

移動物体像の抽出技術†

川 端 敦^{††} 谷 藤 真 也^{††} 諸 岡 泰 男^{††}

本報告で述べる処理装置は、背景のエッジ強調像を画像メモリ上に形成し、その画像を入力画像と比較することによって、高速に移動物体像を抽出する。背景像は、入力画像の不変な部分あるいはゆっくりと変化する部分をとらえることによって形成する。具体的には入力画像の各画素に関する輝度レベルの疑似積分（前回の出力値と今回の入力値の荷重平均を今回の出力値とする演算）を計算することによって得ている。この疑似積分には、低い周波数成分のみを通すローパスフィルタの働きがあるため、大きな輝度レベルの時間変化を引き起こす移動物体像による影響がおさえられ、結果として背景像が形成される。ところがそれだけでは数々の理由により正確な背景像が得られないため、抽出された移動物体像に残像に似たノイズが重畳される。そこで、より正確な背景像を得るため、「背景像をとらえ続けている画素の、輝度レベルの時間変化は小さい」という点に着目した工夫を施してある。つまり入力画像の輝度レベルの時間変化が少ない画素は背景をとらえているものとみなし、背景の更新を行う。輝度レベルの時間変化が大きい画素は移動物体をとらえている可能性が大きいので背景の更新は行わない。この方法を用いた結果、移動物体像の影響をあまり受けない、正確な背景像を構成することができ、したがってより正確に移動物体像を抽出することができる。

1. はじめに

マニピュレータ、ロボット等の発達により近年特にFA (Factory Automation) に関する研究が多くなされている。これは高度に知能化された生産設備に関する研究と言い換えることが可能であるが、このような生産設備は高機能な手足、あるいは触覚を得て初めて実現が可能である。

現在、小型高出力アクチュエータの開発、マイクロプロセッサを使用した制御法の確立等が原動力となつて、マニピュレータ、ロボットの駆動部、あるいは駆動方法に関しては、かなりの成果が得られている。それに比べると各種センサ、特に視覚センサから入力された情報の利用に関する研究はやや遅れており、今後FA化を進めるにあたり大きな障害になると予想されている。

視覚センサは、光が遮られたことを判断する、フォト・トランジスタを使用した簡単なものから、撮像素子を使用したものまで各種あり、最近ではCCD撮像素子、あるいはMOS撮像素子に直接処理回路を付加したスマート・センサと呼ばれているものも考案されている。これらのセンサから入力された信号は、多くの場合デジタル変換され、計算機によって処理される。

センサとして撮像素子を選んだ場合、処理装置に入

力される情報は、膨大な量となるが必要としている情報は、例えばある物が見えているか、いないか、あるいは、あらかじめ記憶してある物のうちのどれが見えたか、といったわずかな情報であることが多い。処理装置は画像情報から上で述べたような情報を抽出しなければならず、そのため極めて大きな処理能力が必要とされる。以前は、このような能力を供えた装置といえば、大型計算機、ミニ・コンピュータ、あるいは高価な専用ハードウェアしかなかったが、ME (Micro Electronics) の発達により、現在極めて高性能で比較的安価な、デバイスが開発されつつある。このようなデバイスは安価なため大型計算機とは根本的に異なった使用法が可能となり、画像処理分野への応用がますます盛んになると予想される。したがって、このようなデバイスを用いることを前提とした画像処理技術の研究は、今後盛んになるものと、予想される。

すでに種々の視覚情報処理技術がFAやロボットの分野で広く用いられているが、一般に画像処理には多くの手間を要するため、大半のシステムでは理想的な照明条件下における静止画しか扱うことができなかつた。本研究は、動画像に対する新しい画像処理装置、特に移動物体像を抽出する処理装置を実現することを目的としている。

2. 移動物体像の抽出方法

移動物体像の抽出方法は、次に示す3種類の方法がよく用いられている。

まず第1に、連続して入力される画像の各画素に関

† An Image Extraction Method for Moving Object by ATSUSHI KAWABATA, SHINYA TANIFUJI and YASUO MOROOKA (Hitachi Research Laboratory, Hitachi, Ltd.).

