

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-257173

(P2012-257173A)

(43) 公開日 平成24年12月27日(2012.12.27)

(51) Int.Cl.	F I	テーマコード (参考)
HO4N 5/232 (2006.01)	HO4N 5/232 C	5C122
GO6T 7/20 (2006.01)	GO6T 7/20 100	5L096

審査請求 未請求 請求項の数 4 O L (全 9 頁)

(21) 出願番号 特願2011-130302 (P2011-130302)  
 (22) 出願日 平成23年6月10日 (2011.6.10)

(71) 出願人 000001007  
 キヤノン株式会社  
 東京都大田区下丸子3丁目30番2号  
 (74) 代理人 100090273  
 弁理士 園分 孝悦  
 (72) 発明者 大矢 崇  
 東京都大田区下丸子3丁目30番2号 キ  
 ヤノン株式会社内  
 Fターム(参考) 5C122 DA03 DA11 EA65 EA66 FH10  
 FH14 GD06 HB01 HB10  
 5L096 AA02 AA06 HA03 JA11

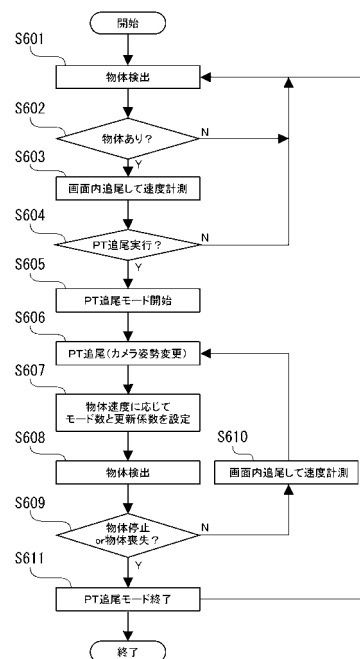
(54) 【発明の名称】 追尾装置、追尾方法及びプログラム

(57) 【要約】

【課題】 移動物体の速度に関係なく、長時間、より正確に移動体の追尾を行うことを目的とする。

【解決手段】 特徴量に対応したパラメータに基づいて撮像手段で撮像された画像中の動体を検出する物体検出手段と、カメラの姿勢を制御する姿勢制御手段と、動体の速度に応じて物体検出手段のパラメータを動的に変更する制御手段と、を有することによって課題を解決する。

【選択図】 図6



**【特許請求の範囲】****【請求項 1】**

特徴量に対応したパラメータに基づいて撮像手段で撮像された画像中の動体を検出する物体検出手段と、

カメラの姿勢を制御する姿勢制御手段と、

前記動体の速度に応じて前記物体検出手段のパラメータを動的に変更する制御手段と、を有する追尾装置。

**【請求項 2】**

前記特徴量とは、画像を構成する画素の輝度成分、又は色成分、又は部分領域に対する周波数成分である請求項 1 記載の追尾装置。

10

**【請求項 3】**

追尾装置が実行する追尾方法であって、

特徴量に対応したパラメータに基づいて撮像手段で撮像された画像中の動体を検出する物体検出ステップと、

カメラの姿勢を制御する姿勢制御ステップと、

前記動体の速度に応じて前記物体検出ステップのパラメータを動的に変更する制御ステップと、

を含む追尾方法。

**【請求項 4】**

コンピュータを、

特徴量に対応したパラメータに基づいて撮像手段で撮像された画像中の動体を検出する物体検出手段と、

20

カメラの姿勢を制御する姿勢制御手段と、

前記動体の速度に応じて前記物体検出手段のパラメータを動的に変更する制御手段と、して機能させるためのプログラム。

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、追尾装置、追尾方法及びプログラムに関する。

**【背景技術】**

30

**【0002】**

監視及びモニタリングシステムでは、動画像処理によって画像中の動体を検出する技術がある。検出した動体が画像上の指定した領域に侵入したことを検知したり、長時間静止したことを検出したりする技術がある。また検知した動体を常に画面中央に捉えるようにカメラの姿勢を制御する技術が追尾技術として知られている。

