

室内監視画像からの物体領域の検出と移動履歴の管理

Detection of Object Regions in Indoor Surveillance Images and Recording of Their Locations

藤井朝子[†], 正会員 目黒光彦[†], 正会員 金子正秀[†]

Asako Fujii[†], Mitsuhiro Meguro[†] and Masahide Kaneko[†]

あらまし 室内に設置された監視カメラからの時系列画像を用いて、物体の移動履歴を記録する手法を提案する。位置変化のあった物体領域の検出、ズームアップ画像を用いての物体認識および物体の移動履歴の記録手法について述べ、実験結果を示す。

キーワード：物体領域検出, 時系列画像, 背景差分, 静的変化領域, 室内画像, 移動履歴

1. ま え が き

日常生活において、モノ(物体)をどこに置いたかわからなくなることは誰もが経験したことがある。本論文では、室内に置かれたモノの位置を自動的に取得することによって、モノの位置をユーザに示すことができるシステムについて検討する。関連した研究としては、ユーザの記憶想起の支援を目的として、ユーザの視点に相等するカメラの映像を用いて、「置く」、「拾う」といった動作を検出し、その動作映像を行為履歴として記録する方法¹⁾²⁾などがある。しかし、物体の認識を行っていないため、どの物体についての「行為」なのかは人間が判断しなければならない。

本論文では、ユーザの視点に相等したカメラの映像ではなく、監視シーン³⁾のような室内に設置したカメラを用いて、物体が「いつどこに置かれたか、どこからなくなったか」という移動履歴を記録するシステムについて提案する。置かれた物体や持ち去られた物体の位置を自動的に取得し、物体の認識を行うことで個々の物体の位置情報を記録し、それを用いてユーザにモノの位置を示すことを目的とする。

2. 全体の処理の流れ

本論文では、図1に示すように室内の上方に設置した2台の監視カメラからの時系列画像を用いる。一台のカメラは対象とするシーン全体を常に監視するため、もう



図1 カメラの設置例
An example of camera setting.

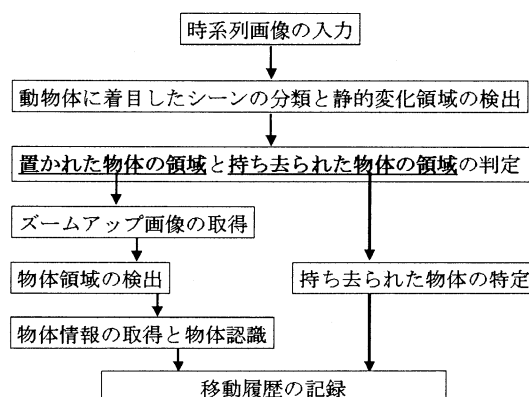


図2 全体の処理の流れ
Overall framework of the proposed method.

一台のカメラは、置かれた物体をズームアップし、物体認識に用いる物体情報をより詳細に得るためのカメラとして用いる。

全体の処理の流れを図2に示す。動物体のように変化が動的ではなく静的なので、置かれた物体や持ち去られた物体の領域を「静的変化領域」と呼ぶ。

2003年2月28日受付, 2003年5月26日再受付, 2003年6月6日採録
[†]電気通信大学 大学院電気通信学研究科

(〒182-8585 調布市調布ヶ丘1-5-1, TEL 0424-43-5216)

[†]Graduate School of Electro-Communications, The University of Electro-Communications

(1-5-1, Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan)

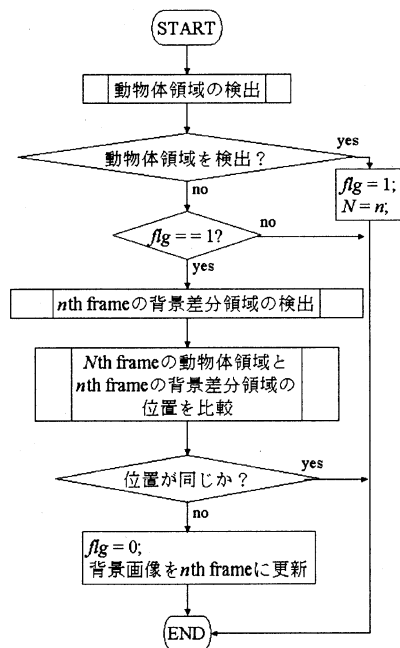


図 3 nth frame の処理の流れ
Processing of the nth frame.

