

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4100146号  
(P4100146)

(45) 発行日 平成20年6月11日(2008.6.11)

(24) 登録日 平成20年3月28日(2008.3.28)

(51) Int.Cl.		F I			
HO4N	7/15	(2006.01)	HO4N	7/15	630Z
HO4N	5/262	(2006.01)	HO4N	5/262	
HO4N	7/26	(2006.01)	HO4N	7/13	Z

請求項の数 12 (全 25 頁)

(21) 出願番号	特願2002-344164 (P2002-344164)	(73) 特許権者	000002185 ソニー株式会社 東京都港区港南1丁目7番1号
(22) 出願日	平成14年11月27日(2002.11.27)	(74) 代理人	100095957 弁理士 亀谷 美明
(65) 公開番号	特開2004-179997 (P2004-179997A)	(74) 代理人	100096389 弁理士 金本 哲男
(43) 公開日	平成16年6月24日(2004.6.24)	(74) 代理人	100101557 弁理士 萩原 康司
審査請求日	平成17年9月9日(2005.9.9)	(72) 発明者	岡田 良平 東京都品川区北品川6丁目7番35号 ソニー株式会社内
		審査官	長谷川 素直

最終頁に続く

(54) 【発明の名称】 双方向コミュニケーションシステム、映像通信装置

(57) 【特許請求の範囲】

【請求項1】

1 又は 2 以上の映像配信ユニットがネットワークにより接続された双方向コミュニケーションシステムであって：

前記映像配信ユニットは、映像データを生成する撮像装置と；

前記映像データから顔領域を検出し、顔領域情報を生成する特徴検出部、前記顔領域情報に基づき符号化パラメータを生成する符号化制御部、前記符号化パラメータに基づき前記映像データを伝送データに圧縮符号化するエンコーダ部及び前記伝送データを前記映像データに伸長するデコーダ部を少なくとも備える映像通信装置と；

前記映像データを表示する出力装置と；

を備え、

送り手側の前記一の映像配信ユニットは、前記映像データのうち、少なくとも前記顔領域と前記顔領域に属さない領域との各領域ごとに圧縮符号化された前記伝送データを、受け手側の前記他の映像配信ユニットに対して配信し、

前記顔領域情報は、前記顔領域の面積情報、前記顔領域の位置情報又は前記顔領域の信頼度情報の少なくともいずれかを含み、

前記符号化制御部は、現フレーム画像から選択された1の顔領域の信頼度情報が、前記現フレーム画像の他の顔領域の信頼度情報、又は、前フレーム画像の顔領域の信頼度情報に比べて低い場合、前記選択された1の顔領域の信頼度情報を、前記前フレーム画像の顔領域の信頼度情報と同程度、又は、前記現フレーム画像の前記他の顔領域の信頼度情報以

上の値に補正することを特徴とする，双方向コミュニケーションシステム。

【請求項 2】

前記符号化制御部は，前記映像データから前記顔領域情報が生成された場合，当該映像データよりも少なくとも 1 フレーム前に圧縮符号化された前記映像データの前記顔領域情報に基づき，当該映像データの前記顔領域情報を補正することを特徴とする，請求項 1 に記載の双方向コミュニケーションシステム。

【請求項 3】

前記映像通信装置は，前記ネットワークの混雑状況を検知する検査部を，さらに備えることを特徴とする，請求項 1 に記載の双方向コミュニケーションシステム。

【請求項 4】

前記符号化制御部は，前記ネットワークの混雑状況に応じて，前記顔領域にかかる前記符号化パラメータと，前記顔領域に属さない領域にかかる前記符号化パラメータとを変更することを特徴とする，請求項 1 に記載の双方向コミュニケーションシステム。

【請求項 5】

前記符号化制御部は，前記顔領域にかかる映像データを，別オブジェクトとして切り出すことを特徴とする，請求項 1 に記載の双方向コミュニケーションシステム。

【請求項 6】

前記映像通信装置は，前記顔領域にかかる映像データを少なくともモザイク変換する特殊処理部を，さらに備えることを特徴とする，請求項 1 に記載の双方向コミュニケーションシステム。

【請求項 7】

映像データを生成する撮像装置と，前記映像データを表示する出力装置とを備えた映像配信ユニットに備わる映像通信装置であって：

前記撮像装置により生成された前記映像データから顔領域を検出し，顔領域情報を生成する特徴検出部と；

前記顔領域情報に基づき符号化パラメータを生成する符号化制御部と；

前記符号化パラメータに基づき前記映像データを伝送データに圧縮符号化するエンコーダ部と；

前記伝送データを前記映像データに伸長するデコーダ部と；

を備え，

前記顔領域情報は，前記顔領域の面積情報，前記顔領域の位置情報又は前記顔領域の信頼度情報の少なくともいずれかを含み，

前記符号化制御部は，現フレーム画像から選択された 1 の顔領域の信頼度情報が，前記現フレーム画像の他の顔領域の信頼度情報，又は，前フレーム画像の顔領域の信頼度情報に比べて低い場合，前記選択された 1 の顔領域の信頼度情報を，前記前フレーム画像の顔領域の信頼度情報と同程度，又は，前記現フレーム画像の前記他の顔領域の信頼度情報以上の値に補正することを特徴とする，映像通信装置。

【請求項 8】

前記符号化制御部は，前記映像データから前記顔領域情報が生成された場合，当該映像データよりも少なくとも 1 フレーム前に圧縮符号化された前記映像データの前記顔領域情報に基づき，当該映像データの前記顔領域情報を補正することを特徴とする，請求項 7 に記載の映像通信装置。

【請求項 9】

前記映像通信装置は，前記ネットワークの混雑状況を検知する検査部を，さらに備えることを特徴とする，請求項 7 に記載の映像通信装置。

【請求項 10】

前記符号化制御部は，前記ネットワークの混雑状況に応じて，前記顔領域にかかる前記符号化パラメータと，前記顔領域に属さない領域にかかる前記符号化パラメータとを変更することを特徴とする，請求項 7 に記載の映像通信装置。

【請求項 11】

10

20

30

40

50

前記符号化制御部は、前記顔領域にかかる映像データを、別オブジェクトとして切り出すことを特徴とする、請求項7に記載の映像通信装置。

【請求項12】

前記映像通信装置は、前記顔領域にかかる映像データを少なくともモザイク変換する特殊処理部を、さらに備えることを特徴とする、請求項7に記載の映像通信装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、映像配信ユニット間において映像データを送受信可能なシステムにかかり、特に双方向コミュニケーションシステム、映像通信装置、映像データ配信方法に関する。

10

【0002】

【従来の技術】

近年、コンピュータなどの情報処理装置の高機能・低価格化による広範な普及と、デジタル回線を始めとするネットワークのブロードバンド化に伴い、例えばデータ、音声、または映像などをやり取りするマルチメディア通信環境が急速に整備され始めている。

【0003】

マルチメディア通信環境は、代表的な例として、音声及び画像を双方向でやり取りすることによりコミュニケーションを図るテレビ電話/テレビ会議システム(双方向コミュニケーションシステム)などのサービスがある(例えば、特許文献1参照)。なお、本願発明に関連する技術文献情報には、次のものがある。

20

【0004】

【特許文献1】

特開平7-67107号公報

【0005】

【発明が解決しようとする課題】

しかしながら、映像データを伝送する場合において、映像データを圧縮符号化する際、1フレーム全体を画一的に圧縮符号化する場合が多く、帯域に大幅な制限があるネットワークを介して、情報量の多い画像データを送信するには、画一的に全体の画質を下げなければならなかった。

【0006】

また、例えば、フレーム内に人間の顔など、映像を把握するのに不可欠な要素となる注目される特徴を有する領域(特徴領域)に対する映像データを個別に検出しても、的確に検出されない場合が多く、したがって、上記特徴領域についても画質が下がる状態で圧縮符号化され、視認性の低い映像データがネットワークを介して、接続先の相手に表示されていた。

30

【0007】

本発明は、上記のような従来の問題点に鑑みてなされたものであり、特徴を有する領域を的確に判断し、各領域に応じて圧縮符号化を制御することが可能な、新規かつ改良された双方向コミュニケーションシステムを提供することを目的としている。

【0008】

【課題を解決するための手段】

上記課題を解決するため、本発明の第1の観点によれば、1又は2以上の映像配信ユニットがネットワークにより接続された双方向コミュニケーションシステムであって：前記映像配信ユニットは、映像データを生成する撮像装置と；前記映像データから顔領域を検出し、顔領域情報を生成する特徴検出部、前記顔領域情報に基づき符号化パラメータを生成する符号化制御部、前記符号化パラメータに基づき前記映像データを伝送データに圧縮符号化するエンコーダ部及び前記伝送データを前記映像データに伸長するデコーダ部を少なくとも備える映像通信装置と；前記映像データを表示する出力装置と；を備え、送り手側の前記一の映像配信ユニットは、前記映像データのうち、少なくとも前記顔領域と前記顔領域に属さない領域との各領域ごとに圧縮符号化された前記伝送データを、受け手側の

40

50

前記他の映像配信ユニットに対して配信し、前記顔領域情報は、前記顔領域の面積情報、前記顔領域の位置情報又は前記顔領域の信頼度情報の少なくともいずれかを含み、前記符号化制御部は、現フレーム画像から選択された1の顔領域の信頼度情報が、前記現フレーム画像の他の顔領域の信頼度情報、又は、前フレーム画像の顔領域の信頼度情報に比べて低い場合、前記選択された1の顔領域の信頼度情報を、前記前フレーム画像の顔領域の信頼度情報と同程度、又は、前記現フレーム画像の前記他の顔領域の信頼度情報以上の値に補正することを特徴とする、双方向コミュニケーションシステムが提供される。

【0009】

本発明によれば、相互に映像データの送受信可能な映像配信ユニット間において、撮影された映像データのうち、視点が注目される特徴を有する領域（特徴領域）が検出されると、上記特徴領域と、特徴領域以外の領域とに区別し、領域に応じて圧縮符号化する。かかる発明によれば、例えば量子化パラメータが映像データ全体につき一律ではなく、特徴領域に対しては量子化パラメータを小さくし、特徴領域以外の領域に対しては量子化パラメータを大きくして圧縮符号化することにより、領域に応じた差別化を図れる。したがって、映像データのストリーム配信時に、画質が低くてもよい特徴領域以外の領域に対してデータ容量の軽減化、および特徴領域に対して視認性の高い画質の維持された映像データを表示させることができる。