動体の検出には隣接フレーム間差分や背景差分が用いられる。隣接フレーム間差分は連続する二つのフレーム間に対応するサンプリング位置の、輝度成分や周波数成分の差分を用いて動体領域を抽出する。処理速度が速く、使用メモリ領域が少なく済む一方で、物体の境界領域が検出され易い。背景差分は物体の存在しないシーンにおいて予め撮影した背景と、現画像との比較によって物体領域を検出する。ここで背景は画素単位の輝度やブロック単位の周波数成分が用いられ、シーンの定常的な変化に対応するために、複数の状態を持つ統計量としてモデル化される例もある。背景差分は物体領域を検出できる一方で、背景モデルの作成に時間を要し、また背景モデルを保持するためにメモリリソースを要する。

40

**【0003】**

姿勢制御可能なカメラを用いて、フレーム間差分によって動体を検知し、追尾を行う技術が特許文献 1 に開示されている。カメラは静止状態でフレーム間差分を行い、検出した動体が画面中央に位置するように姿勢制御を行い、これを繰り返す。特許文献 2 は背景差分による追尾の例である。追尾範囲全体の追尾画像を予め撮影し、追尾時に現在のカメラの方向に一致する背景画像を切り出して背景差分処理に利用する。特許文献 3 は背景差分

50

を用いた追尾において、動きベクトルを用いて画像の位置合わせを行う手法が開示している。

この他の物体検出方法には対象物の画像特徴を格納したテンプレートを用いる手法や、動きベクトル解析を用いる手法がある。動きベクトル解析を用いた物体抽出によってカメラの移動中も物体を追尾する方法が特許文献5に開示されている。また特許文献6は物体の速度によって背景差分と動きベクトル解析を切り替えて、追尾を行う手法である。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2003-219225号公報

【特許文献2】特開平11-284989号公報

【特許文献3】特開平05-145823号公報

【特許文献4】特開平05-153459号公報

【特許文献5】特開2002-374521号公報

【特許文献6】特開2001-092976号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

カメラの視線制御による追尾手法として、テンプレートマッチングを用いた動体検出及び追尾は、物体の見え方の変化に対応するためにテンプレートを更新する必要があり、長時間の追尾に適さない。動きベクトルを用いた動体検出及び追尾も、正確な動きベクトルを安定して求めることは難しい。フレーム間差分を用いた動体検出及び追尾方式は、物体の境界領域を検出するために、物体速度が小さい時に物体を検出できない場合がある。背景差分による物体検出及び追尾方式は、物体の速度が大きいと、背景の生成が間に合わない場合がある。

【0006】

本発明はこのような問題点に鑑みなされたもので、移動物体の速度に関係なく、長時間、より正確に移動体の追尾を行うことを目的とする。

【課題を解決するための手段】

【0007】

そこで、本発明の追尾装置は、特徴量に対応したパラメータに基づいて撮像手段で撮像された画像中の動体を検出する物体検出手段と、カメラの姿勢を制御する姿勢制御手段と、前記動体の速度に応じて前記物体検出手段のパラメータを動的に変更する制御手段と、を有する。

【発明の効果】

【0008】

本発明によれば、移動物体の速度に関係なく、長時間、より正確に移動体の追尾を行うことができる。

【図面の簡単な説明】

【0009】

【図1】追尾システムのシステム構成等の一例を示す図である。

【図2】ネットワークカメラの機能構成等の一例を示す図である。

【図3】物体検出処理の手順を説明する図である。

【図4】物体の速度に応じた背景差分処理の制御方法を説明するための図である。

【図5】物体速度に応じた背景差分パラメータの最適化を説明するための図である。

【図6】追尾処理の処理手順を説明する図である。

【発明を実施するための形態】

【0010】

以下、本発明の実施形態について図面に基づいて説明する。

【0011】

10

20

30

40

50

### < 実施形態 1 >

本実施形態では、カメラ姿勢制御による動体追尾において、物体検出手段の制御パラメータを物体の速度に応じて適応的に変更する方式を説明する。この方式を用いれば、物体速度に関わらず、正確性と即応性とを兼ね備えた動体追尾を実現できる。

