリブレーションにおいて撮影された4枚の画像フレームを示す図解図である。

【図14】図14は視線サーバによる視線方向推定処理を示すフロー図である。

【図15】図15は右上方注視の状態での視線推定結果を示す図解図である。

【図16】図16は上方注視の状態での視線推定結果を示す図解図である。

10

20

30

40

50

【図 17】図 17 は左下方注視の状態での視線推定結果を示す図解図である。

【図 18】図 18 はテンプレートを生成するためのカメラと各被写体との位置関係および各被写体の視線方向の一例を示す図解図である。

【図 19】図 19 はテンプレートを生成するためのカメラと各被写体との位置関係および各被写体の視線方向の他の例を示す図解図である。

【図 20】図 20 はテンプレートを生成するためのカメラと各被写体との位置関係および各被写体の視線方向のその他の例を示す図解図である。

【図 21】図 21 は視線サーバのメモリのメモリマップの一例を示す図解図である。

【図 22】図 22 はテンプレートデータに従うテンプレートの例を示す図解図である。

【図 23】図 23 は視線サーバの全体処理の第 1 の部分を示すフロー図である。

【図 24】図 24 は視線サーバの全体処理の第 2 の部分であって、図 23 に後続するフロー図である。

【図 25】図 25 は視線サーバの全体処理の第 3 の部分であって、図 23 および図 24 に後続するフロー図である。

【図 26】図 26 は視線サーバの全体処理の第 4 の部分であって、図 23 ないし図 25 に後続するフロー図である。

【図 27】図 27 は視線サーバの全体処理の第 5 の部分であって、図 26 に後続するフロー図である。

【図 28】図 28 はカメラの撮影処理の一部を示すフロー図である。