†† (株)日立製作所日立研究所

する輝度レベルの時間微分あるいは類似の演算により得られた値を利用して処理を行う方法である¹⁾。背景像の輝度レベルは変化しないため、時間微分は小さな値を示すが、移動物体像の輝度レベルは、一般に大きく変化するため、時間微分の絶対値は大きな値を示す。そこで時間微分の絶対値が大きい部分だけを抽出することによって移動物体に関する情報を得ることができる。電子回路を用いた時間微分あるいは類似の演算を行う装置は、構成が簡単であり、処理速度も高速である。しかしその演算によって得られた値のみから、移動物体の像を抽出することは困難である。例えば、無地の背景の前で円形の物体が移動した場合、進行方向の前部と後部とに符号の異なった三日月様の領域が、出力されるがこの情報から円形の領域を抽出することは、困難である。

第2の方法としては、連続画像を記憶装置にいったん蓄えて、次に記憶装置から一枚ずつ画像を取り出し、各画像間の対応点を算出するという方法が知られている²⁾。この方法は、画像をセグメントに分割し、それぞれのセグメントを特徴付ける形状パラメータを算出し、画像間でこの形状パラメータの相関をとることによって対応点を算出する。この方法では対応点が見つかる場合には、その移動量を正確に算出できるので、移動物体像を正しく抽出することができるが、その処理に要する時間は長くなる。また、物体の回転、変形等がある場合には、対応点を検出することが難しくなる。

第3の方法は、装置内部に背景像に相当する画像を

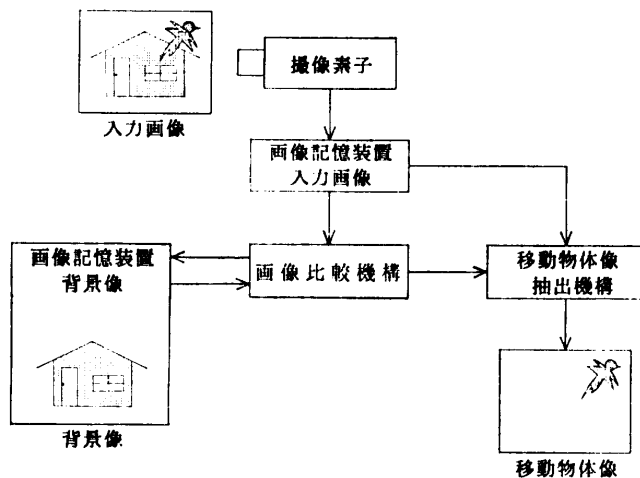


図1 移動物体像を抽出する装置の構成図

Fig. 1 Basic image processing system for extraction of a moving object.

蓄えておき、その画像と入力画像とを各画素ごとに比較することによって移動物体像を抽出する、というものである。本論文で示す方法では、この第3の方法に基づくものである³⁾⁻⁴⁾。入力画像には移動物体像と背景像が含まれており、この画像から背景像を引けば、移動物体だけに関する情報が出力される、という考え方である。この方法を用いた移動物体像抽出装置の構成を図1に示す。固定した撮像素子は静止している背景と移動物体の像をとらえる。図1では家が背景であり、鳥が移動物体である。撮像素子の出力は、A/D変換されて画像記憶装置に蓄えられる。画像比較機構は各画素ごとに入力画像の輝度レベルから装置内部に記憶してある背景像の輝度レベルを差し引き、その値を基にして移動物体に関する情報を出力する。入力画像において背景が占める部分は、一般には差分演算結果が0となり、移動物体像が占める部分の差分演算結果は正あるいは負の大きな値を示す。したがって差の絶対値が0よりも大きな画素あるいはノイズが重畳されていることを考慮して、差の絶対値がある値よりも大きな画素を移動物体像の構成画素と判断してその結果を出力すればよい。移動物体像抽出機構は、画像比較機構から出力された情報に基づき、入力画像から移動物体像だけを切り取り、出力する機構である。この方法で問題となるのは、背景像の形成に関する点である。正しいと思われる背景像をあらかじめ装置内部に入力して固定しておくことも考えられるが、背景が変化した場合には正しく処理を行うことができない。そこでこの方法を用いる場合には背景像を適応的に形成する必要がある。そこで図1の画像比較機構は入力画像を基にして、装置内部に形成した背景像を更新する働きも有する必要がある。

この方式では、いったん正確な背景像を形成することができれば、その後の処理で特に問題となる点はない。本論文は、この背景像の形成方法に関して新しい方法を提案するものである。

3. 背景像の形成方法

本文では、「背景とは撮像素子に対して全く移動しないあるいは変化しない物体か、あるいは移動したり変化したとしてもその速度は極めて小さい物体を指す」と定義する。この定義に従えば背景像は全く変化しないか、あるいは移動物体像に比較してゆっくりと変化するものと仮定することができる。そこで入力画像の不変な部分、あるい

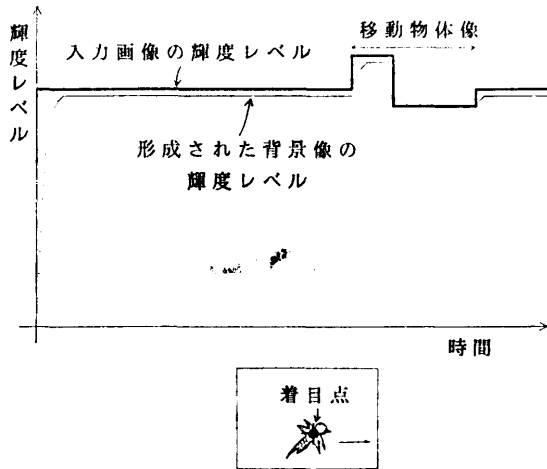


図 2 (1)式により背景が形成される様子
Fig. 2 Typical case of constructing the background image.