10

【0010】

映像通信装置は、特徴領域情報に基づき、圧縮符号化するために必要なパラメータである符号化パラメータを生成する符号化制御部を、さらに備えるように構成することができる。かかる構成により、映像データを圧縮符号化する際に、例えば映像データのフレーム単位であるフレーム画像のうち、検出された顔領域に対しては量子化パラメータを小さくし画質を向上させ、または顔領域以外の領域に対しては量子化パラメータを大きくし画質を落としデータ量を軽減するように、エンコーダ部に指示するための符号化パラメータを生成することができる。なお、映像データのフレーム単位であるフレーム画像に限定されず、例えば、映像データのフィールド単位であるフィールド画像または複数フレームから構成されるシーン単位であるシーン画像などの場合であってもよい。

20

【0011】

エンコーダ部は、符号化パラメータに基づき映像データを伝送データに圧縮符号化するように構成することができる。かかる発明により、例えば、フレーム画像のうちオブジェクトとして特徴領域を切り出し、顔領域に限り圧縮符号化するように符号化パラメータによって制御されることができる。なお、フレーム画像に限定されず、例えば、フィールド画像またはシーン画像などの場合であってもよい。

30

【0012】

特徴領域情報は、少なくとも顔領域の面積情報、顔領域の位置情報、または顔領域の信頼度情報が含まれる顔領域情報であるように構成することができる。かかる構成により、フレーム画像に構成されるマクロブロックのうち顔領域に属すマクロブロックを、信頼度に基づき的確に特定することが可能となる。なお面積情報は、例えば画素単位に示され、位置情報は、XY座標などにより示される。なお、特徴領域は、顔領域に限定されず、その他特徴を有するいかなる領域であってもよい。

40

【0013】

符号化制御部は、映像データから特徴領域情報が生成された場合、当該映像データより少なくとも1フレーム又は1フィールド前に圧縮符号化された映像データの特徴領域情報に基づき、当該映像データの特徴領域情報を補正するように構成することができる。かかる構成により、フレーム画像内に複数の特徴領域が検出された場合に、検出されたフレーム画像よりも、例えば1フレーム、1フィールド、または1シーンなど前に検出された特徴領域情報に含まれる例えば信頼度などの情報に基づき、上記フレーム画像に関する適正な特徴領域情報に補正することができる。なお、フレーム画像に限定されず、例えば、フィールド画像またはシーン画像などの場合であってもよい。

【0014】

50

映像通信装置は、ネットワークの混雑状況を検知する検査部を、さらに備えるように構成することができる。かかる構成により、ネットワークの混雑状況を把握することで、混雑状況に見合った伝送データ容量に基づきネットワークを介して配信することが可能となる。したがって、ネットワークトラフィックに対して負荷を最小限に留め、通信効率の向上を図れる。

【0015】

符号化制御部は、ネットワークの混雑状況に応じて、特徴領域にかかる符号化パラメータと、特徴領域に属さない領域にかかる符号化パラメータとを変更するように構成することができる。かかる構成により、ネットワークトラフィックが混雑してくると、送信可能なデータ容量が限られてくるため、映像データであるフレーム画像のうち特徴領域のオブジェクトを切出して、上記オブジェクトに対しては高画質の状態での圧縮符号化し、伝送する。特徴領域以外の領域に対しては、圧縮符号化せず削除又は無視される。したがって、映像データの視認の上で不可欠な要素である特徴領域だけを切り出して送信するため、少ないデータ容量で、視認性の高い映像データを配信することができる。なお、混雑状況は、1又は2以上の閾値を段階的に設定しておくことで、混雑状況の段階に応じて、柔軟に画質及びデータ容量を変動させ、配信できる。また、フレーム画像に限定されず、例えば、フィールド画像またはシーン画像などの場合であってもよい。

10

【0016】

符号化制御部は、特徴領域にかかる映像データの符号化パラメータと、特徴領域に属さない領域にかかる映像データの符号化パラメータとを、少なくともフレーム、フィールド、またはシーン単位に変更するように構成してもよい。

20

【0017】

符号化制御部は、特徴領域にかかる映像データを、別オブジェクトとして切り出すように構成してもよい。かかる構成により、フレーム画像の特徴領域に属すマクロブロックに限定して圧縮符号化することができる。さらに、特徴領域に属さないマクロブロックに対して圧縮符号化するか否かを制御することができる。したがって、例えばネットワークのトラフィックなどに応じて柔軟に映像データを圧縮符号化できる。なお、フレーム画像に限定されず、例えば、フィールド画像またはシーン画像などの場合であってもよい。

【0018】

エンコーダ部は、少なくともH.263又はMPEG-4の圧縮符号化方式により、映像データを圧縮符号化するように構成することができる。なお、H.263又はMPEG-4に限定されず、ITU-T勧告H.261などの場合でもよい。

30

【0019】

映像通信装置は、特徴領域にかかる映像データを少なくともモザイク変換する特殊処理部を、さらに備えるように構成することができる。かかる構成により、フレーム画像に検出された特徴領域について、モザイク変換または他の画像に置換などの特殊な処理をすることで、特徴領域を正確に認識できないようにすることができる。なお、フレーム画像に限定されず、例えば、フィールド画像またはシーン画像などの場合であってもよい。さらに、特徴領域以外の領域について、モザイク変換または他の画像に置換などの特殊な処理をする場合でもよい。

40

【0020】

映像データは、少なくとも画像データもしくは音声データのうちのいずれか一方又は双方であるように構成することができる。

【0021】

さらに、本発明の別の観点によれば、映像データを生成する撮像装置と、前記映像データを表示する出力装置とを備えた映像配信ユニットに備わる映像通信装置であって：前記撮像装置により生成された前記映像データから顔領域を検出し、顔領域情報を生成する特徴検出部と；前記顔領域情報に基づき符号化パラメータを生成する符号化制御部と；前記符号化パラメータに基づき前記映像データを伝送データに圧縮符号化するエンコーダ部と；前記伝送データを前記映像データに伸長するデコーダ部と；を備え、前記顔領域情報は

50

、前記顔領域の面積情報、前記顔領域の位置情報又は前記顔領域の信頼度情報の少なくともいずれかを含み、前記符号化制御部は、現フレーム画像から選択された1の顔領域の信頼度情報が、前記現フレーム画像の他の顔領域の信頼度情報、又は、前フレーム画像の顔領域の信頼度情報に比べて低い場合、前記選択された1の顔領域の信頼度情報を、前記前フレーム画像の顔領域の信頼度情報と同程度、又は、前記現フレーム画像の前記他の顔領域の信頼度情報以上の値に補正することを特徴とする、映像通信装置が提供される。

【0022】

本発明によれば、相互に映像データの送受信可能な映像配信ユニット間において、撮影された映像データのうち、視認する上で不可欠な要素である特徴を有する領域（特徴領域）が検出されると、ネットワークの混雑状況を勘案し、上記特徴領域と、特徴領域以外の領域とを区別し、各領域に応じて圧縮符号化する。かかる発明によれば、特徴領域に対しては量子化パラメータを小さくし画質を通常の圧縮符号化時よりも向上させ、特徴領域以外の領域に対しては量子化パラメータを大きくして圧縮符号化することにより、ネットワークに負荷のかからない程度データ容量を軽減しつつ、視認性の高い映像データを配信先の出力装置に表示することができる。なお、この映像通信装置は、上記双方向コミュニケーションシステムで採用される映像通信装置とほぼ同様の構成を有する。

10

【0023】

特徴領域情報は、少なくとも顔領域の面積情報、顔領域の位置情報、または顔領域の信頼度情報が含まれる顔領域情報であるように構成することができる。かかる構成により、フレーム画像に構成されるマクロブロックのうち顔領域に属すマクロブロックを、信頼度に基づき的確に特定することが可能となる。なお面積情報は、例えば画素単位に示され、位置情報は、X Y座標などにより示される。なお、特徴領域は、顔領域に限定されず、その他特徴を有するいかなる領域であってもよい。

20

【0024】

符号化制御部は、映像データから特徴領域情報が生成された場合、当該映像データよりも少なくとも1フレーム前に圧縮符号化された映像データの特徴領域情報に基づき、当該映像データの特徴領域情報を補正するように構成してもよい。

【0025】

映像通信装置は、ネットワークの混雑状況を検知する検査部を、さらに備えるように構成してもよく、符号化制御部は、ネットワークの混雑状況に応じて、特徴領域にかかる符号化パラメータと、特徴領域に属さない領域にかかる符号化パラメータとを変更するように構成してもよい。

30

【0026】

符号化制御部は、特徴領域にかかる映像データの符号化パラメータと、特徴領域に属さない領域にかかる映像データの符号化パラメータとを、少なくとも映像データのフレーム、フィールド、またはシーン単位に変更するように構成してもよい。

【0027】

符号化制御部は、特徴領域にかかる映像データを、別オブジェクトとして切り出すように構成してもよく、エンコーダ部は、少なくともH.263又はMPEG-4の圧縮符号化方式により、映像データを圧縮符号化するように構成してもよい。

40

【0028】

映像通信装置は、特徴領域にかかる映像データを少なくともモザイク変換する特殊処理部を、さらに備えるように構成してもよい。

【0029】

さらに、本発明の別の観点によれば、ネットワークに接続され、少なくとも映像データを生成し、映像データを表示する1又は2以上の映像配信ユニットに備わる映像通信装置の映像データ配信方法が提供される。この映像通信装置の映像データ配信方法において、映像通信装置は、映像データから特徴領域情報を生成し；特徴領域情報に基づき符号化パラメータを生成し；符号化パラメータに基づき映像データを伝送データに圧縮符号化することを特徴としている。

50

## 【 0 0 3 0 】

特徴領域情報は、少なくとも顔領域の面積情報、顔領域の位置情報、または顔領域の信頼度情報が含まれる顔領域情報であるように構成してもよい。

## 【 0 0 3 1 】

映像通信装置は、映像データから特徴領域情報が生成された場合、当該映像データよりも少なくとも1フレーム前に圧縮符号化された映像データの特徴領域情報に基づき、当該映像データの特徴領域情報を補正するように構成してもよい。