本実施形態におけるシステム構成の例を図1に示す。図1は、追尾システムのシステム構成等の一例を示す図である。ネットワークカメラ100、表示装置210、録画装置220がネットワーク経由で接続される。ネットワークカメラ100は、表示装置210、録画装置220からの要求に対して、映像150や不図示のメタデータ160を送信する。メタデータとは映像中の物体位置等のシーン情報である。録画装置220は、ネットワークカメラ100から取得した映像150やメタデータ160を記録媒体に保存する。表示装置210は、ネットワークカメラ100から得た映像150や、録画装置220に保存された映像を読み出して表示したり、ネットワークカメラ100の設定を行ったりする。

10

#### 【0012】

ネットワークカメラ100の構成を図2に示す。図2は、ネットワークカメラの機能構成等の一例を示す図である。ネットワークカメラ100は、撮像部・雲台110と制御部120とからなる。表示装置210等のクライアントは、カメラを遠隔制御可能である。通信I/F125は、HTTPやTCP/IPの送受信を担当する。全体制御部132は、ネットワークカメラ100を構成するモジュールの制御を行う。パン、チルト、ズーム制御要求は、通信I/F125及び全体制御部132経由で撮影制御部121に渡される。そして撮影制御部121は、撮像部・雲台110を制御する。

20

撮像部・雲台110から取得された映像は撮影制御部121経由で、映像入力部122によって取得され、映像符号化部123で符号化される。ここで符号化方式としてはJPEGやMPEG-2、MPEG-4、H.264等の方式がある。符号化された映像は映像送信部126で、HTTP、RTP等の通信方式に従い、通信I/F125を通じて送信される。

映像入力部122によって取得された映像は物体検出部127、イベント検知部128にも送信され、物体検出処理及びイベント検知が行われる。物体検出部127では制御パラメータを動的に変更可能な背景情報を利用した背景差分処理を用いて物体情報を検出する。物体検出部127が使用する背景差分処理については後述する。物体検出部127で検出される物体情報の例は、位置、面積、外接矩形、存在時間、静止度、物体領域速度、色成分、輝度成分等である。

30

イベント検知部128は、物体検出部127で検出された物体の状態から、特定画像領域への侵入や置去り等を検出する。イベント検知部128は、イベント検知処理を複数同時に行うこともできる。

#### 【0013】

追尾処理部129では検出した物体が視野の中央に位置するようにカメラの姿勢を変更する。シーン情報生成部130は、物体検出結果、イベント検知結果、追尾結果をシーン情報として生成し、シーン情報送信部131経由でクライアントに送信する。

表示装置210や録画装置220からは、映像配信要求、パン・チルト制御要求、イベント検知設定変更要求、追尾設定変更要求等の様々な要求が送信される。これは例えば、HTTPのGETメソッドや、SOAPを用いて送受信することができる。ここで通信I/F125は、主としてTCP/IPを担当する。そして、全体制御部132は、HTTPやSOAPの構文解析とカメラ内ユニットへの命令生成とを担当する。本実施形態は、図2の構成を実現するために特定のハードウェア構成に限定されるものではない。例えば、撮像部・雲台110、及び、制御部120は分離した構成とすることもできる。

40

また、制御部120内の各部等は、ハードウェアで実装されてもよいし、制御部120内のメモリ等に記憶されたプログラムを制御部120が実行することによって実現されるソフトウェアで実装されてもよい。

#### 【0014】

50

本実施形態では背景モデルを用いた判別による手法を説明する。ここでは観測画像の特徴量と背景モデルの保持する特徴量とを比較することにより物体を検出する。背景モデルは背景の変化に対応するため、モードと称する複数の状態を持つ。各モードは複数の画像特徴量やその統計量を背景の状態として保持する。また、 $8 \times 8$ 画素からなる画像ブロックの周波数成分絶対値和や、隣接フレーム間に対応する周波数成分の差分和を用いる手法もある。本実施形態の処理等は、特定の画像特徴に依存するものではない。

背景モデルを、ある時点  $t$  における、 $i$  番目のモードの特徴ベクトル  $M(i, t)$  として表す。物体の検出は、観測ベクトル  $I(t)$  と背景に属する背景モード  $M(i, t)$  との類似度  $I(t) \cdot M(i, t)$  が低い場合にこれを物体領域と判別する。背景モデルは新規画像が観測されると特徴量が類似したモードを探索し、類似モードの特徴量を更新する。これは観測画像の特徴量と、既存モードの特徴量とを一定比率で混合することによる。この混合比率を更新係数  $K$  とすると、モデルの更新は以下の式で表現される。