はゆっくりと変化する部分をとらえることにより、背景像を形成する。形成された背景像のことを以下では推定背景像と呼ぶ。具体的には、(1)式に示すように入力画像の各画素に関する輝度レベルの疑似積分(前回の出力値と今回の入力値の加重平均を今回の出力値とする演算)を計算することによって背景像を得る。この疑似積分には、低い周波数成分のみを通すローパスフィルタの働きがあるため、大きな輝度レベル変化を引き起こす移動物体像による影響が押さえられ、結果として背景像が形成される。

$$BK(x, y, t) = \frac{(2^n - 1) \cdot BK(x, y, t - 1) + I(x, y, t)}{2^n} \quad (1)$$

$BK(x, y, t)$: t 時点における座標 (x, y) の形成された推定背景像に関する画素の輝度レベル

$I(x, y, t)$: t 時点における座標 (x, y) の入力画像に関する画素の輝度レベル

n : あらかじめ設定した数 ($n \geq 0$)

(1)式右辺分子の第一項は推定背景像を保持する特性を示す項であり、第二項は推定背景像を更新する特性を示す項である。 n は更新の速度を示す項であり、 n を大きくすると更新の速度は遅くなるがノイズの影響を受けにくくなり、 n を小さくすると更新の速度は速くなるがノイズの影響を受けやすくなる。ここで言うノイズとは「移動物体像」のことである。図 2 に(1)式を用いた場合

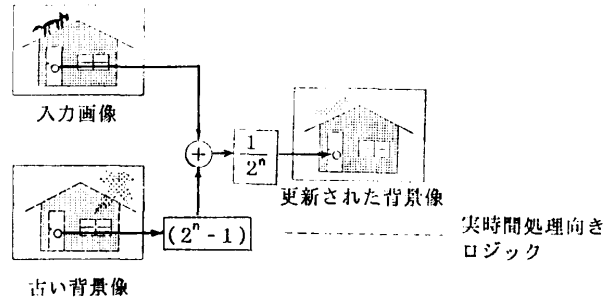


図 3 背景更新機構
Fig. 3 Hardware implementation of revising the background image.

の輝度レベルの変化の様子を示す。グラフは入力画像のある画素に関する輝度レベルの時間変化とその画素に対応して形成される推定背景像の輝度レベルの時間変化を示している。太線は入力される画像の輝度レベルであり、一定値を保っている部分は背景像をとらえ続けており、乱れている部分はその瞬間に移動物体像をとらえたことを示している。細線は装置内部に形成される推定背景像を示している。初期値は 0 であるが、時間とともに一定値(背景像)に収束している様子が分かる。この演算を実行する装置は、図 3 に示すように加減算と 2^n の乗算あるいは除算のみで構成することができる。 2^n の乗算、除算は電子回路を用いた場合シフト演算に置き換えることが可能であるから全体としてこの演算は、高速演算が可能となっている。

(1)式によって推定背景像が形成される様子を模式的に図 4 に示す。(A1), (A2), (A3) は入力される画

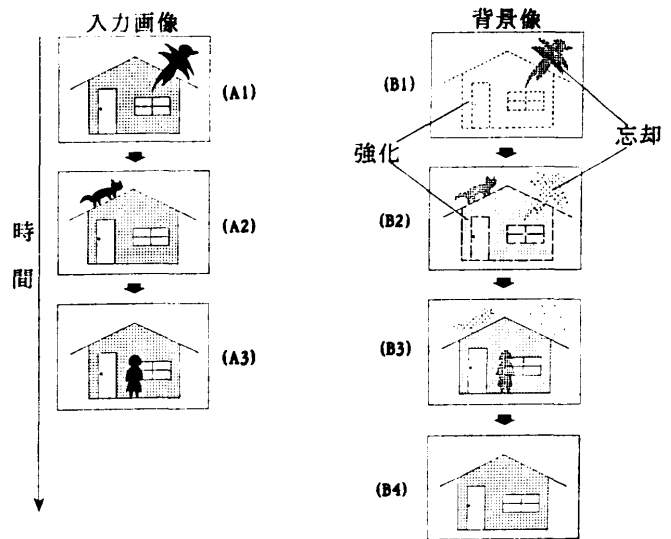


図 4 背景像の形成モデル
Fig. 4 The process of constructing the background image.