## 【 0 0 3 2 】

映像通信装置は、ネットワークの混雑状況を検知する検査部を、さらに備えるように構成してもよく、映像通信装置は、ネットワークの混雑状況に応じて、特徴領域にかかる符号化パラメータと、特徴領域に属さない領域にかかる符号化パラメータとを変更するように構成してもよい。

10

## 【 0 0 3 3 】

映像通信装置は、特徴領域にかかる映像データの符号化パラメータと、特徴領域に属さない領域にかかる映像データの符号化パラメータとを、少なくとも映像データのフレーム、フィールド、またはシーン単位に変更するように構成してもよい。

## 【 0 0 3 4 】

映像通信装置は、特徴領域にかかる映像データを、別オブジェクトとして切り出すように構成してもよく、映像通信装置は、少なくともH.263又はMPEG-4の圧縮符号化方式により、映像データを圧縮符号化するように構成してもよい。

20

## 【 0 0 3 5 】

映像通信装置は、さらに、特徴領域にかかる映像データを少なくともモザイク処理又は他の映像データに置換処理するように構成してもよい。

## 【 0 0 3 6 】

## 【 発明の実施の形態 】

以下、本発明の好適な実施の形態について、添付図面を参照しながら詳細に説明する。なお、以下の説明及び添付図面において、略同一の機能及び構成を有する構成要素については、同一符号を付することにより、重複説明を省略する。なお、本発明にかかる特徴検出部は、例えば、本実施の形態にかかる顔検出ブロック203などに該当する。

## 【 0 0 3 7 】

## ( 1 . システム構成 )

まず、図1を参照しながら、本実施の形態にかかる双方向コミュニケーションシステムについて説明する。図1は、本実施の形態にかかる双方向コミュニケーションシステムの概略的な構成を示すブロック図である。

30

## 【 0 0 3 8 】

図1に示すように、双方向コミュニケーションシステムは、1又は2以上の映像配信ユニット101(a, b, ..., n)がネットワーク105に接続されている。

## 【 0 0 3 9 】

上記映像配信ユニット101(a, b, ..., n)により、使用者106(a, b, ..., n)は、ネットワーク105を介して、お互いの画像又は音声をやりとすることで例えばテレビ会議システムなどのサービスを受けることができる。

40

## 【 0 0 4 0 】

映像配信ユニット101(a, b, ..., n)は、ビデオカメラなどの撮像装置102(a, b, ..., n)と、上記撮像装置102の撮影により生成された、映像データを送受信する映像通信装置104(a, b, ..., n)と、映像データを表示する出力装置103(a, b, ..., n)とが備えられている。なお、本実施の形態にかかる映像データは、少なくとも音声データ又は画像データのうちのいずれか一方又は双方からなる。

## 【 0 0 4 1 】

撮像装置102は、映像データを生成可能なビデオカメラであり、例えば、テレビ会議、監視・モニタリングなどに適用される低ビットレート通信用のビデオカメラであるが、か

50

かる例に限定されず，本実施の形態にかかる撮像装置 102 は，放送用のニュース番組の取材や，スポーツなどの試合の様相などを撮影するカムコーダなどの場合であっても実施可能である。

【0042】

出力装置 103 は，映像データを表示することが可能な例えば，TV 装置又は液晶ディスプレイ装置などが例示され，さらにスピーカを備えることにより，音声および画像を出力することが可能な装置である。

【0043】

映像通信装置 104 は，上記撮像装置 102 により生成された映像データに基づき，使用者 106 の顔である顔領域を検出し，上記顔領域から生成される顔領域情報に基づき，映像データを圧縮符号化し，上記圧縮符号化された伝送データを，ネットワーク 105 を介して送信する。また送信された伝送データを受信し，上記伝送データを伸長する。上記伸長された映像データは，出力装置 103 に送信される。さらに，ネットワーク 105 を介して伝送データを送信する際に，ネットワーク 105 のトラフィックの混雑状況に応じて伝送データを制御する。

10

【0044】

なお，本実施の形態にかかる顔領域に基づく圧縮符号化は，少なくとも H.263，または MPEG-4 に基づき行われるが，後程詳述する。さらに，ネットワーク 105 のトラフィックの混雑状況の検知についても後程詳述する。

【0045】

次に，本システムの典型的な動作例について説明する。

20

【0046】

ある使用者 106 との間で，例えば，使用者 106a と使用者 106b との間で，テレビ会議をする場合，映像配信ユニット 101a に備わる撮像装置 102a により，使用者 106a の映像データが生成され，ネットワーク 105 を介して映像配信ユニット 101b に映像データが送信される。

【0047】

したがって映像配信ユニット 101b に備わる出力装置 103b は，ネットワーク 105 を介して送信された映像データを表示する。また，撮像装置 102b により，使用者 106b の映像データが生成されて，ネットワーク 105 を介して映像配信ユニット 101a に送信され，出力装置 103a に表示される。

30

【0048】

映像配信ユニット 101a と映像配信ユニット 101b との間で，遠隔地であってもネットワーク 105 を介して映像データを送受信することで，お互いの使用者 106a と使用者 106b とのコミュニケーションを図ることができる。

【0049】

なお，本実施の形態にかかる映像配信ユニット 101 には，撮像装置 102，出力装置 103，および映像通信装置 104 とがそれぞれ備わっている場合を例にあげて説明したが，かかる例に限定されず，例えば，1 の映像配信ユニット 101 には，映像通信装置 104 及び出力装置 103 を備え，他の映像配信ユニット 101 には，撮像装置 102 及び映像通信装置 104 を備える場合であっても実施可能である。この場合，例えば，駐車場などに駐車された乗用車又は自動二輪車などのナンバープレート撮像装置 102 により監視する監視システムとしても適用可能である。

40

【0050】

(2 双方向コミュニケーションシステムの各コンポーネントの構成)

次に，本実施の形態にかかる双方向コミュニケーションシステムの各コンポーネントの構成について説明する。

【0051】

(2.1 ネットワーク 105)

ネットワーク 105 は，映像配信ユニット 101 (a, b, ..., n) に備わる映像通信装

50

置 104 ( a , b , ... , n ) を相互に双方向通信可能に接続するものであり , 典型的にはインターネットなどの公衆回線網であるが , WAN , LAN , IP - VPN などの閉鎖回線網も含む。また接続媒体は , FDDI ( Fiber Distributed Data Interface ) などによる光ファイバケーブル , Ethernet (登録商標) による同軸ケーブル又はツイストペアケーブル , もしくは IEEE 802 . 11 b など , 有線無線を問わず , 衛星通信網なども含む。

【 0052 】

( 2 . 2 映像配信ユニット 101 )

映像配信ユニット 101 ( a , b , ... , n ) は , 撮像装置 102 ( a , b , ... , n ) , 上記撮像装置 102 の撮影により生成された映像データを送受信する映像通信装置 104 ( a , b , ... , n ) , もしくは映像データを表示する出力装置 103 ( a , b , ... , n ) のうちいずれか一つ又は任意の組み合わせとが備えられている。

10

【 0053 】

( 2 . 2 . 1 撮像装置 102 )

図 1 に示す撮像装置 102 は , 少なくとも 1 又は 2 以上の撮像素子 ( 撮像デバイス ) が備わる撮像部 ( 図示せず ) と , 音声が入力されるマイク部 ( 図示せず ) と , 映像通信装置 104 に映像入力信号として映像データを出力する出力部 ( 図示せず ) とを備えている。

【 0054 】

上記撮像素子は , 受光面に 2 次的に設けられた光電変換素子からなる複数の画素により , 被写体から受光した光学像を光電変換して画像データとして出力することが可能である。例えば , 撮像素子は , 多種からなる CCD などの固体撮像デバイスが挙げられる。

20

【 0055 】

出力部は , 撮像部により生成された画像データおよびマイク部から生成された音声データに基づき , 映像データを生成し , 映像通信装置 104 に映像入力信号として出力する。

【 0056 】

なお , 本実施の形態にかかる撮像装置 102 に備わる出力部は , 映像データを映像通信装置 104 にアナログデータとして出力するが , かかる例に限定されず , A / D 変換部 ( A / D コンバータ ) を備えることにより , デジタルデータとして出力する場合であっても実施可能である。

【 0057 】

( 2 . 2 . 2 映像通信装置 104 )

次に , 図 2 を参照しながら , 本実施の形態にかかる映像通信装置 104 について説明する。図 2 は , 本実施の形態にかかる映像通信装置の概略的な構成を示すブロック図である。

30

【 0058 】

図 2 に示すように , 映像通信装置 104 は , 撮像装置 102 により送出された映像データを A / D 変換する変換部 201 と , 映像データを一時的に記憶保持するメモリ部 202 と , 映像データに基づき顔領域を検出する顔検出ブロック 203 と , 映像データのうち , 上記顔領域について少なくともモザイク変換又は他の画像に置換する特殊処理部 204 と , 少なくとも顔検出ブロック 203 の検出結果により生成される顔領域情報に基づき符号化パラメータを生成する符号化制御部 205 と , 上記符号化パラメータに基づき映像データを圧縮符号化するエンコーダ部 206 と , 圧縮符号化された伝送データを送受信する通信部 207 と , 通信部 207 により受信された伝送データを伸長するデコーダ部 208 と , 映像データを D / A 変換し , 出力装置 103 に送出する変換部 209 とを備える。なお , 上記顔検出ブロック 203 及び通信部 207 については , 後程詳述する。以下 , 顔領域は , 後程詳述するが , 図 7 に示す顔領域 700 または顔領域 702 である。

40

【 0059 】

( 2 . 2 . 3 出力装置 103 )

出力装置 103 は , 図 2 に示すように , 変換部 209 により D / A 変換された映像データを表示する。また , 出力装置 103 は , 上記説明の通り , 例えば , TV 装置又は液晶ディスプレイ装置などが例示され , 音声又は画像を出力することが可能な装置である。