【図 29】図 29 はカメラの撮影処理の他の一部であって、図 28 に後続するフロー図である。

【図 30】図 30 はカメラの撮影処理のその他の一部であって、図 28 および図 29 に後続するフロー図である。

【図 31】図 31 はカメラの合成処理の一部を示すフロー図である。

【図 32】図 32 はカメラの合成処理の他の一部を示すフロー図である。

【図 33】図 33 は他の実施例のテンプレートデータに従うテンプレートの例を示す図解図である。

【符号の説明】

【0213】

10 ... 画像処理システム

12 ... カメラ

14 ... 視線サーバ

16 ... スピーカ

20 ... CPU

24 ... カメラ処理回路

26 ... JPEGコーデック

28 ... 距離センサ

30 ... メモリ制御回路

32 ... LCDドライバ

38 ... RAM

40 ... イメージセンサ

42 ... 光学系

44 ... SDRAM

48 ... 外部記憶媒体

50 ... 入力装置

52 ... モータドライバ

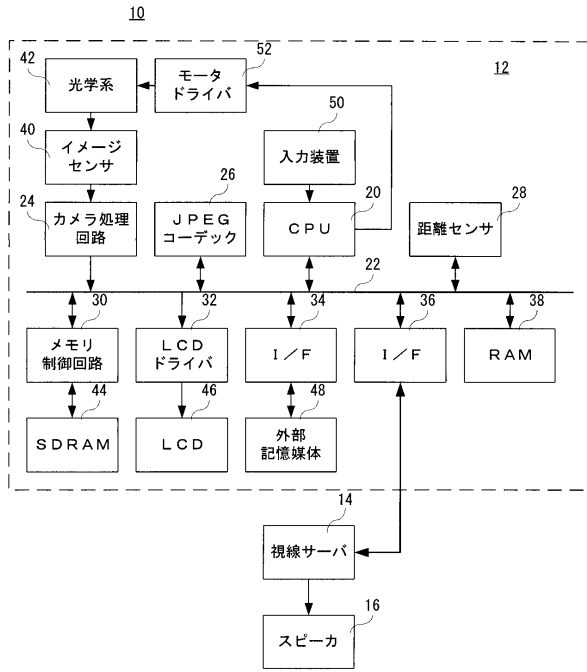
10

20

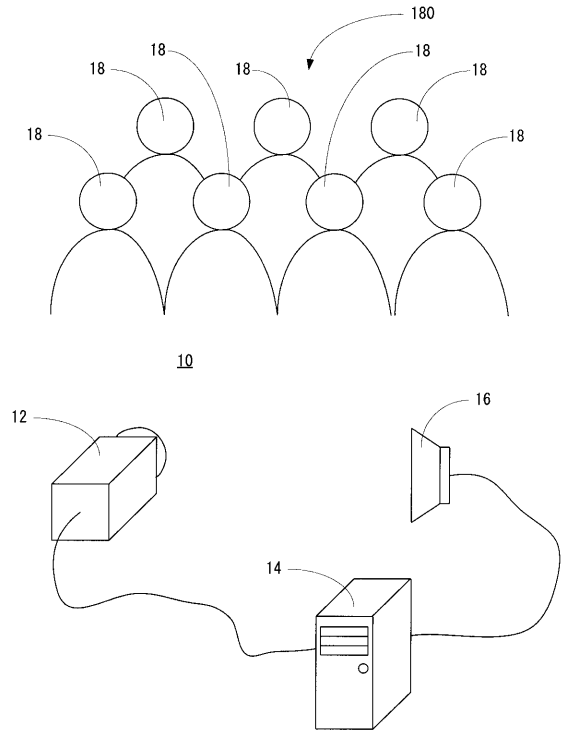
30

40

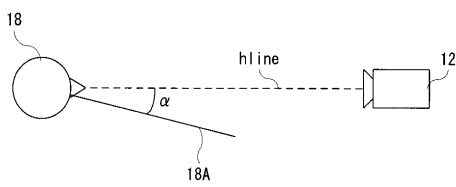
【図1】



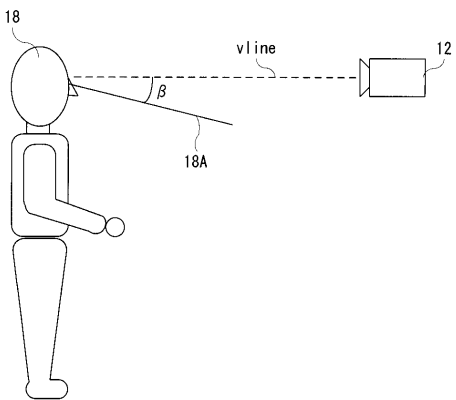
【図2】



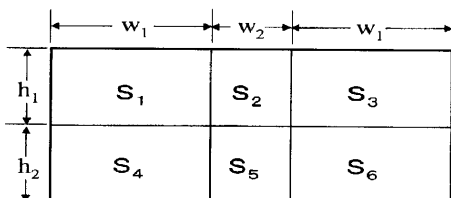
【図3】



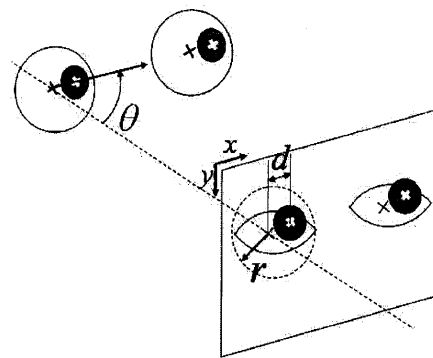
【図4】



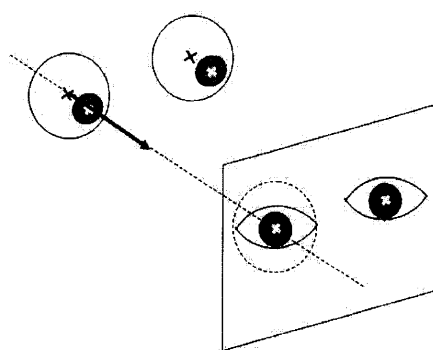