【0015】

$$M(i, t+1) = K * F(t) + (1 - K) * M(i, t) \quad (\text{数式 1})$$

$M(i, t)$  : 時刻  $t$  における  $i$  番目のモードの保持する特徴量ベクトル

$F(t)$  : 時刻  $t$  における観測画像の特徴ベクトル

$K$  : 更新係数  $0 < K < 1$

数式 1 において、 $K$  が大きい場合は短時間で背景モデルが更新される。特に  $K$  が 1 の場合は、直前のフレーム情報からのみ背景情報が生成されるため、隣接フレーム間差分に等しくなる。

【0016】

初期状態において、背景モデルの生成には一定時間を要する。これを背景モデルの初期化と呼ぶ。背景モデルの初期化にかかる時間は更新係数に依存し、例えば  $1/K$  フレームと設定できる。例として初期期間に観測される特徴ベクトルの平均値によって初期化する場合、初期背景モデルの生成は数式 2 のように表現できる。なお初期背景の生成時間中は物体検出を行うことはできない。

【0017】

$$M(i, T_{ini}) = 1 / T_{ini} * \sum_{t=1}^{T_{ini}} F(i) | t = 1 \dots T_{ini} \quad (\text{数式 2})$$

$T_{ini} = 1 / K$  初期背景生成時間 (フレーム)

【0018】

本実施形態における物体検出処理の手順を説明する。図 3 は、物体検出処理の手順を説明する図である。図 3 の (a) において、制御部 120 は、開始後 S301 で画像の取得を行い、次に S310 で周波数成分 (DC T 係数) を生成する。次に S311 で制御部 120 は、周波数成分から画像特徴を抽出する。S312 において制御部 120 は、S311 で生成した複数の特徴量が既存の背景モデルと合致するかを判別する。

原画像との比較は前述したように特徴量ベクトルの差分演算によって行われる。制御部 120 は、S312 では既存モードと比較して、もし類似モードが存在する場合は Y 分岐し、S314 に対応するモードの特徴量情報を更新する。

制御部 120 は、S313 において類似モードが存在しない場合には、N 分岐し、S315 に進んで新規モードを生成する。制御部 120 は、新規モードを生成する際に、既存モードのうち長時間使用していない古いモードを破棄する。制御部 120 は、S314、S315 終了後、S316 に進み、全てのブロックで処理が終了後、S319 に進んで物体の抽出処理を行う。

【0019】

図 3 の (b) の S321 から S325 までは物体抽出処理である。S321 は前景判別処理であり、制御部 120 は、各ブロックの状態が前景か背景かを判別する。これは入力した画像から抽出した特徴量と合致するモードが、前景か背景かを判別する処理である。ここで背景となるモードは、生成されてからの経過時間が、所定の閾値以上であるモードである。逆に前景となるモードは、生成されてからの時間が所定の閾値未満であるモード

10

20

30

40

50

である。

制御部 1 2 0 は、次に S 3 2 2 で前景ブロックの連結処理を行い、次に S 3 2 3 で小領域をノイズとして除去する。制御部 1 2 0 は、最後に S 3 2 4、S 3 2 5 で全ての物体に対して、物体情報の抽出を行い終了する。図 3 の方式によれば、背景モデルを逐次更新しながら安定して物体情報を抽出することができる。

#### 【 0 0 2 0 】

移動と観測とを繰り返す方式の追尾処理の場合、新規観測位置において迅速に物体領域を抽出することが求められる。しかしながら物体は移動しているため、速度によっては背景を生成する時間を十分に確保できない。そこで物体の速度に応じた背景差分処理の制御方法を図 4 を用いて説明する。

10

図 4 の ( a ) は背景モデルのモード数  $N_m$  が多く、更新係数  $K$  が小さい場合である。これは物体の移動速度が遅い場合に用いられる。この場合初期背景生成に時間がかかるが、生成された背景は一定時間のシーン情報を含むため、信頼度の高いものとなる。またモード数を増やすことによって、動的に変化する背景にも対応できる。一方で処理速度は低下する。また新規に現れた物体が背景化するまでに時間を要するため、速度の小さい物体であっても、物体領域情報を正確に抽出することができる。