像の系列であり、(B1), (B2), (B3), (B4)は推定背景像が次第に形成される様子を示している。最初、背景としては何も記憶されていない状態であると仮定する。鳥と家を含む画像(A1)が入力されると、その像に関する情報が(B1)のように少しだけ推定背景像として記憶される。次に猫と家を含む入力画像(A2)が入力されると、推定背景像においては(B2)のように引き続き存在している家の像は強調され、猫の像が少しだけ記憶され、今回は存在していない鳥の像が弱まっている。同様に人と家を含む像(A3)が入力された場合、(B3)のように家の像はさらに強く記憶され、他の像は家の像に比べて弱く記憶される。最終的には背景像として(B4)に示すような家の像のみが形成される。

4. より正確な推定背景像の形成方法

(1)式の入力画像 $I(\cdot)$ は背景と移動物体の像により構成されているが、移動物体の像は推定背景像を形成するという立場からすると単なるノイズにすぎない。つまり一度形成された推定背景像は、移動物体像の影響を全く受けず真の背景の輝度レベルを保持し続けるのが望ましい。ところが n の値が小さい場合には、形成された推定背景像が移動物体像の影響を強く受けて図2に示すように乱れる。装置内部の推定背景像が乱れると、この推定背景像を基にして算出される移動物体像にノイズが重畳される。逆に n の値を大きくした場合には形成された推定背景像は、図5に示すように移動物体像の影響を受けにくくなるが、この場合には別の原因によるノイズが新たに重畳される。

(1)式は図3に示すような整数演算機構によって実行されるため、 2^n による除算の結果から小数点以下

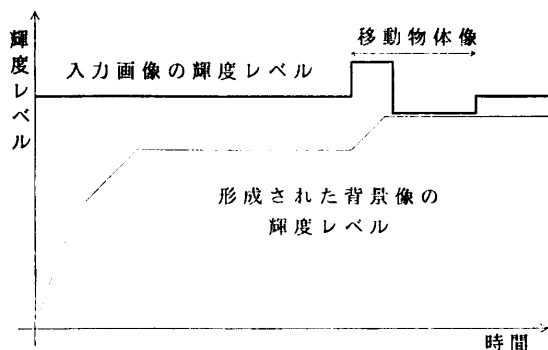


図5 n が大きい場合に背景が形成される様子
Fig. 5 Typical case of constructing the background image with large n .

が切り捨てられる。したがって(1)式を厳密に表すと次式のように表現される。

$$BK(x, y, t) = \left[\frac{(2^n - 1) \cdot BK(x, y, t-1) + I(x, y, t)}{2^n} \right] \quad (2)$$

$BK(x, y, t), I(x, y, t), n$: (1)式と同じ

$[m]$: m を越えない最大の整数

(2)式より下式を満足する整数 Bias が必ず存在する。

$$2^n \cdot BK(x, y, t) = (2^n - 1) \cdot BK(x, y, t-1) + I(x, y, t) - \text{Bias} \quad (3)$$

Bias: (2)式より定まる数 ($0 \leq \text{Bias} \leq 2^n - 1$)

入力には移動物体像は含まれず背景像のみが与えられ、しかも十分に時間が経過して装置内部の推定背景像 $BK(x, y, t)$ がある値 $BK(x, y)$ に収束したと仮定する。

この場合(3)式は次式に示すとおりになる。

$$2^n \cdot BK(x, y) = (2^n - 1) \cdot BK(x, y) + I(x, y) - \text{Bias} \quad (4)$$

$BK(x, y)$: 座標 (x, y) の形成された推定背景像に関する画素の輝度レベル

$I(x, y)$: 座標 (x, y) の入力画像に関する画素の輝度レベル

したがって

$$BK(x, y) = I(x, y) - \text{Bias} \quad (5)$$

となる。上式によると、(1)式によって形成された推定背景像には、各画素に最大 $(2^n - 1)$ のノイズが重畳される可能性のあることがわかる。したがって n が大きくなると Bias が大きくなり各画素に多くのノイズが重畳される。図5によると n が大きい場合にはオフセット状のノイズが多く重畳されている様子が分かる。以上をまとめると、 n が小さい場合には、形成された推定背景像が移動物体像の影響を強く受け、 n が大きい場合には(5)式によるノイズが重畳される。つまり(1)式をそのまま用いたのでは、 n を大きくしても小さくしても形成された推定背景像がノイズの影響を受けることになる。

そこで入力画像 $I(\cdot)$ に含まれる移動物体像が形成された推定背景像 $BK(\cdot)$ に影響を与えないような工夫をした。その結果 n を小さくすることができ、 $BK(x, y, t)$ に重畳されるノイズ量も減少させることができ、より高精度な推定背景像が得られた。以下に新たに開発した輝度レベル差検出法と呼ぶ工夫について述べる。