50

## 【 0 0 6 0 】

なお、本実施の形態にかかる出力装置 1 0 3 は、D / A 変換された映像データを表示する場合を例に挙げて説明したが、かかる例に限定されず、例えば、D / A 変換せずに、デジタルデータのまま映像データを表示する場合でも実施可能である。

## 【 0 0 6 1 】

( 2 . 2 . 4 顔検出ブロック 2 0 3 )

次に、図 2 を参照しながら、メモリ部 2 0 2 に記憶された映像データに含まれる顔領域を検出する顔検出ブロック 2 0 3 及び顔領域検出処理について説明する。

## 【 0 0 6 2 】

顔検出ブロック 2 0 3 は、メモリ部 2 0 2 に記憶された映像データをフレーム単位に、映像データから人間の顔画像である顔領域を検出する。したがって、顔検出ブロック 2 0 3 には、複数の工程により上記顔領域を検出するために、各部がそれぞれ備わっている。

10

## 【 0 0 6 3 】

なお、本実施の形態にかかる顔検出ブロック 2 0 3 は、人間の顔領域を検出する場合を例に挙げて説明したが、映像データのうち特徴的な領域を有する場合であれば、かかる例に限定されず、例えば、乗用車のナンバープレート、時計、またはパソコンなどの画像領域を検出する場合であっても実施可能である。

## 【 0 0 6 4 】

顔検出ブロック 2 0 3 は、図 2 に示すように、リサイズ部 2 3 0 と、ウィンドウ切出部 2 3 1 と、テンプレートマッチング部 2 3 2 と、前処理部 2 3 3 と、SVM ( サポートベクタマシン; Support Vector Machine ) 識別部 2 3 4 と、結果判定部 2 3 5 とが備わる。

20

## 【 0 0 6 5 】

リサイズ部 2 3 0 は、撮像装置 1 0 2 により生成された映像データを、メモリ部 2 0 2 からフレーム単位に読み出して、当該フレーム単位に読み出された映像データ ( 以下、フレーム画像 ) を縮小率が相異なる複数のスケール画像に変換する。

## 【 0 0 6 6 】

例えば、本実施の形態にかかるフレーム画像が、NTSC方式 ( National Television System Committee 方式 ) の 7 0 4 × 4 8 0 画素 ( 横 × 縦 ) からなる場合、0 . 8 倍ずつ順次縮小して 5 段階 ( 1 . 0 倍、0 . 8 倍、0 . 6 4 倍、0 . 5 1 倍、0 . 4 1 倍 ) のスケール画像に変換する。なお以下、上記複数のスケール画像は、1 . 0 倍のスケール画像を第 1 のスケール画像とし、順次縮小するごとに、第 2 ~ 第 5 のスケール画像とする。

30

## 【 0 0 6 7 】

ウィンドウ切出部 2 3 1 は、まず第 1 のスケール画像に対して、画像左上を起点として順にフレーム画像の右下まで、例えば 2 画素ずつなど、適当な画素ずつ右側又は下側にずらしながらスキャンするようにして、2 0 × 2 0 画素の矩形領域 ( 以下、ウィンドウ画像と呼ぶ ) を順次切出す。なお、本実施の形態にかかるスケール画像の起点は、画像左上である場合に限らず、例えば画像右上などであっても実施可能である。

## 【 0 0 6 8 】

上記第 1 のスケール画像から切出された複数のウィンドウ画像は、順次、ウィンドウ切出部 2 3 1 により、後続のテンプレートマッチング部 2 3 2 に送出される。

40

## 【 0 0 6 9 】

テンプレートマッチング部 2 3 2 は、ウィンドウ切出部 2 3 1 により送出されたウィンドウ画像について、例えば正規化相関法、誤差二乗法などの演算処理を実行してピーク値をもつ関数曲線に変換した後、当該関数曲線に対して認識性能が落ちない程度に十分に低い閾値を設定し、当該閾値を基準として当該ウィンドウ画像の領域が顔領域であるか否かを判断する。

## 【 0 0 7 0 】

上記テンプレートマッチング部 2 3 2 には、予め、例えば 1 0 0 人程度の人間の顔画像の

50

平均から生成される平均的な人間の顔領域をテンプレートデータとして登録されている。

【 0 0 7 1 】

ウィンドウ画像の領域が顔領域であるか否かの判断は、上記テンプレートマッチング部 2 3 2 に顔領域のテンプレートデータとして登録することにより、かかる顔領域か否かの判断基準となる閾値が設定され、当該ウィンドウ画像について、テンプレートデータとなる平均的な顔領域との簡単なマッチングをすることにより判断される。

【 0 0 7 2 】

テンプレートマッチング部 2 3 2 は、ウィンドウ切出部 2 3 1 により送出されたウィンドウ画像について、テンプレートデータによるマッチング処理を行い、テンプレートデータとマッチングし、顔領域であると判断された場合には、当該ウィンドウ画像をスコア画像（顔領域と判断されたウィンドウ画像。）として後続の前処理部 2 3 3 に送出する。

10

【 0 0 7 3 】

また、上記ウィンドウ画像について、顔領域でないと判断された場合には、当該ウィンドウ画像そのまま結果判定部 2 3 5 に送出する。なお、上記スコア画像には、顔領域と判断された度合いがどの程度確からしいのかを示す信頼度情報が含まれる。例えば、信頼度情報は、スコア値が“ 0 0 ”～“ 9 9 ”の範囲内の数値を表し、数値が高いほど、より顔領域であることが確からしいことを表す。なお、信頼度情報は、例えば結果判定部 2 3 5 に備わるキャッシュ（図示せず。）などに格納される場合でもよい。

【 0 0 7 4 】

上記説明の正規化相関法、誤差二乗法などの演算処理は、後続の前処理部 2 3 3 および S V M 識別部 2 3 4 における演算処理と比較すると、演算処理量が 1 0 分の 1 から 1 0 0 分の 1 程度で済むとともに、テンプレートマッチング部 2 3 2 によるマッチング処理時点で、8 0 ( % ) 以上の確率で顔領域であるウィンドウ画像を検出することが可能である。つまり、明らかに顔領域でないウィンドウ画像を、この時点で除去することが可能となる。

20

【 0 0 7 5 】

前処理部 2 3 3 は、テンプレートマッチング部 2 3 2 から得られたスコア画像について、矩形領域でなる当該スコア画像から人間の顔領域とは無関係な背景に相当する 4 隅の領域を抽出するべく、当該 4 隅の領域を切り取ったマスクを用いて、2 0 × 2 0 画素あるスコア画像から 3 6 0 画素分を抽出する。なお本実施の形態にかかるスコア画像は 4 隅を切り取った 3 6 0 画素分を抽出する場合を例に挙げて説明したが、かかる例に限定されず、例えば、4 隅を抽出しない場合であっても実施可能である。

30

【 0 0 7 6 】

さらに前処理部 2 3 3 は、撮像時の照明などにより濃淡で表される被写体の傾き条件を解消するために、例えば平均二乗誤差（R S M : R o o t M e a n S q u a r e）などによる算出方法を用いて当該抽出された 3 6 0 画素のスコア画像の濃淡値に補正をかける。

【 0 0 7 7 】

続いて、前処理部 2 3 3 は、当該 3 6 0 画素のスコア画像のコントラストが強調された結果のスコア画像を、ヒストグラム平滑化処理を行うことにより、撮像装置 1 0 2 の撮像素子のゲイン又は照明の強弱に左右されないスコア画像を検出させることが可能となる。

40

【 0 0 7 8 】

またさらに、前処理部 2 3 3 は、例えばスコア画像をベクトル変換し、得られたベクトル群をさらに 1 本のパターンベクトルに変換するため、ガボア・フィルタリング（G a b o r F i l t e r i n g）処理を行う。なお、ガボア・フィルタリングにおけるフィルタの種類は必要に応じて変更可能である。

【 0 0 7 9 】

S V M 識別部 2 3 4 は、前処理部 2 3 3 からパターンベクトルとして得られたスコア画像に対して顔領域の検出を行う。そして検出された場合、顔領域検出データとして出力する。検出されない場合は、顔領域未検出データとして追加され、さらに学習する。

【 0 0 8 0 】

50

SVM識別部234は、前処理部233により送出されたスコア画像に基づいて生成されたパターンベクトルについて、当該スコア画像内に顔領域が存在するか否かを判断し、顔領域が検出された場合、当該スコア画像における顔領域の左上位置（座標位置）、顔領域の面積（縦×横の画素数）、顔領域であることの確からしさを表す信頼度情報、当該スコア画像の切出しの元となるスケール画像の縮小率（第1～第5のスケール画像に該当する縮小率のうちいずれか一つ。）とからなる顔領域情報を、例えば結果判定部235に備わるキャッシュ（図示せず。）に格納することにより、スコア画像ごとにリスト化する。なお、本実施の形態にかかる顔領域の位置（起点）は、画像左上である場合に限らず、例えば画像右上などであっても実施可能である。

【0081】

SVM識別部234により、例えば、第1のスケール画像のうち最初のウィンドウ画像の顔領域の検出が終了すると、ウィンドウ切出部231により第1のスケール画像の中の次にスキャンされたウィンドウ画像がテンプレートマッチング部232に送出される。

【0082】

次にテンプレートマッチング部232は、当該ウィンドウ画像がテンプレートデータにマッチングした場合のみスコア画像として、前処理部233に送出する。前処理部233は、上記スコア画像をパターンベクトルに変換してSVM識別部234に送出する。SVM識別部234は、パターンベクトルに基づき顔領域を検出した場合、上記スケール画像に関する顔領域情報を生成し、上記結果判定部235に備わるキャッシュに格納する。

【0083】

上記記載のように、第1のスケール画像について、ウィンドウ切出部231により順次スキャンされたウィンドウ画像について、以降後続のテンプレートマッチング部232、前処理部233、及びSVM識別部234による各処理が実行され、当該第1のスケール画像から顔領域が含まれるスコア画像を複数検出することが可能となる。

【0084】

さらに、ウィンドウ切出部231による第1のスケール画像のスキャンが全て終了し、後続のテンプレートマッチング部232、前処理部233、及びSVM識別部234による各処理についても終了すると、第2のスケール画像について、上記説明の第1のスケール画像とほぼ同様に顔領域の検出するための各処理が実行される。第3～第5のスケール画像についても、第1のスケール画像とほぼ同様にして顔領域の検出処理が実行される。