【図7】



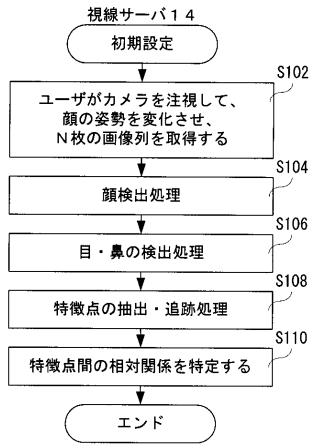
【図10】



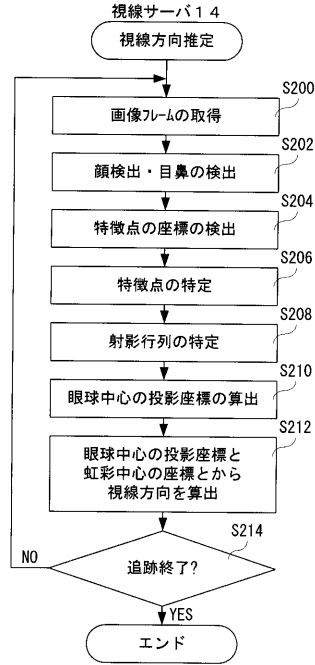
【図11】



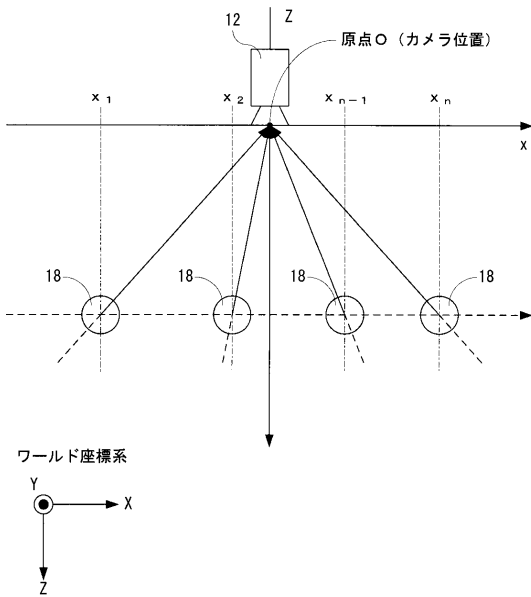
【図12】



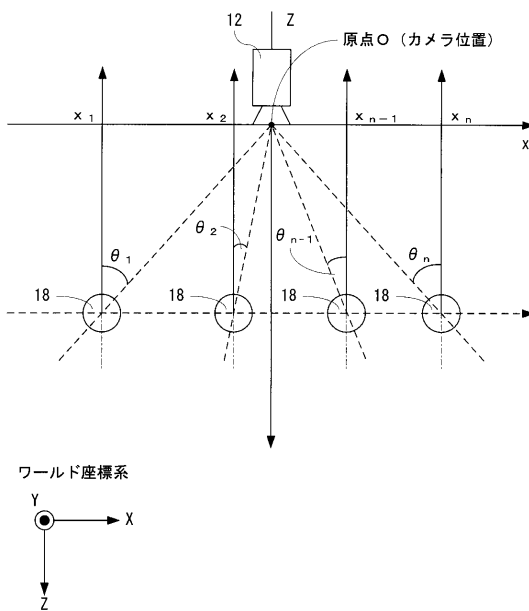
【図14】



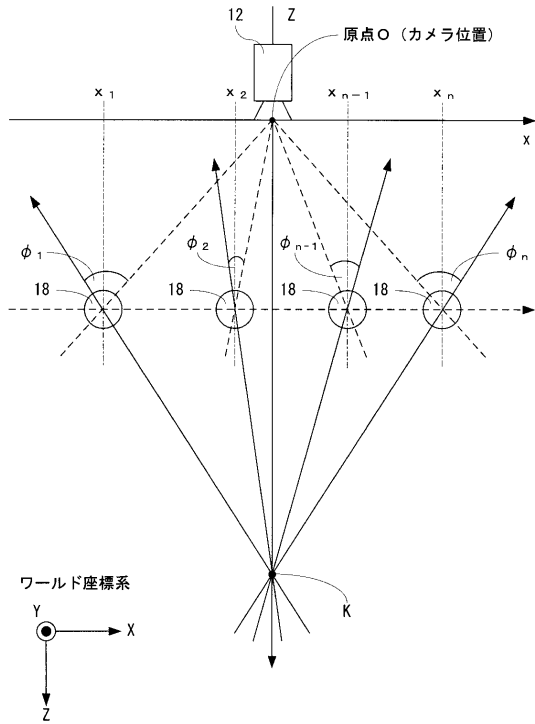
【図18】



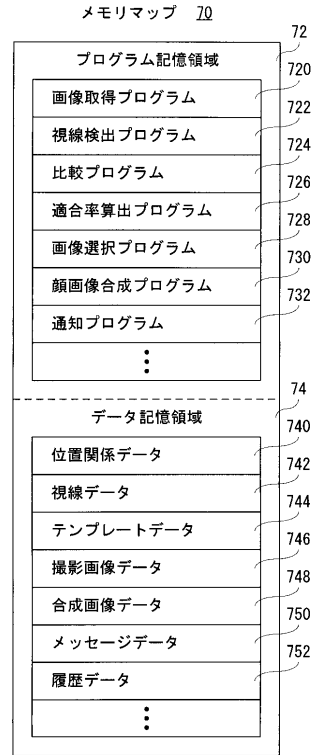
【図19】



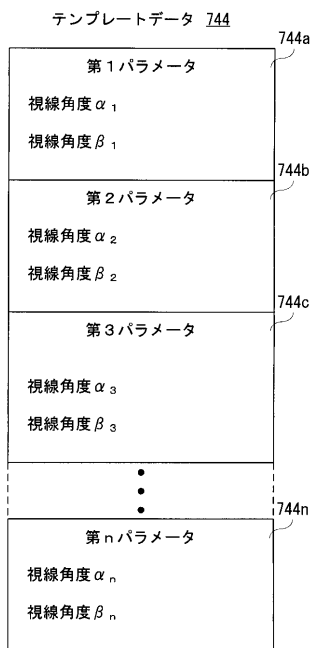
【図20】



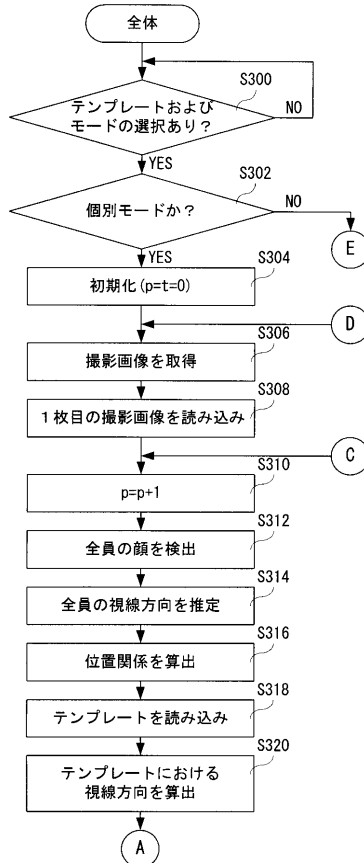