3. シーンの種類と静的変化領域の検出

3.1 動物体に着目したシーンの分類

時系列画像における変化には、動物体 (主に人物) による変化と置かれた物体などによる静的変化がある。2 種類の変化を区別するために、動物体による変化領域の検出を行うことで、動物体に着目したシーンの分類を行う。ここで、動物体による変化領域としては、動いている動物体だけではなく、シーン内で静止している動物体も含めた動物体による変化領域とし、静止物体による変化領域とは区別する。動物体による変化領域が検出されるシーンは動物体を含むシーン、検出されないシーンは動物体を含まないシーンと分類する。

動物体に着目したシーンの分類の流れを図 3 に示す。図 3 中の fg が 1 のときは、nth frame が動物体を含むシーンに、0 のときは含まれないシーンに分類される。N は最後に動物体領域が検出されたときのフレーム番号を格納する変数である。図 3 中の動物体領域は実際に動いている物体領域であり、動物体による変化領域とは異なる。nth frame の動物体領域の検出は、(n-1)th frame と nth frame の差分領域と、nth frame と (n+1)th frame の差分領域との論理積をとった領域とする。動物体による変化領域は、動物体領域が検出されたときは動物体領域と同じ領域とし、動物体領域が検出されなくても、動物体がシーンに留まっているときは、以前に検出された動物体領域と同じ領域とする。

3.2 シーンの切り替り時における静的変化領域の検出

動物体がシーンの外へ移動したと判断されたときをシーンの切り替り時と呼ぶこととする。また、動物体が

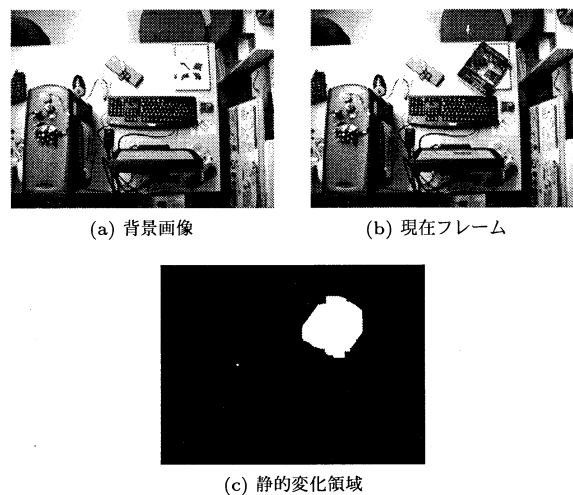


図 4 シーンの切り替り時における静的変化領域の検出
Detection of static change regions at a scene change.

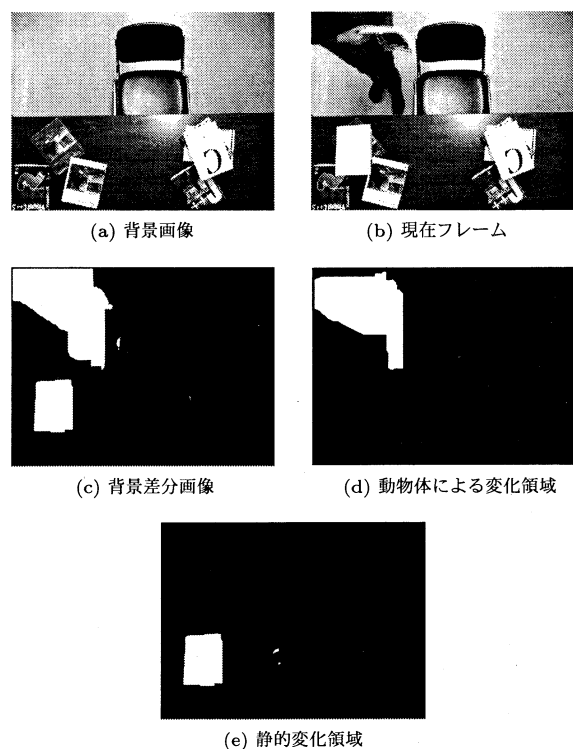


図 5 動物体を含むシーンでの静的変化領域の検出
Detection of static change regions in a scene including moving objects.