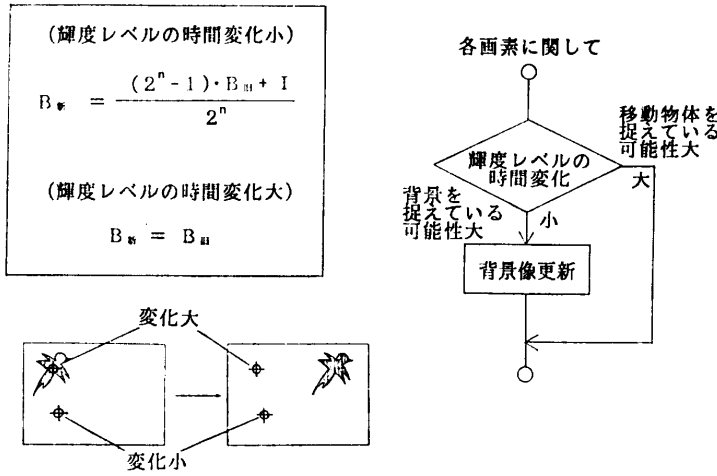


図 6 輝度レベル差検出法の原理
 Fig. 6 Principle of the intensity-difference detecting method.

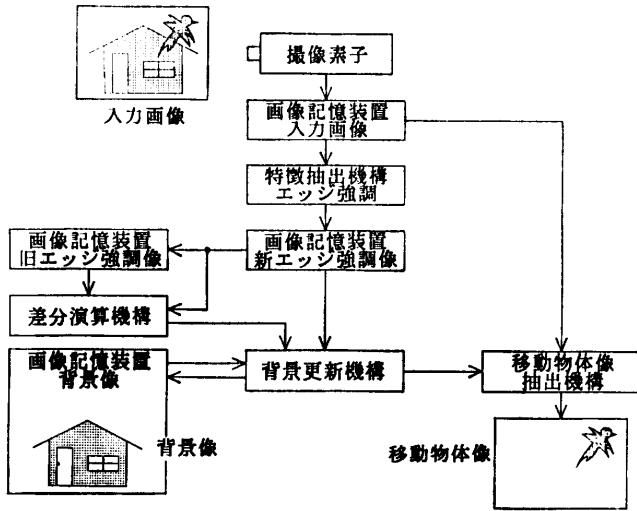


図 7 輝度レベル差検出法を用いた装置の構成図
 Fig. 7 Image processing system for extraction of a moving object with using the intensity-difference detecting method.

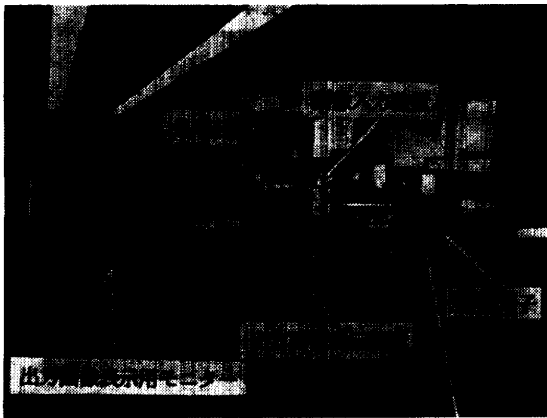


図 8 装置全景
 Fig. 8 Panoramic view of the system.

入力画像 $I(x, y, t)$ には移動物体像と背景像が含まれる。背景をとらえている画素の輝度レベルに関する時間変化は小さいが、移動物体をとらえている画素の時間変化は一般に大きい。つまり相続く2組の入力画像 $I(x, y, t)$ の画素ごとの差分を計算すると、背景をとらえ続けている画素の差分値は小さな値を示し、移動物体をとらえた画素の差分値は大きな値を示す。そこで図6に示すようにその差分値の小さな画素は、背景をとらえているものとみなし、入力値を利用して推定背景像を更新するが、差分値の大きな画素は、移動物体をとらえているものとみなし、推定背景像を更新しない。この方式を用いると、移動物体像が形成された推定背景像に与える影響を抑えることができるので、 n を小さくすることができる。結果として移動物体像の影響を受けにくく、また(5)式によるノイズの影響も少ない推定背景像を形成することが可能となる。

5. 移動物体像を抽出する装置

図7に装置の全体構成図を、図8に装置の全景を示す。撮像素子は一般に使用されているI.T.V.(MOS撮像素子)を用いた。撮像素子により入力された画像は1/60秒ごとに128×128画素、各画素に関して64階調のデジタル信号に変換され画像記憶装置に蓄えられる。

特徴抽出機構は図9に示すような空間フィルタを入力画像に作用させてエッジの強調を行う。一般に、物体像を表示した場合、もし輪郭線が完全に表示されなかったならば、物体像は著しくみにくいものとなり、その程度によっては認識が困難になることが知られている。したがって画像処理後の移動物体像から輪郭線が脱落することを防止するために、エッジを強調する必要がある。また撮像素子を通して入力した画像は、フレーム間の同期信号のずれによりしば

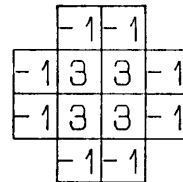


図 9 エッジ強調用空間フィルタ
 Fig. 9 Spatial filter for enhancement of edges.