【図21】



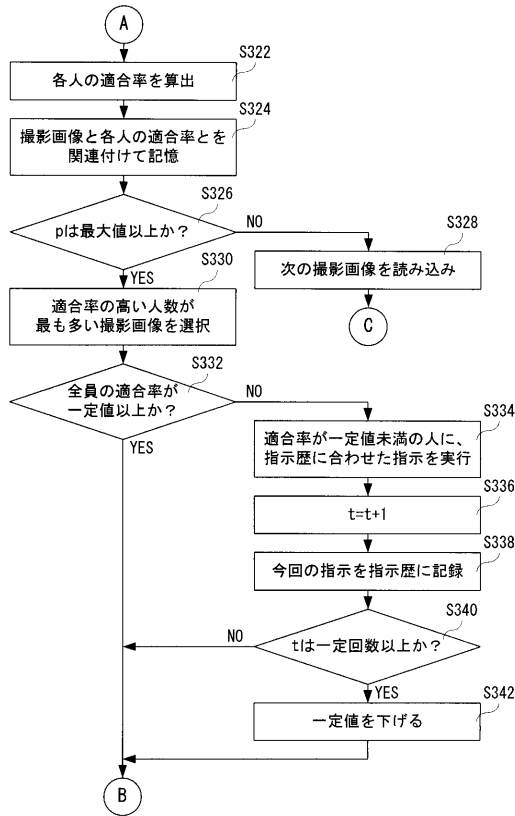
【図22】



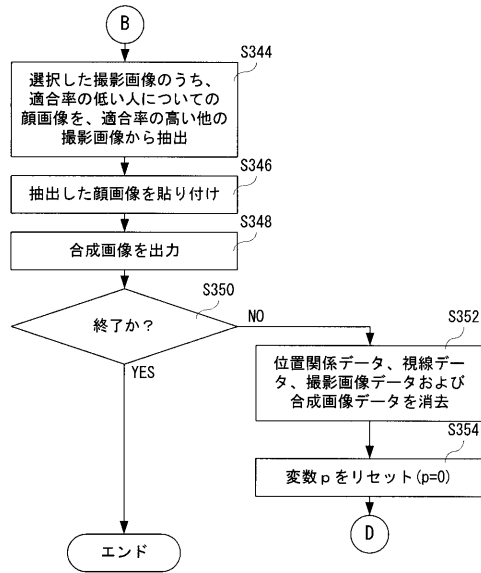
【図23】



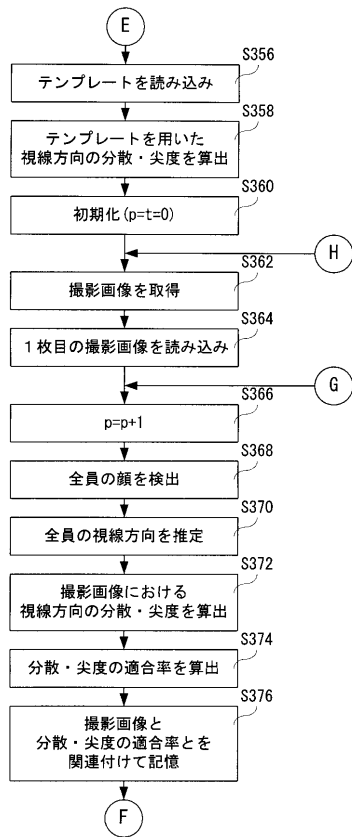
【図 24】



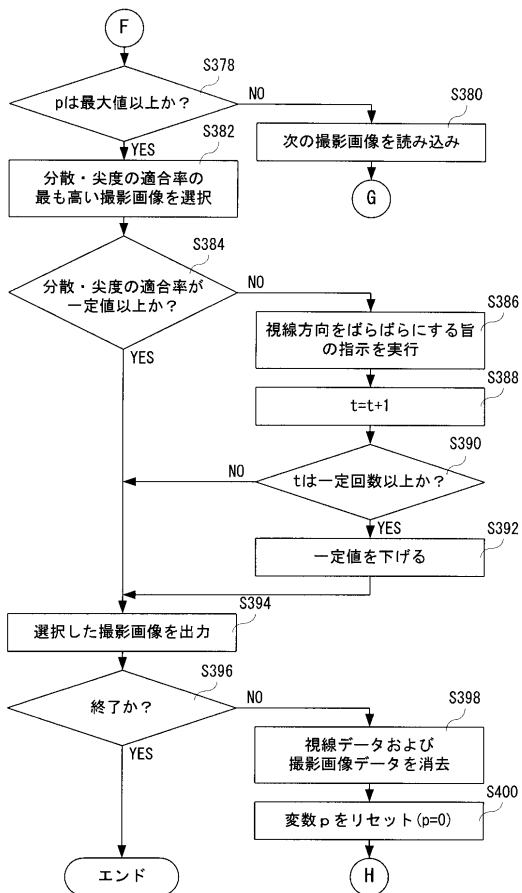
【図 25】



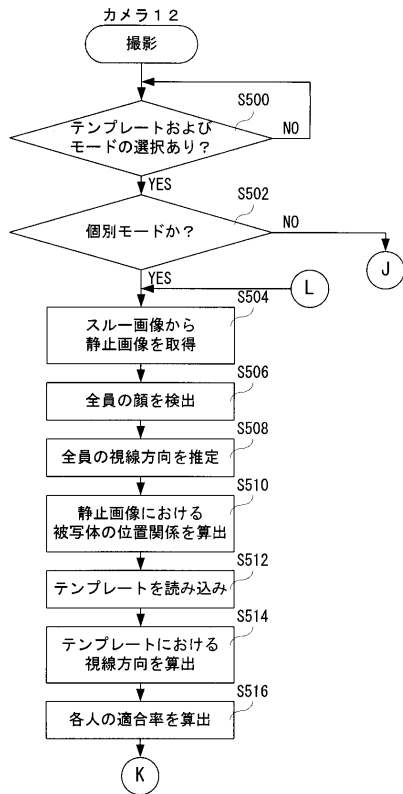
【図 26】



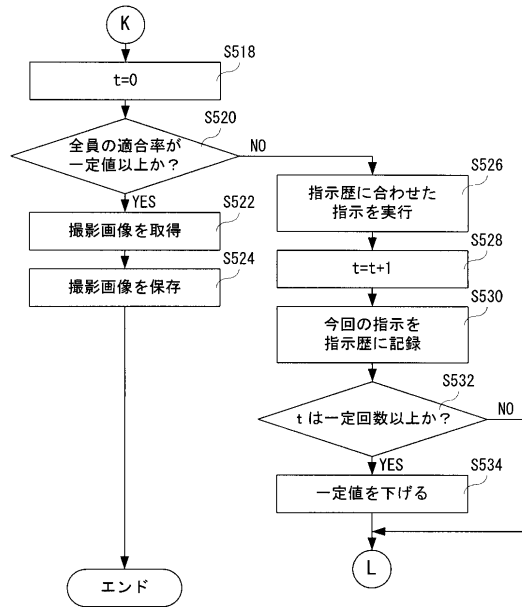
【図 27】



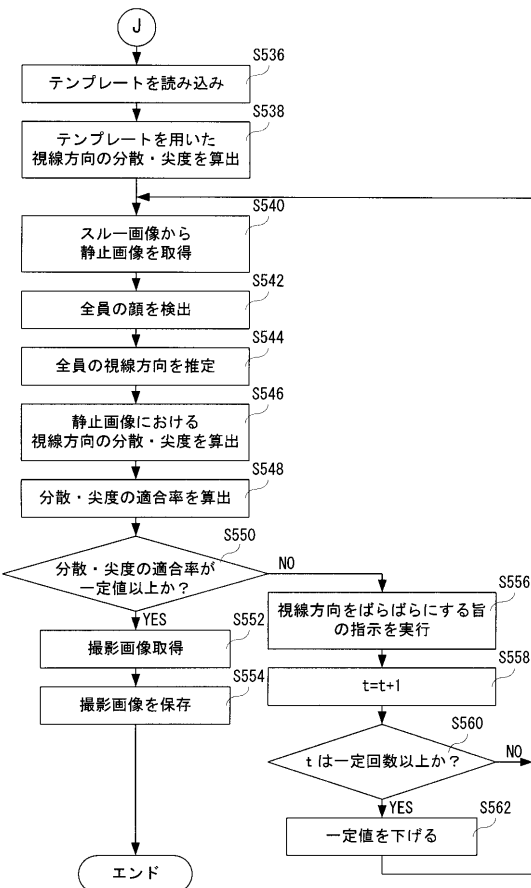
【図 28】



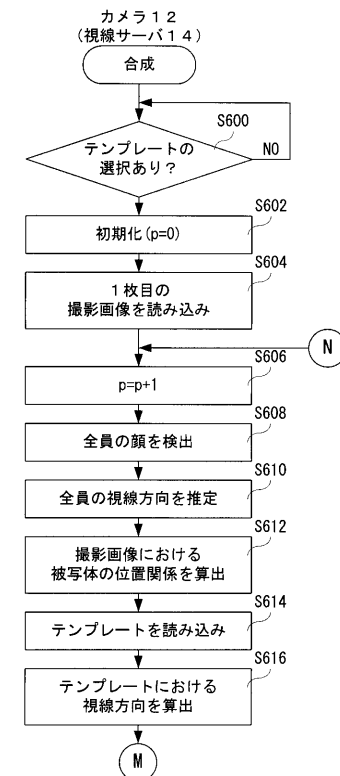