シーン内に入る直前のシーンの画像を背景画像とする。背景画像は、シーンの切り替りの都度更新されていく。動物体がシーンの外へ移動したときのフレームの画像と背景画像との差分を検出することで、動物体がシーン内に置いた物体や持ち去った物体による変化領域を検出することが可能である。図 4(a), (b) に、各々背景画像、現在フレームを示す。また、背景差分が静的変化領域の検出結果になり、図 4(c) に示す。

3.3 動物体を含むシーンでの静的変化領域の検出

3.2 の方法では動物体がシーン内で物体を次々に重ねてからいなくなったときなどに、置かれた個々の物体領

域を検出することができない。このため、動物体がシーン内にいる間も静的変化領域を検出する必要がある。

動物体を含むシーンでの静的変化領域の検出方法を説明する。まず、背景画像と現在フレーム画像の差分（背景差分）画像を生成する。画像内のほぼ同じ場所で、連続して背景差分が生じるフレーム数をカウントする。そのカウントが閾値を超えたときの背景画像、現在フレームの画像、背景差分画像の例をそれぞれ図5(a), (b), (c)に示す。また、動物体による変化領域として検出された領域を図5(d)に示す。図5(c)において、図5(d)に含まれるかたまりを除いた結果が、静的変化領域であり、図5(e)に示す。

4. 置かれた物体の領域と持ち去られた物体の領域の判定

背景画像と現在フレームのエッジ情報の差が生じる位置関係をもとに、検出された静的変化領域が置かれた物体による変化領域なのか、持ち去られた物体による変化領域なのかの判定を行う。

図4(a)の背景画像より図4(b)の現在フレームの微分値を引いた値について、正の場合を図6(a)に、負の場合を図6(b)に示す。図6(a)を見ると隠された部分のエッジが検出され、図6(b)を見ると置かれた物体のエッジが検出されていることがわかる。図6(a)と(b)の論理和が図6(c)である。図6(c)において、画像の横軸方向で各水平ラインの一番左と右にある画素（最大で2画素）、また、縦軸方向で各垂直ラインの一番上と下にある画素（最大で2画素）を抽出した結果を図6(d)に示す。抽出された画素のほとんどは、図6(b)のエッジに含まれているため、検出された変化領域は置かれた物体による領域と判定する。なお、本手法では、物体の外輪郭の情報だけでなく、物体内の絵柄の情報も用いている。このため、同じ形状であっても、絵柄が異なっていたり、位置が少しでもずれていれば判定が可能である。

5. ズームアップ画像の取得と物体領域の検出

5.1 ズームアップ画像の取得

物体の認識に必要なより詳細な情報を得るために、対象とする物体をズームアップして撮影した画像を得る。本論文で使用するカメラは、プログラムによって制御可能であり、置かれた物体による変化領域が生じた場合、自動でズームアップ画像を撮影する。図7(a)に示すシーンに対して、図7(b)に示す変化領域が検出された場合に、取得したズームアップ画像を図7(c)に示す。ズームアップ率としては、4～6倍程度としている。

5.2 ズームアップ画像からの物体領域の検出

得られたズームアップ画像から対象とする物体領域を検出する。2通りの手法によってそれぞれ物体候補領域を抽出し、それらの結果を統合することで、物体領域の

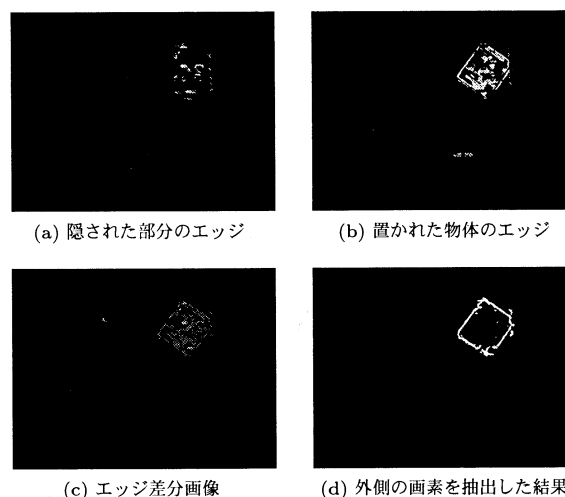


図6 置かれた物体の領域と持ち去られた物体の領域の判定
Detection of a placed object and a removed object.