しば画像が1画素程度ずれたり、あるいは孤立した1画素に各種の要因による非定常ノイズが重畳される。図9に示した空間フィルタは物体像のエッジを強調すると同時に輪郭線内部の情報も出力し、かつ大きさが1画素程度であるノイズを除去する機能がある。この空間フィルタを作用させる演算は画像処理装置を用いることにより、1/60秒で処理される。

入力画像のエッジ強調像はいったん画像記憶装置に蓄えられたのちに、もう一つの画像記憶装置に転送される。したがって本装置には入力画像のエッジ強調像を蓄える画像記憶装置が新旧の2組あることになる。(本装置では、実際には画像記憶装置間でのデータ転送は行わずに、2組の画像記憶装置を切り替えて新エッジ強調像用、あるいは旧エッジ強調像用に使用している。)

輝度レベル差検出機構はこれら2組の画像記憶装置の各画素に関する差を計算しその差が定めた値より小さな画素だけが背景のエッジ強調像の更新に使用されるよう背景更新機構に指示を出す。つまり相続く2枚の画像を比べた場合輝度変化の大きな画素は背景ではなくそれとは逆に輝度変化の小さい画素は背景を構成している画素と判断しその画素データを使用して背景の更新を行うように指示する。この各画素ごとに更新の要、不要を決定する演算は1/60秒周期で処理される。

背景更新機構は、輝度レベル差検出機構から指示のあった画素に関してのみ(1)式を作用させそれ以外の画素に関しては更新を行わない。また、入力画像のエッジ強調像と背景のエッジ強調像の輝度レベルを各画素ごとに比較し、その差が定めた値よりも大きい画素だけが移動物体像の抽出に使用されるように移動物体像抽出機構に指示する。この推定背景像を更新する演算は1/30秒周期で処理される。移動物体像抽出機構は、背景更新機構から指示のあった画素を入力画像から切り取る。

実験装置は、コントローラ部と、画像処理用ハードウェア部と、画像入出力部に分かれる。コントローラ部は8MHzの8086を内蔵しており、画像処理用ハードウェア部の制御、およびデータのコントロールを行う。画像処理用ハードウェア部は、ISPを4個含んでおり、本装置の中で中核的な役割を果たしている。200.66mm(width)×250.00mm(height)のボード2枚で構成している。画像入出力部は256×256画素64階調の画像を1/60秒周期で記憶装置に取り込んで表

示できる。以上により構成された装置は128×128画素の画像を1秒間に約10枚処理することが可能であった。

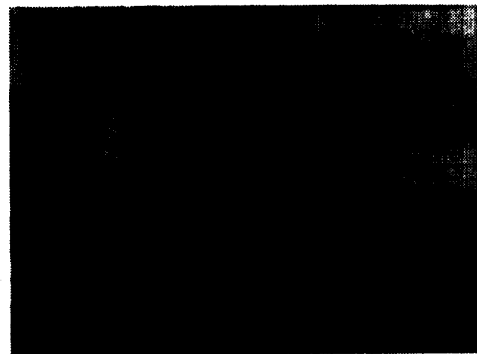
6. 実験結果

実験条件は次に示すとおりである。実験は自然な採光の下で行った。移動物体としては電池で移動する玩具を使用し、机上で動かした。背景としては室内の風景をそのまま用いた。なお(1)式における n の値は2と設定した。

装置に入力した像は図10に示すような像である。装置内部の推定背景像を蓄えるメモリは実験開始前にあらかじめクリアしておいた。図11には全画素に関して常に推定背景像を更新する手法を用いた場合の出力結果を示す。4枚の画像はそれぞれ処理を開始してから1秒後、3秒後、4秒後の画像であり、最後の画像は10秒後の画像である。実験開始直後には装置内部に推定背景像が形成されていないため、入力画像全体が移動物体像であるものと認識されて出力されてい



移動物体



背景

図10 入力画像
Fig. 10 Input image.