【図 29】



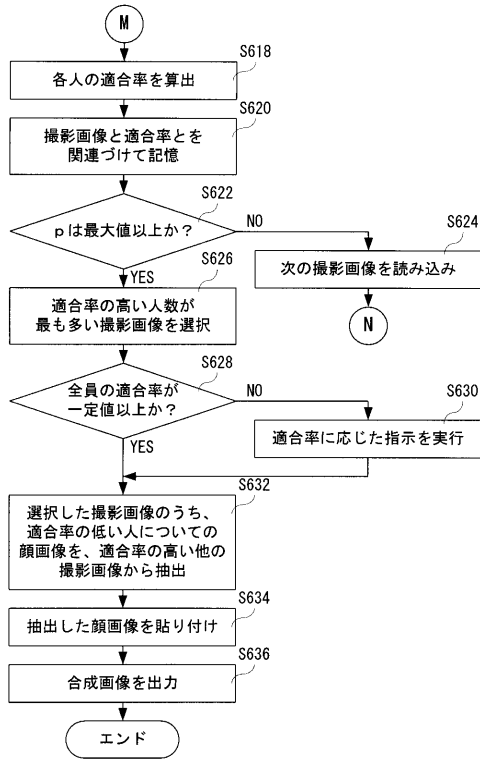
【図 30】



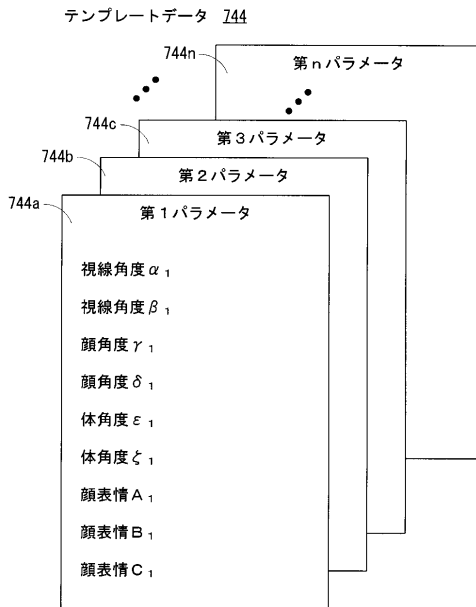
【図 31】



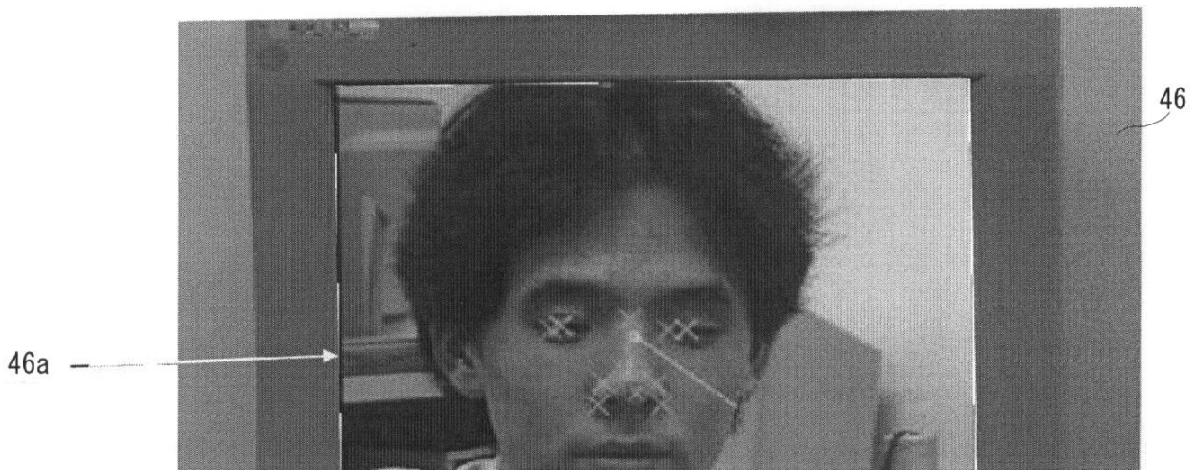
【図32】



【図33】

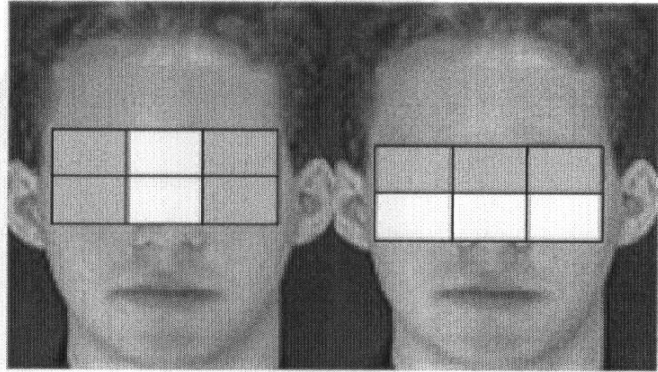


【図5】



【 図 6 】

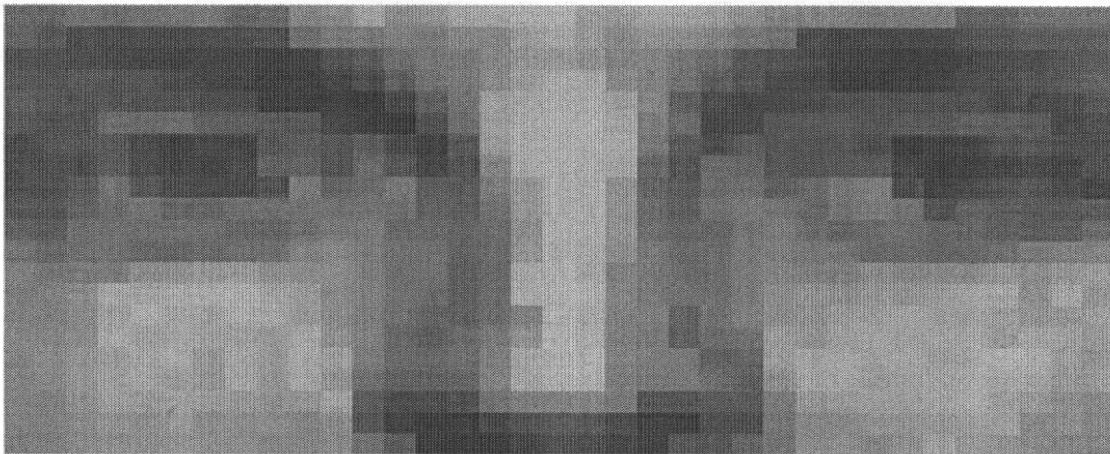
S1	S2	S3
S4	S5	S6



(a)

(b)

【 図 8 】



【 図 9 】



(a)

(b)

【図13】

キャリブレーション



(a)

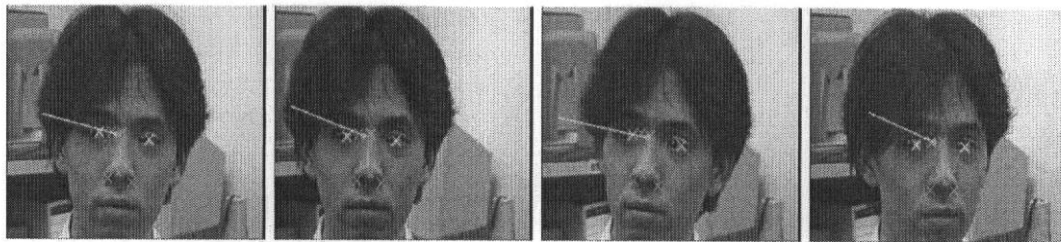
(b)

(c)

(d)

【図15】

右上方注視



(a)

(b)

(c)

(d)

【図16】

上方注視



(a)

(b)

(c)

(d)

【図17】

左下方注視



(a)

(b)

(c)

(d)

フロントページの続き

(72)発明者 安部 伸治

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 板垣 有紀

(56)参考文献 特開2006-319610(JP,A)

特開2007-096440(JP,A)

特開2006-203600(JP,A)

特開2005-020446(JP,A)

特開2007-299297(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/60

G06T 1/00

H04N 1/387

H04N 5/232