【0085】

SVM識別部234は、メモリ部202から読み出した映像データであるフレーム画像を5段階の相異なる縮小率から構成される第1～第5のスケール画像について、顔領域が検出されたスコア画像をそれぞれ複数検出し、その結果、生成される顔領域情報を、上記結果判定部235に備わるキャッシュ（図示せず。）に格納する。なお、本実施の形態にかかるキャッシュは、結果判定部235に備わる場合を例に挙げて説明したが、かかる例に限定されず、例えば、顔検出ブロック203内に単独で備わる場合などであっても実施可能である。さらに、顔領域が検出されずスコア画像が全く得られない場合もあるが、少なくとも1個など、所定の個数だけスコア画像が得られれば、顔検出処理は続行される。

【0086】

上記第1～第5のスケール画像において顔領域が検出されたスコア画像は、ウィンドウ切出部231におけるスキャンが所定画素（例えば、2画素など。）ずつ移動しながら実行されているため、前後のスコア画像の間では、近傍領域において高い相関性があり、相互に重なり合う領域を有する場合が多い。

【0087】

結果判定部235は、上記重複する領域を除去するため、2つのスコア画像の位置、スコア画像の画素数、および所定の数式に基づき、重複しているか否かを判定する。

【0088】

例えば、上記2つのスコア画像の位置として左上角の位置を、 $X$ ・ $Y$ 座標により $(X_A, Y_A)$ 、 $(X_B, Y_B)$ とそれぞれ表し、スコア画像の画素数（縦×横）を、それぞれ $H$

10

20

30

40

50

$L_A \times L_A, H_B \times L_B, dX (= X_B - X_A), dY (= Y_B - Y_A)$  とすると、以下に示す(1)式および(2)の関係が同時に成り立つ場合、2つのスコア画像は重なり合うと判定される。

【0089】

$$(L_A - dX) \times (L_B + dX) > 0 \dots\dots (1)$$

【0090】

$$(H_A - dY) \times (H_B + dY) > 0 \dots\dots (2)$$

【0091】

結果判定部235は、当該判定結果に基づいて、複数のスコア画像のうち重なり合う領域を除くことにより、重なり合わない最終的な顔領域を取得し、最終的に確定となる顔領域情報を生成し、上記キャッシュに格納されていた顔領域情報を更新する。なお、本実施形態にかかる格納されていた顔領域情報は、確定された顔領域情報に更新される場合を例に挙げて説明したが、かかる場合に限らず、別途新規に確定された顔領域情報を格納する場合であっても実施可能である。

10

【0092】

重なり合う領域が存在する場合、結果判定部235は、キャッシュ(図示せず。)に格納されたスコア画像に対応する信頼度情報に基づき、信頼度の高い、つまり顔領域である確からしさが高いスコア画像の方の顔領域情報を生成し、上記キャッシュに格納された当該顔領域情報を信頼度の高い顔領域情報に更新する。

【0093】

結果判定部235は、上記顔領域が検出されない場合、キャッシュに格納処理を行わず、さらに重なり合う顔領域が存在しない場合は、顔領域情報の更新は行わない。

20

【0094】

以上から、顔検出ブロック203は、撮像装置102により撮影された映像データから、信頼性の高い顔領域に対して顔領域情報を生成することが可能となる。したがって、複数の顔領域が検出されても、より確実に、例えば使用者106の顔領域を検出することが可能となる。

【0095】

上記生成された顔領域情報は、図2に示す符号化制御部205に送信されて、顔領域情報に基づき、映像データを圧縮符号化するための符号化パラメータが生成される。

30

【0096】

なお、本実施の形態にかかる結果判定部235による重複領域の判定処理は、(1)式に定められた場合を例に挙げて説明したが、かかる例に限定されず、他の数式を用いた場合であっても実施可能である。

【0097】

また、本実施の形態にかかるスケール画像をはじめとする画像の位置は、左上隅を基準に表される場合を例に挙げて説明したが、かかる例に限定されず、他の位置を基準とした場合であっても実施可能である。

【0098】

また、本実施の形態にかかる顔領域の検出される映像データは、フレーム単位に読み込まれて、顔領域が検出処理される場合を例に挙げて説明したが、かかる例に限定されず、例えば、フィールド単位又は複数フレームからなるシーンごとに顔領域の検出処理を行う場合などであっても実施可能である。

40

【0099】

また、本実施の形態にかかるテンプレートマッチング232に登録されるテンプレートデータは、平均的な人間の顔を示す顔領域が登録される場合を例にあげて説明したが、かかる例に限定されず、例えばテンプレートデータとして、乗用車のナンバープレート、時計、またはペットなどの動物の顔の画像領域が登録される場合であっても実施可能である。

【0100】

(2.2.5 通信部207)

50

次に、本実施の形態にかかる通信部 207 について説明する。通信部 207 は、ネットワーク 105 と接続され、ネットワーク 105 を介して圧縮符号化された伝送データを送信、または伝送データを受信する。

#### 【0101】

通信部 207 には、ネットワーク 105 のトラフィックの混雑状況を検知する検査部 210 を備える。検査部 210 は、ネットワーク 105 のトラフィックの混雑状況を検知するため、所定時間ごとに、例えば“ping”を利用した ICMP などにより、接続先の映像通信装置 104、または任意のホストに対し動作確認を要求（エコー検査）する。

#### 【0102】

検査部 210 は、ping コマンドにより、少なくとも接続相手先のアドレス情報を設定し、ICMP パケットを送信する。接続相手先の例えばホストなどは、上記 ICMP パケットを受信すると、ping コマンド発行元の検査部 210 に対し、正常に受信された旨の応答（Reply）パケットを送信する。なお、正常に接続相手先に受信されない場合（または、制限時間内に ICMP パケットが受信されなかった場合）は、エラーとなる。

#### 【0103】

したがって、検査部 210 は、ICMP パケット送信してから上記応答パケットを受信するまでの時間を取得し、トラフィックの混雑状況を検知する。例えば、ネットワーク 105 のトラフィックが平常時において、通信速度が 128 KByte/sec 及び上記応答パケットを受信するまでの時間（以下、応答時間）を 40 msec の場合、ある時点の検査部 210 の検査で、上記応答時間が 80 msec と検知されると、検査部 210 は、ネットワーク 105 のトラフィックは混雑していると判断する。

#### 【0104】

検査部 210 は、ネットワーク 105 のトラフィックの混雑を検知すると、混雑情報を生成し、符号化制御部 205 に上記混雑情報を送信する。混雑情報はネットワーク 105 のトラフィックの混雑状況を示すデータであり、例えば、応答時間などの情報が含まれる。

#### 【0105】

符号化制御部 205 は、上記混雑情報を受信すると、ネットワーク 105 のトラフィックの混雑状況に応じて、マクロブロック単位に映像データの圧縮符号化を制御させるため、符号化パラメータを設定する。例えば、所定時間内の複数フレームの映像データについては圧縮符号化せず、伝送データを送信しないように制御させる、または所定時間内の複数フレームの映像データについては、顔領域に属すマクロブロックだけを圧縮符号化し、伝送データを送信するよう制御させる符号化パラメータが例示される。なお、以下に記載されるマクロブロックは、図 5 に示す MB503 を示すこともある。マクロブロックについては、後程詳述する。

#### 【0106】

なお、本実施の形態にかかる検査部 210 は、ICMP（Internet Control Message Protocol）によりトラフィックの混雑状況を検知する場合を例に挙げて説明したが、かかる例に限定されず、例えば TCP（Transmission Control Protocol）セグメントなどデータの再送信処理の際に、接続先の相手側から確認応答が返ってくるまでの時間（RTT：Round Trip Time）を取得する、または接続先の例えばホストなどに、まとめてデータを送受信することが可能なウィンドウ・サイズ（受信可能なデータサイズ）の変動により混雑状況を検知する場合であっても実施可能である。

#### 【0107】

##### （3．双方向コミュニケーションシステムの動作）

次に、図 3 を参照しながら、上記のように構成された双方向コミュニケーションシステムの動作の実施形態について説明する。図 3 は、本実施の形態にかかる双方向コミュニケーションシステムの動作の概略を示すフローチャートである。

#### 【0108】

図 3 に示すように、本実施の形態にかかる双方向コミュニケーションシステムにおいて、

10

20

30

40

50

例えばテレビ会議などにより複数の使用者 106 が打ち合わせをする場合、打ち合わせされる時間内は絶えず複数の映像配信ユニット 101 間で、相互に映像データをやりとりし、双方向コミュニケーションシステムの動作が継続される。

【0109】

したがって、打ち合わせ時間が終了（撮影処理が終了）するまで、映像配信ユニット 101 間で、映像データの配信処理（S301）が続行（配信ループ）される。

【0110】

（3.1 映像配信ユニット 101 からの映像データ配信処理）

次に、図4を参照しながら、本実施の形態にかかる映像データ配信処理について説明する。図4は、本実施の形態にかかる映像データ配信処理の概略を示すフローチャートである。なお、以下の説明は、ITU-T勧告 H.263 の場合である映像データ配信処理について説明するが、MPEG-4 についても準拠する。

10

【0111】

映像データ配信処理（S301）は、撮像装置 102 の撮影処理により、映像データが生成されると、例えば、RS-232C または RS-422 などを通じて、映像通信装置 104 の変換部 201 に送出される。

【0112】

変換部 201 は、上記映像データを A/D 変換し、メモリ部 202 に送出する。映像データが、メモリ部 202 に送出されると、図4に示すように、顔検出ブロック 203 により顔検出処理（S401）が行われる。なお、本実施の形態にかかる顔検出処理は、上記説明したのとほぼ同様の構成であるため省略する。

20

【0113】

顔検出処理（S401）は、メモリ部 202 に送出される映像データのフレーム単位に行われるが、かかる例に限らず、フィールド単位の場合でもよい。また、フレーム単位の映像データであるフレーム画像（ピクチャ）内に顔領域が存在しない、検出されない（S402）場合は、再度顔検出処理（S401）が行われる。