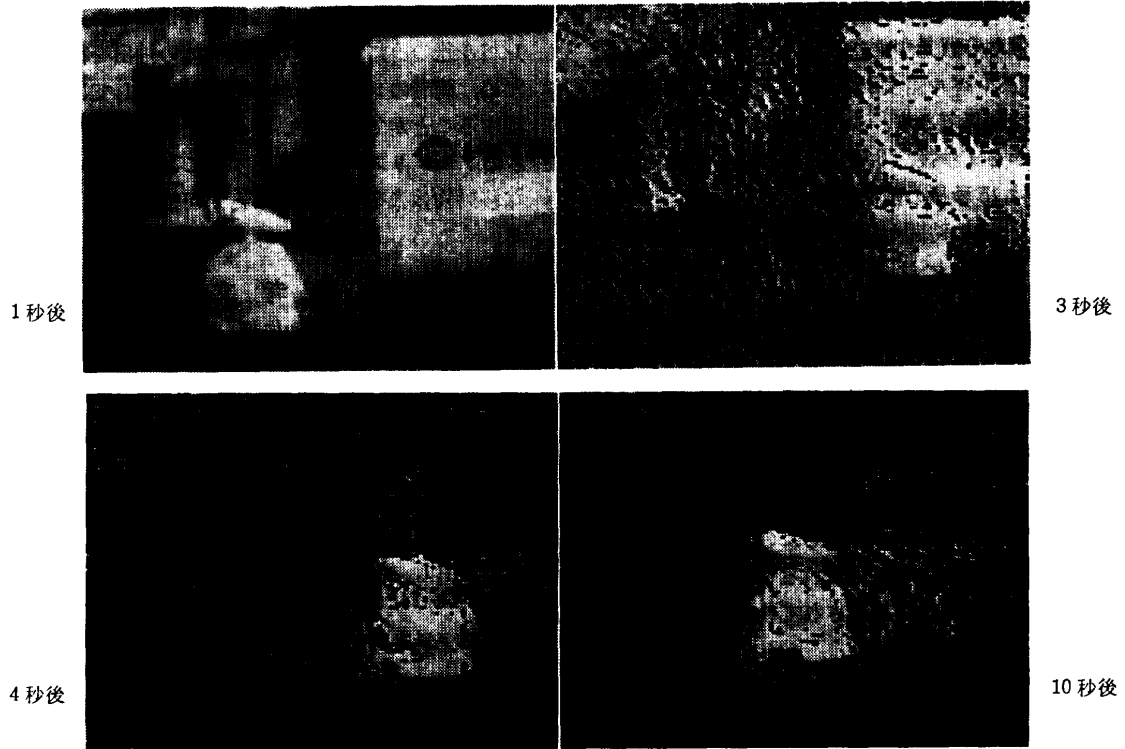


図 11 実験結果
Fig. 11 Results.

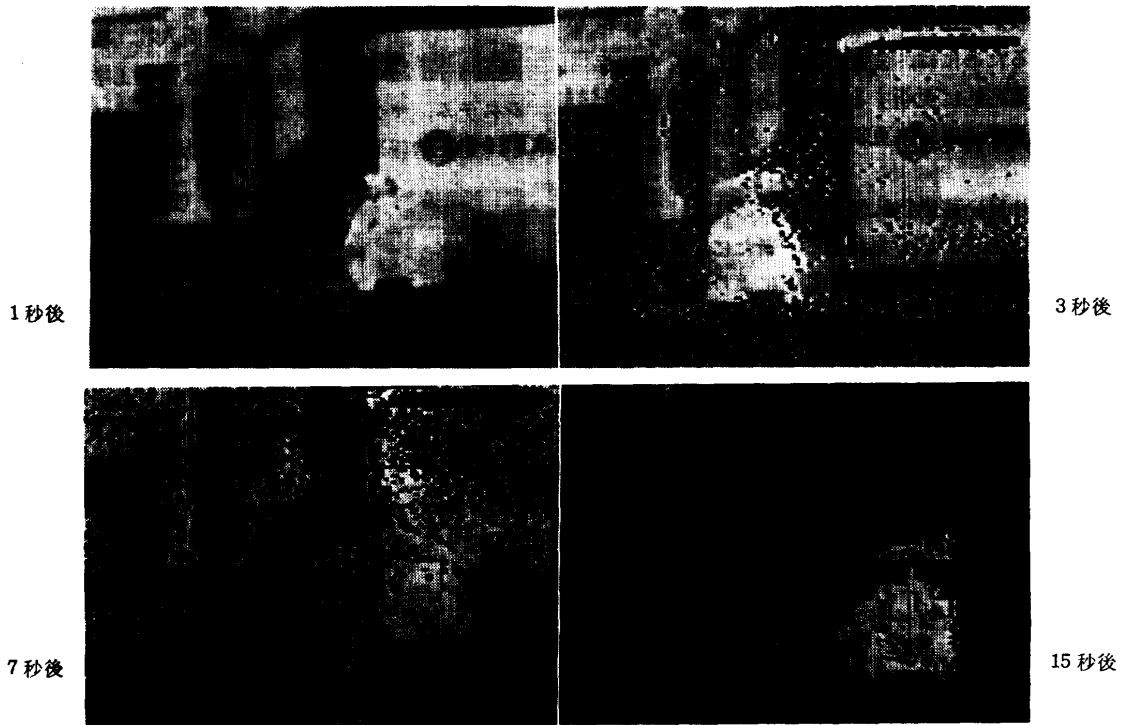


図 12 輝度レベル差検出法を用いた実験結果
Fig. 12 Results of the intensity-difference detecting method.