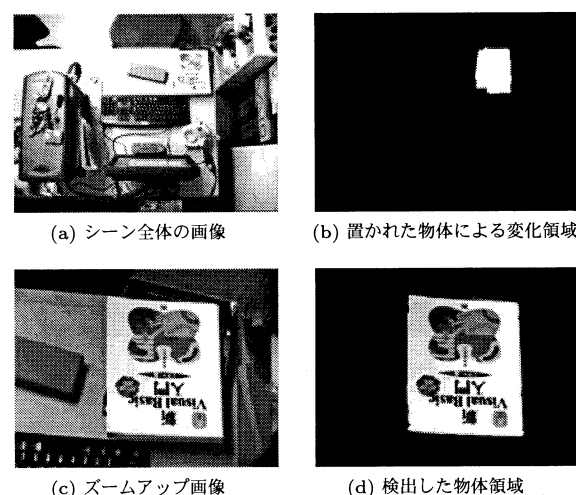


図7 ズームアップ画像と物体領域の検出
Zooming and detection of object region.

検出を行う。一つは、シーン全体の画像とズームアップ画像から求めたアフィン変換パラメータを利用して、物体候補領域を抽出する方法である。もう一つは、ズームアップ画像のエッジ情報を用いて領域分割を行い、物体候補領域を抽出する方法である。後者の物体候補領域は、領域分割によって複数の領域に分けられている。その一つ一つの領域の面積の半分以上が、前者の物体候補領域に含まれるか否かを調べ、含まれればその領域を物体領域として検出する。このような手順によって、図7(c)から物体領域を検出した結果が図7(d)である。

6. 物体情報の取得と物体の認識

本論文では、予め登録してある物体領域の参照画像と、入力画像から検出された物体領域について、形と色の情報を照合し、同一物体であるか否かを判断することで物体の認識を行う。面積、モーメント、外接矩形のアスペクト比、外接矩形との面積比の四つを形の情報として照合

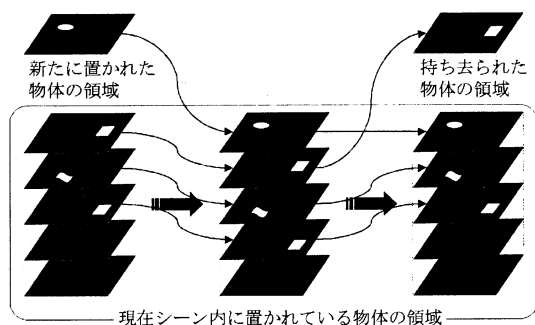


図 8 物体位置情報の記録
Recording of location of each object.

に用いる。また、 $L^*a^*b^*$ 色空間の色相ヒストグラムと、二つの物体領域中で一番大きな面積を占める色の配置情報を色の情報として照合に用いる。対象物体数が増加したときに処理の効率化を図るため、2物体領域の照合は、すべての情報を用いて一度に判断するのではなく、段階的に照合を行い、「似ている」物体領域のみ次の段階の照合を行う。容易に得られる情報から順に「形の照合」、「色相ヒストグラムの照合」、「色の配置情報の照合」を行う。

7. 移動履歴の記録と持ち去られた物体の特定

置かれた物体の領域が検出されたときには、そのときの静的変化領域を記録し、持ち去られた物体の領域が検出されたときには、そのときの静的変化領域と現在置かれている物体の領域とを比較し、どの物体が持ち去られたかを特定する。特定した物体の領域を削除することで、現在のシーン内に置かれている物体領域の記録を行う。新たに置かれた物体の領域を記録するときには、図8に示すように置かれた順番を保つため、置かれた物体の領域を一番上に記録する。一方、上から2番目の物体が持ち去られたときは、一番上にある物体の領域はそのままとし、上から3番目以降にある物体の領域を一つずつ上にずらすことにより、領域の削除を行う。このように記録することで、同じ位置にある物体でもどの物体が上にあるかということを記録することが可能であり、隠れてしまった物体も記録しておくことが可能となる。

図8に示すように、記録されている現在置かれている物体領域と、持ち去られた物体として検出された静的変化領域とを比較する。最後に置かれた物体の領域から比較していき、領域の面積が50%程度以上重なった物体領域を持ち去られた物体領域として特定する。

8. 実験結果

8.1 シーンの種類と静的変化領域の検出

人間がシーン内に移動し、物体を置くなどした後、シーンの外に移動するという2915フレーム分の時系列画像を実験に使用した。入力シーンの中に物体が置かれたり、

持ち去られたり物体の位置が変化した回数が74回あった。それに対し、86%の静的変化領域の検出に成功した。残りの10%については、3で述べた背景差分が正しく検出できていないため、また、4%については動物体を含むシーンの判定が的確に行えていないため、検出に失敗している。