【0114】

顔検出処理（S401）の結果、顔領域が検出された（S402）場合は、映像通信装置 104 に備わる結果判定部 235 のキャッシュに格納された顔領域情報が符号化制御部 205 に送信される（S403）。

30

【0115】

符号化制御部 205 は、上記顔領域情報を受信すると、符号化制御部 205 内に備わる記憶部（図示せず。）に格納された少なくとも 1 フレーム前のフレーム画像にかかる顔領域情報を取得する。なお、取得されるフレーム画像は、1 フレーム前に限らず、例えば、複数フレーム前、または 1 フィールド前などであってもよい。

【0116】

上記 1 フレーム前のフレーム画像（前フレーム画像）にかかる顔領域情報が格納されている場合は、上記受信した現フレーム画像の顔領域情報と、前フレーム画像にかかる顔領域情報とを比較し、補正処理を行う（S405）。

【0117】

上記前のフレーム画像にかかる顔領域情報が記憶部に格納されて無い場合（S404）、つまり前フレーム画像において顔領域が検出されない場合（S404）には、顔領域情報の補正処理（S405）は実行されない。

40

【0118】

上記補正処理（S405）は、前フレームおよび現フレーム画像にかかる顔領域情報の顔領域の面積情報、位置情報、または信頼度情報のうち少なくとも一つを比較することにより現フレーム画像にかかる顔領域情報を補正する。

【0119】

本実施の形態にかかる補正処理（S405）は、例えば、前フレーム画像において 1 の顔領域のみ存在し、現フレーム画像において 2 の顔領域が存在し、現フレーム画像において

50

も前フレーム画像で検出された顔領域を選択する場合、現フレーム画像に含まれる前フレーム画像にかかる顔領域情報を、選択するため正確に顔領域情報を判別する必要がある。

【 0 1 2 0 】

前フレーム画像および現フレーム画像間の時間差は極めて短く、人間の動作によりフレーム画像内を移動可能な範囲は極めて限られているため、符号化制御部 205 は、顔領域情報の面積情報と位置情報とに基づき、現フレーム画像にかかる顔領域のうち、前フレーム画像にかかる顔領域の近傍に存在する顔領域の顔領域情報を選択する。

【 0 1 2 1 】

選択された顔領域情報のうち信頼度情報が、現フレーム画像の他の信頼度情報または前フレーム画像の信頼度情報に比べて低い場合、前フレーム画像の信頼度情報と同程度もしくは現フレーム画像にかかる他の信頼度情報以上の値に補正する (S 405)。したがって、例えば、信頼度情報が最も高い顔領域情報を選択すれば、前フレーム画像の顔領域を現フレーム画像においても正確に選択することが可能となる。なお、本実施の形態にかかる補正処理は、かかる例に限定されない。

10

【 0 1 2 2 】

符号化制御部 205 は、補正された現フレーム画像にかかる顔領域情報に基づき、信頼度情報の最も高い顔領域に対してオブジェクトの切出処理 (S 406) をする。なお、本実施の形態にかかるオブジェクトの切出処理は、信頼度情報の最も高い顔領域に限定されることなく、例えば、信頼度情報に依存しない全ての顔領域、または最も低い信頼度情報を除く他の顔領域全てについて、オブジェクトの切出処理 (S 406) をする場合であっても実施可能である。

20

【 0 1 2 3 】

( 3 . 1 . 1 映像フォーマット )

ここで、オブジェクトの切出処理 (S 406) を説明する前に、図 5 を参照しながら、本実施の形態にかかる映像フォーマットについて説明する。図 5 は、本実施の形態にかかる映像フォーマットの概略的な構成を示す説明図である。

【 0 1 2 4 】

撮像装置 102 により、NTSC 方式又は PAL 方式にて撮影された映像データは、フレーム画像単位に、例えば ITU-T 勧告に定める H. 261, H. 263, または ISO/IEC 14496 に定める MPEG-4 などの場合において、予め共通フォーマットとして定められた CIF 画面, QCIF 画面, または SQCIF 画面などのフレーム画像に変換され、さらに圧縮符号化され、伝送データとしてネットワーク 105 を介して送信される。

30

【 0 1 2 5 】

図 5 に示すように、画面 501 は、上記 CIF 画面, QCIF 画面, または SQCIF 画面のいずれかに該当し、グループ・オブ・ブロックと呼ばれる複数の GOB (502A, 502B, 502C, ...) から構成されている。

【 0 1 2 6 】

例えば、本実施の形態にかかる GOB 502 は、H. 261 の場合、CIF 画面では、12 個の GOB 502 から構成され、QCIF 画面では 3 個の GOB 502 から構成される。

40

【 0 1 2 7 】

また、GOB 502 は、さらにマクロブロック (MB) と呼ばれる、複数の MB (503A, 503B, 503C, ...) から構成され、各 MB 503 は、16 × 16 画素の輝度マクロブロックである MB 503-1 と、8 × 8 画素の C<sub>B</sub> 色差マクロブロックである MB 503-2 と、8 × 8 画素の C<sub>R</sub> 色差マクロブロック 503-3 とから構成されるが、GOB 502 に構成される MB 503 の個数は、例えば H. 261, H. 263, または MPEG-4 などに応じて変動し、H. 261 の場合、1 の GOB 502 に、33 個の MB 503 から構成されている。

【 0 1 2 8 】

50

また、MB503は、さらに8×8画素からなる最小単位のブロック(504A, 504B, 504C, 504D)から構成されている。したがって、1のMB503には、4個の輝度ブロック(504A, 504B, 504C, 504D)と、2個の(C<sub>B</sub>, C<sub>R</sub>)色差ブロック(504E, 504F)とから構成されている。

【0129】

(3.1.2 マクロブロックのデータ構造)

次に、図6を参照しながら、本実施の形態にかかるマクロブロックのデータ構造について説明する。図6は、本実施の形態にかかるマクロブロックのデータ構造の概略的な構成を示す説明図である。

【0130】

図6に示すように、マクロブロックのデータ構造は、マクロブロックヘッダと、ブロックデータとからなり、上記マクロブロックヘッダは、“COD”と、“MCBPC”と、“MODB”と、“CBPB”と、“CBPY”と、“DQUANT”と、“MVD”と、“MVD<sub>2</sub>”と、“MVD<sub>3</sub>”と、“MVD<sub>4</sub>”と、“MVDB”とから構成される。

【0131】

なお、本実施の形態にかかるマクロブロックのデータ構造は、H.263にかかるデータ構造である場合を例にあげて説明したが、かかる例に限定されず、例えば、H.261、またはMPEG-4などの場合であっても、H.263に準拠する。

【0132】

上記“DQUANT”は、2ビット又は可変長データであり、QUANTの変化を定義する。QUANTは、マクロブロックに対する量子化パラメータであり、1～31の範囲の値を取り得る。なおQUANTは、予め任意の値に設定されている。

【0133】

したがって、“DQUANT”は、差分値を表すことから、例えば、“DQUANT”が2進数表示で“00”の場合、差分値は“-1”であり、“01”の場合、差分値は“-2”であり、“10”の場合、差分値は“1”であり、“11”の場合、差分値は“2”と表すことができる。

【0134】

“DQUANT”の差分値が変化することにより、QUANTの値が変化するが、量子化パラメータであるQUANTが大きくなると、該当するマクロブロックの画質は落ちて、ぼんやりと精細を欠いた画像になる、QUANTが小さくなると画質は向上して、圧縮符号化しても、ほぼ元の原画に近い状態の画像になる。つまりマクロブロックごとに、“DQUANT”の変化を制御することにより、映像データの任意領域の画質を制御することが可能となる。上記“DQUANT”の変化は、符号化制御部205により生成される符号化パラメータに基づいて、制御される。

【0135】

図6に示すように、H.263にかかる“COD”は、符号化マクロブロックインジケータであり、1ビットからなるデータである。“COD”が“0”である場合、圧縮符号化される対象のマクロブロックであることを示し、“1”である場合、圧縮符号化されず削除または無視されるマクロブロックであることを示す。

【0136】

したがって、H.263の場合において、符号化制御部205は、マクロブロックを圧縮符号化するか否かを制御するため、上記マクロブロックの“COD”に値を指示するための符号化パラメータを生成する。

【0137】

ここで、図4に示すように、顔領域情報の補正処理(S405)が終了し、符号化制御部205は、上記顔領域情報を受信すると、上記顔領域情報に含まれる顔領域の面積情報または顔領域の位置情報に基づき、オブジェクトとしてフレーム画像の顔領域の切出処理(S406)を実行する。

【0138】

10

20

30

40

50

さらに、図7(A)及び図7(B)を参照しながら、本実施の形態にかかるオブジェクトについて説明する。図7(A)は、本実施の形態にかかる初期形成時の顔領域ブロックの概略的な構造を示す説明図であり、図7(B)は、本実施の形態にかかる最終決定時の顔領域ブロックの概略的な構造を示す説明図である。

【0139】

図7(A)および図7(B)に示す映像データのフレーム画像701は、36個(6×6)のマクロブロックから構成されている。

【0140】

まず図7(A)に示すように、符号化制御部205は、受信する顔領域情報に含まれる面積情報または位置情報に基づき、顔領域700の領域を初期形成する。図7(A)に示す顔領域700は、人間の顔が全て含まれる4つのマクロブロックの範囲内に収まっている。つまり顔領域700上から3ブロック、左から3ブロックを左上隅とする3×3マクロブロックの範囲内に収まっている。

【0141】

しかし、圧縮符号化はマクロブロック単位に行われるため、図7(B)に示すように、符号化制御部205は、顔領域700を、拡大又は縮小する割合が最小限であるマクロブロック単位領域の顔領域702に補正する。圧縮符号化する場合はマクロブロック単位に行われるため、顔領域702のように補正されて、顔領域として最終決定される。

【0142】

図7(B)に示す補正された顔領域702により、符号化制御部205は、顔領域702に属すマクロブロックと、顔領域702に属さないマクロブロックと、別の領域として、オブジェクト単位に切出す(S406)。したがって、顔領域702のオブジェクトに対して、量子化パラメータを小さくするなど、オブジェクトごとに圧縮符号化させるよう、符号化パラメータで指示することができる。