図 4 の ( b ) は背景モデルのモード数  $N_m$  が少なく、更新係数  $K$  が大きい場合である。これは物体の移動速度が速い場合に用いられる。この場合初期背景生成には時間を要しないが、生成されたモデル特徴に含まれるシーン情報量は減ずるため、信頼度の低いものとなる。ここで更新係数  $K = 1$ 、モード数  $N_m = 1$  の場合は、毎フレーム背景が更新されることを意味する。これはフレーム間差分と同一である。この場合、検出した物体がすぐに背景化されるため、物体領域の内側は差分量がなく、検出されにくい。結果として物体の境界領域周辺が検出されるため、追尾位置を正確に決定できない。しかし高速に物体を検出可能であるため、速度の大きな物体に対応できる。

20

#### 【 0 0 2 1 】

次に図 5 を用いて、背景更新係数  $K$ 、背景モード数  $N_m$  をどのように制御すればよいかを説明する。図 5 は、物体速度に応じた背景差分パラメータの最適化を説明するための図である。初期状態では、画面上の物体位置は問わない。物体の移動速度は画像中の見かけ上の移動量である。図 5 の ( a ) に示すように、最初の物体速度は 2 回連続して物体を検出し、そのときの物体位置の差分から求めることができる。また M P E G - 2 や H . 2 6 4 符号化の過程で利用する動きベクトルの結果を解析して、類似の動きベクトル集合を求める手法によることもできる。これはパン、チルトの制御を伴わない画面内の物体位置の対応付け、という意味で画面内追尾と呼ぶ。

30

最初の追尾動作は求めた物体位置がカメラの中央に移動するように、パン、チルト方向のカメラ制御量  $P a n$ 、 $T i l t$  を以下のように設定することで実現できる。 $G p a n$ 、 $G t i l t$  は物体の大きさと速度から、移動後の物体が画面中に収まるように大きく取り過ぎないように設定する。

#### 【 0 0 2 2 】

$$P a n = G p a n * ( X o b j - X c )$$

$$T i l t = - G t i l t * ( Y o b j - Y c ) \quad ( 数 式 3 )$$

40

カメラの最初の追尾動作後に、物体速度に応じて背景差分による物体検出処理制御パラメータを変更する。図 5 の ( b ) は物体速度が小さい場合の処理例である。物体速度が小さい場合のモード数  $N_m$  を  $N_1$ 、背景更新係数を  $K_1$  とすると、 $N_1 > 1$ 、 $K_1 < 1$  である。ここで  $1 / K_1$  フレーム後に初期化が終了するので、物体検出処理を開始する。初期背景の生成後、更に  $1 / K_1$  フレーム待ってから、処理を行ってもよい。その結果、画面中で物体領域 5 1 1 を検出できる。したがって、上記物体位置に対して ( 数式 3 ) を適用し、カメラ制御を行えばよい。

#### 【 0 0 2 3 】

ここで背景生成時に物体が存在している領域は物体自体が背景化する場合がある。これは背景差分の性質上避けることができない。しかし実際には追尾後の物体は画面中央にあ

50

ることが想定されるため、物体が移動を続ける場合には、周辺部で検出が可能であり、結果として、検出ミスが生じる可能性は低い。即ち、画面中央の領域 5 1 0 で検出漏れが生じても、その周辺領域で物体領域 5 1 1 を検出可能である。

次に図 5 の ( c ) は、物体速度が大きい場合の例である。この場合のモード数を  $N_m = N_2$  に、背景更新係数  $K = K_2$  とすると、例えば、 $N_1 = 1$ 、 $K_1 = 1$  である。この場合フレーム毎に背景の持つ情報を全てを更新するため、フレーム間差分と同一であり、高速処理が可能である。即ち、追尾動作後にカメラが静止状態になり、2 フレーム目には物体検出結果を取得できる。但し、検出した物体領域は 5 2 1, 5 2 2 のようになり、物体の境界付近が検出される。追尾目標の定め方は、部分領域 5 2 1, 5 2 2 のうち面積の大きい領域の重心か、両者の重心位置の平均を定める方法がある。