8.2 置かれた物体の領域と持ち去られた物体の領域の判定

動物体がシーン内に現れる前と動物体がシーンから出て行った後の2枚の画像を一組とし、実験には39組、計78枚の画像を用いた。一組2枚のうち、1枚はもう1枚の画像よりシーン内に物体が一つ少ない画像である。これらの画像は、8.1において静的変化領域の検出が成功した画像の組である。39例のうち、置かれた物体の領域と持ち去られた物体の領域の判定に失敗した例は一例のみであった。失敗した原因は、背景画像と物体の色が酷似して物体の輪郭が検出できなかったためである。

8.3 物体領域の検出と物体認識

5.2の方法で、シーン全体の画像、検出された静的変化領域の画像、ズームアップ画像の3種類を用いて27の物体領域を検出した。そして、物体領域の検出結果の良し悪しに関わらず、検出された27の物体領域すべてについて、物体認識の実験を行った。実験に用いた物体の種類は、本やペン、ハサミなどの10種類であり、その中で物体領域の検出が比較的良好であった7領域(7種類)と、他の領域の照合をそれぞれ行った。7種類の参照物体と同一の物体が、それぞれ1から4領域あり、総領域数は17である。同一物体が同じ種類の物体と認識された率は71%である。失敗した原因のほとんどは、物体領域の検出がうまくいっていなかった点にある。物体領域の検出の失敗自体は、8.1、8.2で述べたような原因による。次に、参照物体とは異なる物体がそれぞれ22から25領域あり、総領域数は165である。異なる物体であるのに同じ種類の物体として判定された誤検出率は1%未満となった。ここでは三段階の照合を行ったが、それぞれの段階で1/3程度に物体領域を絞ることができ、三段階の照合がそれぞれ有効であったと考えられる。

9. むすび

本論文では、室内時系列画像を用いてシーン内で位置変化のあった物体領域を検出し、物体の認識を行うことで、個々の物体の移動履歴を記録する手法を提案した。また、実際にカメラにより撮影した時系列画像を用いて実験を行い、本手法の有効性を示した。今後の課題としては、人間が頻繁に物体の位置を変えるなどの複雑なシーン、あるいは、室内照明や外光などの照明の変化があるシーンへの適用などが挙げられる。

最後に、日頃ご指導頂く、電気通信大学大学院電気通信学研究科電子工学専攻、樽松 明教授に感謝致します。

〔文 献〕

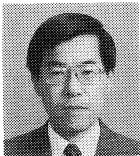
- 1) 飯島俊匡, 石上陽一, 川嶋稔夫, 青木由直: “日常動作のモニタリングによる行為履歴の記録”, 信学技報, PRMU98-187(1999)
- 2) 川崎広一, 久野義徳: “日常生活のモニタリングによる記憶支援システム”, 電子情報通信学会情報・システムソサイエティ大会, SD-4-1(2001)
- 3) E. Stringa and C. S. Regazzoni: “Real-time video-shot detection for scene surveillance application”, IEEE Trans. on Image Processing, 9, 1, pp.69-79 (2000)



ふじい あさこ
藤井 朝子 2001 年, 電気通信大学電気通信学部電子工学科卒業. 2003 年, 同大学院電気通信学研究科電子工学専攻博士前期課程修了. 同年, 富士通 (株) 入社. 在学中, 画像処理に関する研究に従事.



めぐろ みつひこ
目黒 光彦 1995 年, 武蔵工業大学工学部電気電子工学科卒業. 1997 年, 同大学院修士課程修了. 2000 年, 慶應義塾大学大学院博士課程修了. 同年, 電気通信大学大学院電気通信学研究科助手, 現在に至る. 工学博士. 非線形信号処理, 画像信号処理の研究に従事. 電子情報通信学会, 映像情報メディア学会, 情報処理学会, IEEE 各会員.



かねこ まさひで
金子 正秀 1976 年, 東京大学工学部電子工学科卒業. 1981 年, 同大学院博士課程修了. 同年, 国際電信電話 (株) 入社, 研究所に勤務. 1994 年, 東京大学工学部助教授. 1997 年, 国際電信電話 (株) に戻り研究所に勤務. 1998 年, 電気通信大学大学院電気通信学研究科助教授, 現在に至る. 画像符号化, 3 次元画像処理, 知的ヒューマンインタフェース, 顔画像処理などの研究に従事. 工学博士. 電子情報通信学会, 映像情報メディア学会, 情報処理学会, 日本 ME 学会, 日本顔学会, IEEE 各会員.