【0143】

さらに、例えば、符号化制御部205は、顔領域702に属すマクロブロックに対しては、“COD”に“0”が設定されるよう、符号化パラメータで指示し、顔領域702に属さないマクロブロックに対しては、“COD”に“1”が設定されるよう、符号化パラメータで指示することで、顔領域702だけが圧縮符号化されて、伝送データとしてネットワーク105を介して送信されることができる。

【0144】

(3.1.3 顔領域変換処理)

図2に示す特殊処理部204は、メモリ部202に格納される映像データのフレーム単位に、検出された顔領域に対して、例えばモザイク処理、または動物の画像など他の画像に置換するなどの顔領域変換処理(S407)を実行する。

【0145】

上記顔領域変換処理(S407)は、例えば、映像通信装置104に備わるモザイク処理設定ボタン及び置換処理設定ボタン(図示せず。)などにより、モザイク処理または置換処理が設定された場合、実行される。なお、本実施の形態にかかる顔領域変換処理(S407)は、撮影処理前に予め設定する場合、または撮影処理中に設定する場合のどちらであっても実施可能である。

【0146】

ここで、図8を参照しながら、本実施の形態にかかる顔領域変換処理について説明する。図8は、本実施の形態にかかる顔領域変換処理の概略を示すフローチャートである。

【0147】

図8に示すように、モザイク処理または置換処理からなる顔領域変換処理が設定されていると(S801)、特殊処理部204は、メモリ部202に格納された映像データをフレーム単位に読み出し、さらに置換処理が設定されている場合には、置換するための適当な置換画像データを読み出す。

【0148】

10

20

30

40

50

さらに特殊処理部 204 は、顔検出ブロック 203 から送信される顔領域情報に基づき、上記映像データにおけるフレーム画像の顔領域に対し、モザイク処理または置換処理 (S802) して、エンコーダ部 206 にフレーム画像を送出する。

【0149】

モザイク処理または置換処理 (S802) が終了することにより、図 4 に示す顔領域変換処理 (S407) が終了する。なお、本実施の形態にかかる顔領域変換処理は、モザイク処理または置換処理から構成される場合を例にあげて説明したが、かかる例に限定されず、例えば、シャープネス処理、フレーム画像の明度を上げる明度処理などの場合であっても実施可能である。

【0150】

また本実施の形態にかかる顔領域変換処理は、顔領域に対してモザイク処理または置換処理が実行される場合を例にあげて説明したが、かかる例に限定されず、顔領域以外の領域に対してモザイク処理又は置換処理を実行する場合であっても実施可能である。

【0151】

次に、図 4 に示すように、特殊処理部 204 において顔領域変換処理 (S407) が終了すると、符号化制御部 205 は、特殊処理部 204 から送られるフレーム画像に対する符号化パラメータを生成する (S408)。

【0152】

符号化制御部 205 は、エンコーダ部 206 に、少なくとも顔領域 702 に属すマクロブロックに対する量子化パラメータの設定、顔領域 702 に属さないマクロブロックに対する量子化パラメータの設定、またはオブジェクト単位に圧縮符号化するか否かの設定などを指示するための符号化パラメータを生成する (S408)。

【0153】

さらに、上記説明したように検査部 210 により、ネットワーク 105 のトラフィックの混雑状況の検知処理 (S409) を実行する。検知処理 (S409) の結果、トラフィックの混雑状況が所定の閾値を超えて、検査部 210 により混雑していると判断されると (S410)、混雑情報を生成し、符号化制御部 205 に送信する。

【0154】

符号化制御部 205 は、上記混雑情報を受信すると、例えば、顔領域 702 であるオブジェクトに限定して圧縮符号化させるようにエンコーダ部 206 に符号化パラメータを送信し、圧縮符号化を制御する。

【0155】

フレーム画像の顔領域 702 だけを圧縮符号化させるのは、上記説明の通り、顔領域 702 に属すマクロブロックの“COD”に“0”を設定し、顔領域 702 に属さないマクロブロックには、“COD”に“1”を設定することで、ネットワーク 105 には顔領域 702 にかかる伝送データが送信される。

【0156】

したがって、符号化制御部 205 は、エンコーダ部 206 に上記顔領域 702 のオブジェクトのみを圧縮符号化させるため、符号化パラメータ生成処理 (S408) で生成された符号化パラメータを変更処理 (S411) し、上記符号化パラメータをエンコーダ部 206 に送信する。

【0157】

上記符号化パラメータの変更処理 (S411) により、エンコーダ部 206 の圧縮符号化するか否かを制御することが可能となり、ネットワーク 105 のトラフィックに負荷を最小限に留めることが可能となる。

【0158】

次に、エンコーダ部 206 は、符号化パラメータに基づき、特殊処理部 204 から送られる映像データであるフレーム画像を圧縮符号化 (S412) し、通信部 207 に伝送データとして送る。したがって、例えば、顔領域 702 に属すマクロブロックに対しては画質を落とさず圧縮符号化し、顔領域 702 に属さないマクロブロックに対しては画質

10

20

30

40

50

を落として圧縮符号化させることが可能である。さらにまた、顔領域702に属すマクロブロックだけを圧縮符号化することも可能である。

【0159】

したがって、フレーム画像全体を圧縮符号化せずに、フレーム画像内の顔領域702に対するマクロブロックのみを切り出して圧縮符号化することが可能であり、ネットワーク105に送出するデータ容量を節約することが可能となり、さらに人間の顔画像の画質は落ちないため、視認性の高い映像データを表示することができる。

【0160】

ここで、MPEG-4の場合における本実施の形態にかかる圧縮符号化について説明すると、MPEG-4の圧縮符号化(S412)は、H.261及びH.263の圧縮符号化(S412)とは、エンコーダ部206に形状符号化部(図示せず。)およびテクスチャ符号化部(図示せず。)を備えることで実施される点で相違する。

10

【0161】

上記形状符号化部は、上記顔領域702であるオブジェクトの形状を符号化するために、まず符号化すべき領域を図7(A)または(B)に示すフレーム画像701にバウンディングレクタングルを設定し、図7(B)に示すマクロブロックと同じ位置に16×16画素のブロック(2値形状ブロック:BAB)を設定する。

【0162】

図9に示すように、形状符号化部は、符号化パラメータに基づき、2値形状ブロックを設定すると、顔領域702であるオブジェクトに属す2値形状ブロックは、“1”で表され、オブジェクトに属さない2値形状ブロックは、“0”で表される。図9は、本実施の形態にかかる2値形状ブロックの概略的な構成を示す説明図である。

20

【0163】

図9に示す2値形状ブロックのように、顔領域702であるオブジェクトの内部と外部とを区別するために、2値で表示されると、形状符号化部は、2値形状ブロックごとに当該フレーム画像701の形状符号化をする。

【0164】

また、形状符号化されるとともに、テクスチャ符号化部は、上記顔領域702であるオブジェクトに属すマクロブロックに対してパディング処理などを行い、テクスチャ(画素値)の圧縮符号化が行われる。形状符号化及びテクスチャ符号化されることにより、圧縮符号化処理(S412)が処理終了し、エンコーダ部206は、伝送データを通信部207に送出する。なお、本実施の形態にかかるテクスチャ符号化部は、オブジェクトに属さないマクロブロックに対して、圧縮符号化する場合であっても実施可能である。

30

【0165】

したがって、フレーム画像全体を圧縮符号化せずに、顔領域702に対するマクロブロックのみを切り出して圧縮符号化することが可能であり、ネットワーク105に送出するデータ容量の軽減化が図れ、人間の顔画像の画質は落ちないため、視認性の高い映像データを表示することができる。

【0166】

送出された伝送データは、通信部207により多重化され、ネットワーク105を介して、配信される(S413)。以上から構成される映像データ配信処理(S401~S413)は、撮影処理が終了するまで継続される。

40

【0167】

なお、本実施の形態にかかる配信後の映像データの受信処理については、ネットワーク105を介して送信された伝送データが、通信部207により受信され、デコーダ部208により伸長されるとメモリ部202に順次、映像データが格納される。

【0168】

以後の処理については、図4に示す顔検出処理(S401)~顔領域変換処理(S407)が行われ、映像データは、変換部209によりD/A変換される。D/A変換後、出力装置103は、映像データを表示する。なお本実施の形態にかかる映像データの受信処理

50

の顔検出処理（S401）～顔領域変換処理（S407）における処理は、映像データの配信処理の顔検出処理（S401）～顔領域変換処理（S407）の処理とほぼ同様な構成であるため詳細な説明は省略する。

【0169】

以上、添付図面を参照しながら本発明の好適な実施形態について説明したが、本発明はかかる例に限定されない。当業者であれば、特許請求の範囲に記載された技術的思想の範疇内において各種の変更例または修正例を想定し得ることは明らかであり、それらについても当然に本発明の技術的範囲に属するものと了解される。

【0170】

上記実施形態においては、映像配信ユニットが複数台から構成される場合を例にあげて説明したが、本発明はかかる例に限定されない。例えば、映像配信ユニットが1台から構成される場合であっても実施することができる。この場合には、監視システムとして実施することが可能である。

10

【0171】

また、上記実施の形態においては、人間の顔領域である場合を例にあげて説明したが、本発明はかかる例に限定されない。例えば、乗用車のナンバープレートの画像などを特徴を有する領域として実施する場合であってもよい。

【0172】

また、上記実施の形態においては、映像データの配信処理および受信処理はフレーム単位に行われる場合を例に挙げて説明したが、本発明はかかる例に限定されない。例えば、映像データのフィールド単位、または、映像データの複数フレームから構成されるシーン単位で行われる場合でも実施可能である。

20

【0173】

また、上記実施の形態においては、映像配信ユニットは、テレビ会議に用いられる場合を例にあげて説明したが、本発明は、かかる例に限定されない。例えば、携帯電話、携帯端末、またはパソコン（Personal Computer）などに用いる場合であっても実施可能である。

【0174】

【発明の効果】

以上説明したように、本発明によれば、複数の特徴領域が存在する場合でも過去の特徴領域の情報によりの確に特徴領域を判断し、特徴領域のみ画質を落とさず切出して圧縮符号化することにより、ネットワークのトラフィックに依存せず視認性の高い画像を表示することができる。