10

物体が画面から外れる前に背景を作成する必要がある。したがって、 $K$  は画面上の物体速度  $V_{obj}$  と画像サイズとから、数式 4 のように決定できる。ここで  $G_k$  は更新係数  $K$  を決定するための係数であり、通常 1 より小さい正の値を設定する。 $N_m$  は処理速度が更新率  $K$  に影響しない最大の値を設定すればよい。

【 0 0 2 4 】

$$K = G_k / (S_{img} / V_{obj}) \quad (\text{数式 4})$$

$S_{img}$  : 画像サイズ (幅又は高さの平均)

【 数 1 】

$$V_{obj} = \sqrt{(X_{obj} - C_x)^2 + (Y_{obj} - C_y)^2}$$

20

$V_{obj}$  は画面上の物体速度

$G_k$  : 速度決定係数

【 0 0 2 5 】

以上の追尾方式をネットワークカメラ 1 0 0 上で実施する場合の、実行手順を図 6 に示す。図 6 は、追尾処理の処理手順を説明する図である。

制御部 1 2 0 は、開始後、S 6 0 1 で物体検出処理を行う。次に、制御部 1 2 0 は、S 6 0 2 で検出物体の有無を判別する。制御部 1 2 0 は、もし物体が検出された場合には Y 分岐し、S 6 0 3 で画面内追尾を行って物体の速度を計測する。次の S 6 0 4 で、制御部 1 2 0 は、パンチルト追尾を行うかどうかを判別する。これは一定以上の速度を持つ物体、或いは、視野外への移動が短期間に予想される物体に対して追尾を行うためである。

30

制御部 1 2 0 は、追尾を行う場合は Y 分岐し、S 6 0 5 でパンチルト追尾モードを開始して、S 6 0 6 でカメラの姿勢を変更する。続いて S 6 0 7 で、制御部 1 2 0 は、物体速度に応じて、既に述べた手法によって、背景モデルの更新係数  $K$ 、モード数  $N_m$  を決定する。更に S 6 0 8 で、制御部 1 2 0 は、再び物体検出処理を行う。次に S 6 0 9 で、制御部 1 2 0 は、追尾対象物体が停止したかどうか、或いは追尾対象物体が静止したかどうかを判別する。もし物体が移動を続けている場合には、制御部 1 2 0 は、S 6 1 0 に進んで追尾処理を継続する。もし物体の停止ないしは物体の喪失を検出した場合には、制御部 1 2 0 は、S 6 1 1 に進んでパンチルト追尾モードを終了する。

40

以上説明した手法によれば、背景差分処理によって画像中の動体領域を求めて、追尾処理を行う観測・移動方式の追尾方式において、物体の速度に応じて初期背景生成時間、背景更新係数、背景モード数を変更する。これにより、即応性と正確性とを両立する追尾を実現できる。

【 0 0 2 6 】

< 実施形態 2 >

別の実施形態としては、ネットワークカメラ 1 0 0 から表示装置 2 1 0 や録画装置 2 2 0 等のクライアントにライブ画像を送信し、クライアント側で背景差分処理による物体検出処理を行うこともできる。この場合、図 3 で説明した物体検出処理、及び図 6 に示した追尾制御処理をクライアント側で行い、クライアントからはカメラのパンチルト制御情報を送信する。この場合、図 6 の追尾制御フローにおいて、S 6 0 6 の P T 追尾はクライア

50

ントからカメラへのパン・チルト位置制御命令であり、カメラは受信した位置制御命令に従って、パン、チルトの制御を行う。

このような方式によれば、カメラ側の処理速度に制限がある場合にも移動物体の追尾を行うことができる。

【0027】

<実施形態3>

その他の実施形態としては、カメラ移動の前後で、重複する領域については移動する前の背景モデルを適用することもできる。この場合、カメラのパン、チルトに応じて背景モデルの位置をシフトする手法がある。

【0028】

<その他の実施形態>

また、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア(プログラム)を、ネットワーク又は各種記憶媒体を介してシステム或いは装置に供給し、そのシステム或いは装置のコンピュータ(又はCPUやMPU等)がプログラムを読み出して実行する処理である。