る。時間の経過とともに次第に推定背景像が形成されて、出力画像から背景が消えていく様子が分かる。図12には、輝度レベル差検出法を用いて推定背景像を更新した実験結果であり、4枚の画像はそれぞれ処理を開始してから1秒後、3秒後、7秒後の画像であり最後の画像は15秒後の画像である。

図11では十分に時間が経過したのちでも、移動物体の通過したあとには多くのノイズが重畳されている。これは推定背景像が移動物体像によって乱されるために生ずると考えられる。一方図12に示した実験例は、出力画像から背景が消えるまでに多くの時間を要している。これは入力画像に多くの非定常ノイズが重畳されているために、輝度レベル差検出機構が働き、背景更新式を作用させないためだと考えられる。しかしながら十分に時間が経過したのちには、推定背景像が移動物体像によって乱されることがないために、正確な移動物体像が抽出されている。

7. おわりに

装置内部に背景像に相当する画像を蓄えておき、その画像と入力画像とを各画素ごとに比較することによって移動物体像を抽出する画像処理に関して、次に示すような結論を得た。

- (1) 背景が変化した場合でも、正しく移動物体を抽出するためには、装置内部に蓄えておく推定背景像を適応的に更新する必要がある。(1)式により、背景を適応的に更新することが可能となった。
- (2) より良い移動物体像を得るためには、正確な背景像を形成する必要がある。新たに開発した輝度レベル差検出法を用いることにより移動物体像の影響を受けにくく、また(5)式によるノイズの影響も少ない推定背景像を形成することが可能となった。
- (3) 移動物体像としては、輪郭線の脱落が少ない中抜けのない移動物体像を形成する必要がある。また入力画像に重畳されているノイズを除去する必要がある。図9に示したエッジ強調フィルタを用いることによって物体像のエッジを強調すると同時に輪郭線内部の情報も出力し、かつ大きさが1画素程度であるノイズを除去することができた。

以上の結論に基づき実験装置を試作しその有効性を確かめた。

なお、本方式を用いると推定背景像と入力画像とのレベル差のみに着目して移動物体の抽出を行っているため、両者のレベル差がない画素に関しては正確に処理することができない。しかし輪郭とその内部を認識するアルゴリズム等を併用することによってより正確な移動物体像を抽出することが可能となると考えられる。

参 考 文 献

- 1) 辻 三郎：動画像処理，情報処理，Vol. 21, No. 6, pp. 633-638 (1980).
- 2) 川端，谷藤，諸岡：移動物体像の抽出技術，第32回情報処理学会全国大会論文集，pp. 1441-1442 (1986).
- 3) 川端，谷藤，諸岡：移動物体像の抽出技術，情報処理学会，コンピュータビジョン研究会資料，34-5 (1985, 1).
- 4) 山口，宮沢：視覚機能の機械化，自動化技術，Vol. 14, No. 6, pp. 30-36 (1982).

(昭和61年8月4日受付)

(昭和62年1月14日採録)



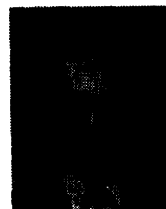
川端 敦

昭和32年生。昭和56年東京大学工学部計数工学科卒業。昭和58年同大学大学院修士課程修了。同年(株)日立製作所入社。現在日立研究所において生体情報工学，分散処理システム，マン・マシンインタフェースの研究に従事。電子情報通信学会，計測自動制御学会各会員。



谷藤 真也 (正会員)

昭和22年生。昭和48年早稲田大学大学院修了。同年(株)日立製作所日立研究所入社。鉄鋼関係の制御システムの研究・開発に従事。現在は知識処理応用およびワークステーションの開発を担当。電気学会，計測自動制御学会各会員。



諸岡 泰男 (正会員)

昭和17年生。昭和40年九州工業大学制御工学科卒業。同年(株)日立製作所入社。以来、日立研究所にて鉄鋼等プラントの計算機制御システムの開発に従事。現在に至る。主として、モデリング，情報処理，大規模分散システムに興味を持つ。電気学会，計測自動制御学会各会員。