30

【図面の簡単な説明】

【図1】図1は、本実施の形態にかかる双方向コミュニケーションシステムの概略的な構成を示すブロック図である。

【図2】図2は、本実施の形態にかかる映像通信装置の概略的な構成を示すブロック図である。

【図3】図3は、本実施の形態にかかる双方向コミュニケーションシステムの動作の概略を示すフローチャートである。

40

【図4】図4は、本実施の形態にかかる映像データ配信処理の概略を示すフローチャートである。

【図5】図5は、本実施の形態にかかる映像フォーマットの概略的な構成を示す説明図である。

【図6】図6は、本実施の形態にかかるマクロブロックのデータ構造の概略的な構成を示す説明図である。

【図7】図7（A）は、本実施の形態にかかる初期形成時の顔領域ブロックの概略的な構造を示す説明図であり、図7（B）は、本実施の形態にかかる最終決定時の顔領域ブロックの概略的な構造を示す説明図である。

【図8】図8は、本実施の形態にかかる顔領域変換処理の概略を示すフローチャートであ

50

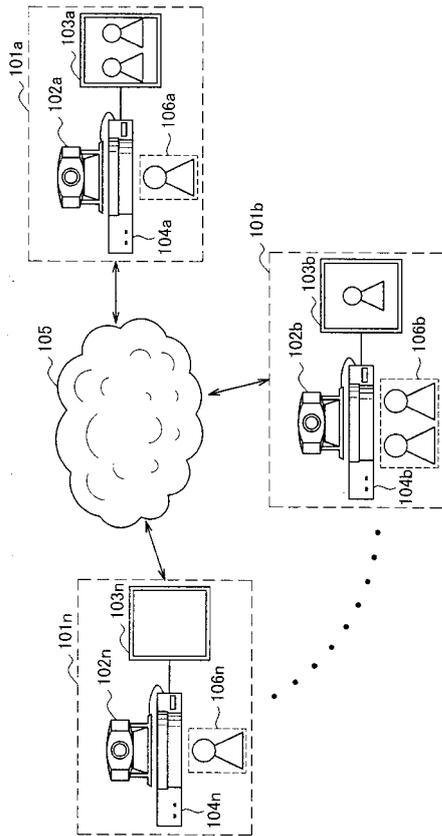
る。

【図9】図9は、本実施の形態にかかる2値形状ブロックの概略的な構成を示す説明図である。

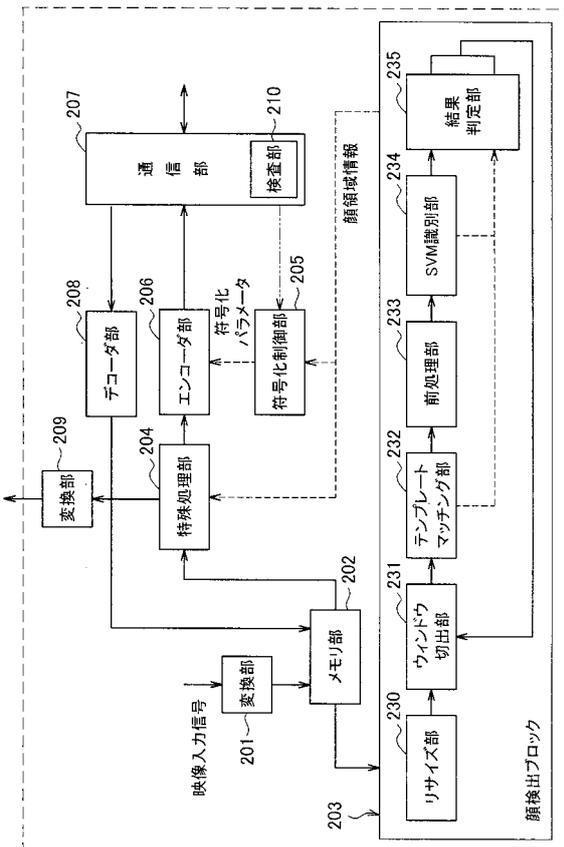
【符号の説明】

- 101 : 映像配信ユニット
- 102 : 撮像装置
- 103 : 出力装置
- 104 : 映像通信装置
- 105 : ネットワーク
- 106 : 使用者

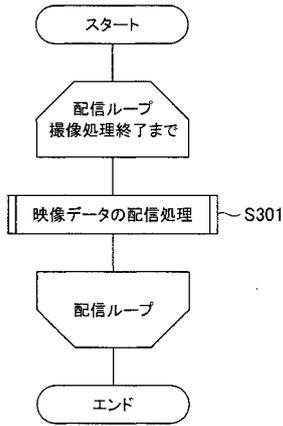
【図1】



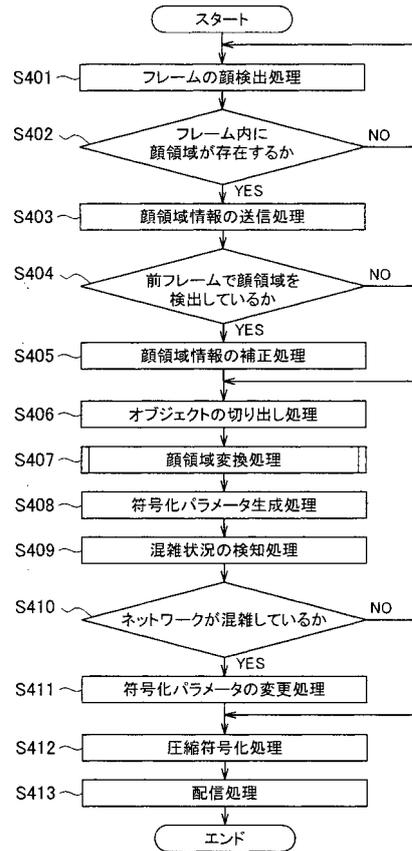
【図2】



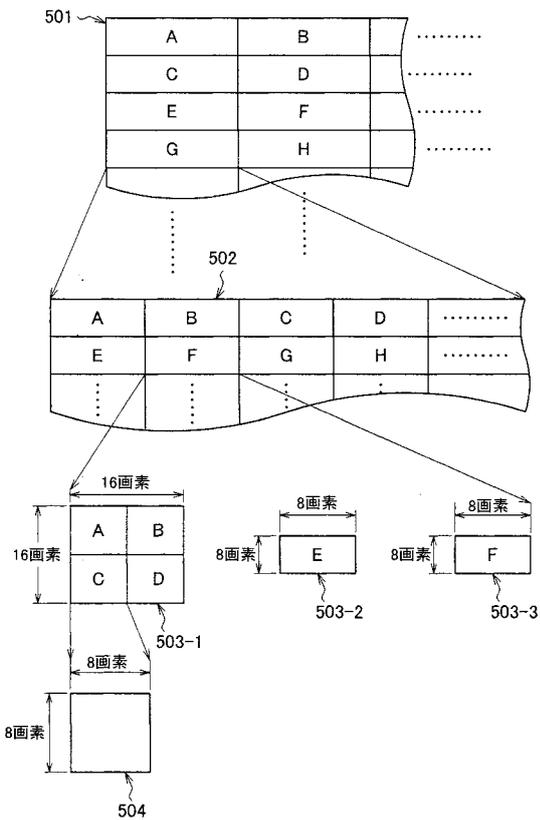
【図3】



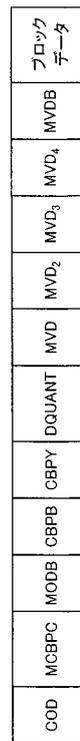
【図4】



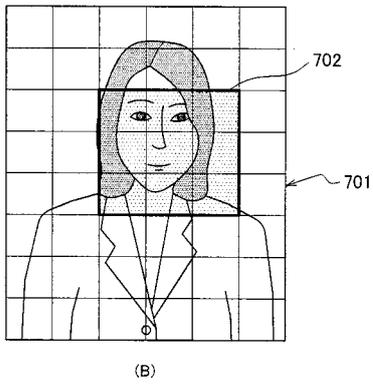
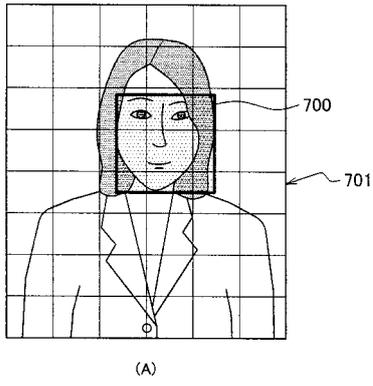
【図5】



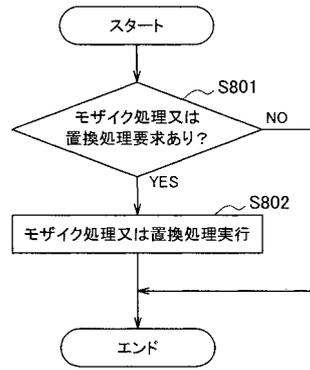
【図6】



【図7】



【図8】



【図9】

0	0	0	0	0	0
0	0	0	0	0	0
0	0	1	1	1	0
0	0	1	1	1	0
0	0	1	1	1	0
0	0	0	0	0	0
0	0	0	0	0	0
0	0	0	0	0	0

## フロントページの続き

- (56)参考文献 特開平07-203436(JP,A)  
特開平07-288789(JP,A)  
特開平05-095541(JP,A)  
特開平06-217276(JP,A)  
特開平06-350989(JP,A)  
特開平08-181992(JP,A)  
特開平09-036752(JP,A)  
特開平09-275564(JP,A)  
特開平11-328309(JP,A)  
特開平10-051770(JP,A)  
特開平06-046414(JP,A)  
特開2000-078588(JP,A)  
特開2000-197050(JP,A)  
特開2002-010259(JP,A)  
特開2001-078193(JP,A)  
特開2000-172852(JP,A)  
特開2002-007294(JP,A)  
特開平09-298749(JP,A)  
特開2002-252848(JP,A)  
特開2000-30066(JP,A)  
特開2001-92970(JP,A)  
特開平5-314320(JP,A)

## (58)調査した分野(Int.Cl., DB名)

- H04N 7/14-7/15,  
H04N 5/262,  
H04N 7/26-7/68