【0029】

以上、上述した各実施形態によれば、移動物体の速度に関係なく、長時間、より正確に移動体の追尾を行うことができる。

【0030】

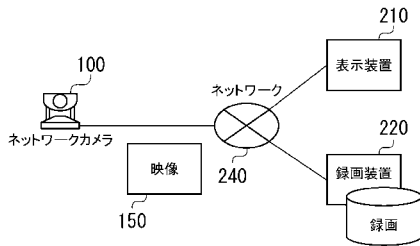
以上、本発明の好ましい実施形態について詳述したが、本発明は係る特定の実施形態に限定されるものではなく、特許請求の範囲に記載された本発明の要旨の範囲内において、種々の変形・変更が可能である。

【符号の説明】

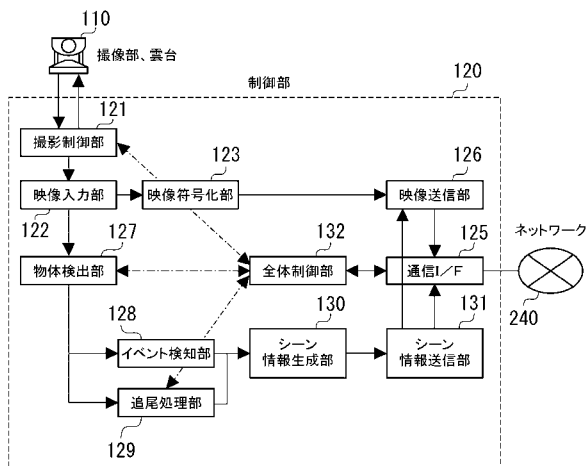
【0031】

120 制御部

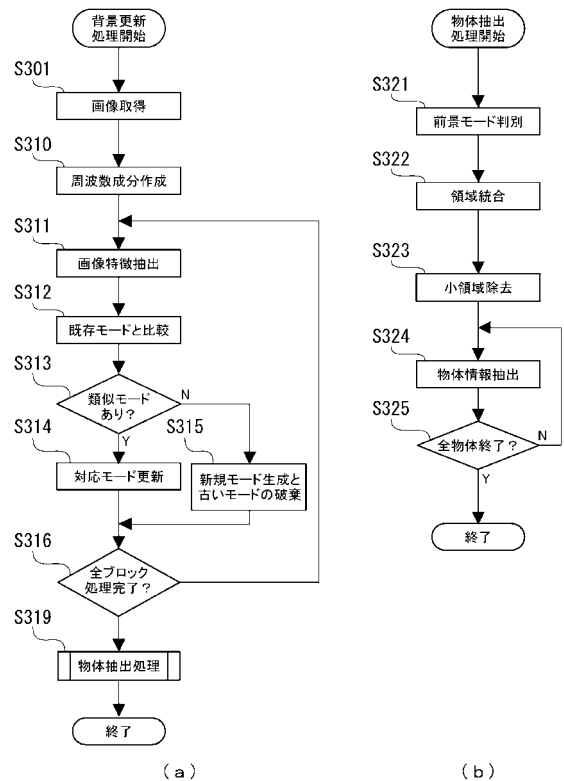
【図1】



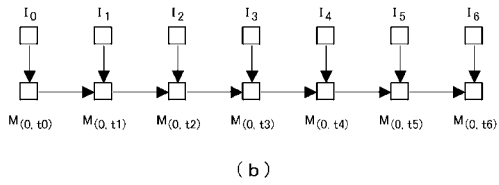
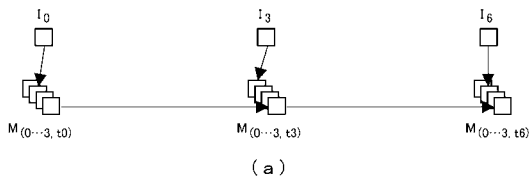
【図2】



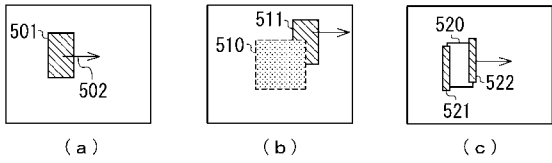
【図3】



【 図 4 】



【 図 5 】



【 図 6 】

