

安心・安全社会を支える画像認識による 広域監視技術

Image Recognition Wide-Area Surveillance Technology for Safety and Security in Society

あらまし

安心・安全社会を構築するためには、「社会を見守る目」が必要である。とくに、公園や公共施設など、多くの人が集まる広域での場所では、人の目が行き届かない死角が多く、犯罪が発生する可能性が高い。著者らは、「社会を見守る目」として、画像認識という人の目に代わる技術を使うことで、24時間365日広域での場所を監視する技術を開発した。

この技術は、75 m四方もの広い範囲で、画像中で小さく写る5画素程度の人でも認識し、その位置を通知する技術である。具体的には、「周波数パターン強調差分法」と呼ぶ、背景と人などの物体との空間周波数の違いによって、小さな物体でも抽出する技術である。さらに、物体の持つ周波数成分を強調して抽出することで、実用に不可欠な天候や日照変化への耐性も高い。実地での評価では、30～75 m四方の広域で95%以上の物体の抽出率で実現した。

本稿では、人の目に代わって自動監視する上での課題と、その課題の解決技術を紹介する。

Abstract

Surveillance is needed to build a safe and secure society. To meet this need, we developed a small-object detection technology for 24/7 wide-area outdoor surveillance based on an image recognition technology that is functionally similar to human vision. It can detect people and other small objects that are represented by as little as 5 pixels in an area of up to 75 m × 75 m and reports their locations. This technology, called frequency-pattern emphasis subtraction (FPES), detects a small object by analyzing the differences between the spatial frequencies of the background and the object. Because FPES emphasizes the spatial frequency component of objects, the detection precision is unaffected by changes in the weather. In experiments, this technology detected 95% or more of the objects in various areas of 30 m × 30 m to 75 m × 75 m. This paper describes FPES and some of the problems of using automatic surveillance and their solutions.



中山收文（なかやま おさふみ）
画像パターン処理研究部 所属
現在、人物認識およびITS車載技術など屋外画像認識技術の研究に従事。



三浦真樹（みうら まさき）
企画開発統括部ソフトウェア技術部 所属
現在、画像認識およびセキュリティシステムの企画開発業務に従事。

まえがき

2001年9月に起こった米国での同時多発テロ以来、いつどこで起こるか分からないテロの恐怖に世界中が巻き込まれ、今でもその恐怖と背中合わせの状況にある。日本では、民族的な性質によるものか、直面しない問題に対しての危機意識は低いと言われていた。しかし、国内では実感として捕らえどころのないテロの恐怖よりも身近な問題として、無差別殺人、放火などの事件が多く発生しており、危機管理への意識は少しずつではあるものの、高まりつつある。

このような背景を受けて、セキュリティに対する関心が高まっており、とくに、河川、港湾、空港、道路など、不特定多数の人が利用する大規模な公共施設では犯罪発生確率が高い。このように、24時間365日施設内の人や車の出入りなどを「社会を見守る目」として監視システムへの期待が一層高くなっている。この監視システムを支える要素技術の一つが、画像認識技術である。

本稿では、人の目に代わって自動監視する上での課題と、その課題の解決技術を紹介する。

自動監視への期待と課題

従来の監視システムでは、人の目に頼っていることが多い。とくに、遊技場などの大規模な施設では、警備員の巡回では対応しきれず、施設内に多数のカメラを、また監視を集中して行う専門の部屋にはモニタを配置し、数人の専任係員がモニタに映し出される施設内の映像を監視する。しかし、係員が24時間365日見落とさずに監視するのは難しく人による目視監視の限界がここにある。また、広域になればなるほど監視する箇所が増えるだけでなく、カメラの死角が増え、たとえ係員の集中力が続いたとしても広範囲に監視することは難しい。そこで、人の目に代わる技術として、画像認識と呼ばれる映像の内容を自動的に理解する技術を使い、安心・安全な社会を支える監視インフラが注目されている。

このような自動監視の実現には、映像から侵入者や侵入物体を発見する技術や自動抽出技術の研究開発が必要である。しかし、河川や港湾、公園のように50 m四方以上の広い場所がかつ屋外での自動抽出には、技術的に大きな課題がある。一つ目は、物

体が小さく、背景と識別しにくいことである。例えば、50 m四方での広域では、映像に写る人は5画素程度の大きさにしかならず、従来の技術では、映像に含まれるノイズと区別がつけられなかった。そのため、せいぜい10 m四方での抽出にとどまっている。二つ目の課題は、天候、日照などの撮影環境の変動や風などによるカメラの揺れによって、撮影条件が変動し、人や物体を安定に抽出することが難しいことである。とくに、撮影条件の変動は、実用段階では必須課題であり、従来技術では、10 m四方の監視においてもこの課題を解決したものはほとんどない。

広域で物体を抽出する従来技術は、大きく分けて3通りある。複数のカメラを連携して物体を抽出する方法とPTZ（パン・ティルト・ズーム）カメラで物体を追従する方法⁽¹⁾ 1台のカメラの映像から物体を抽出する方法である。前者二つの複数カメラ連携やPTZカメラ追従の方式では、カメラによるコスト高や追従することで複数の侵入物体に対応できないなどの問題がある。一方、3番目の方法である1台のカメラで物体を抽出する方法では、オプティカルフローを用いて動き領域を物体として抽出する方法^{(2),(3)}があるが、広域での撮影では、物体の動きが微小であるので、広域監視には向かない。背景との差異から物体を抽出する背景差分法⁽⁴⁾では、広域のように物体が小さいので物体と背景との輝度差が少なく抽出が困難である。これら1台のカメラによる方法では、屋外での環境変動やカメラの振動によって撮影条件の変化に追従できず、物体の安定な抽出は難しい。

広域監視の課題と新技術への取り組み

著者らは、安心・安全な社会実現のための強力な監視インフラとして、1台のカメラで、75 m四方の広域を屋外で24時間365日物体を安定に抽出できる技術を開発した。本章では、広域監視する場合の課題とそれを解決した新技術への取り組みについて述べる。

広域監視の課題

この技術を開発するために解決しなければならない課題は前章でも述べた以下の二つである。

(1) 天候や日照などの環境変動が起こる中でも、小さくぼけて写る物体をノイズと区別し抽出す

ること。

(2) 風などによりカメラが揺れる場合でも、小さくぼけて写る物体を抽出すること。

(1) の課題では、NTSCカメラを用いて監視範囲50m(画像中央での横幅)を撮影した場合は、図-1のように、画像中央の人物が5×10画素程度と小さくぼけて写るため、物体とノイズとの区別が難しい。また、屋外では、天候や日照などの環境変動により撮影条件が安定しない。

(2) の課題では、風などの影響でカメラが揺れる場合には、映像がぶれる。高さ10mの位置にカメラを設置する一般的な監視システムの場合、台風並みの風速60m/sの風が吹いたと仮定すると、VGAサイズ(640画素×480画素)の画像では、無風時に比べ最大で20画素程度のずれが生じ、安定な物体抽出が難しい。

新技術への取組み

前記の課題(1)を解決するため、屋外で起こり得る状況を、対象物体が存在する場合、日照変化など環境変化が生じる場合、画像ノイズがある場合の3通りに分類し、それぞれについて、背景と前景の両映像の周波数特性を調査した。結果を図-2に示す。物体の場合は、物体境界の有無の違いや物体内部の

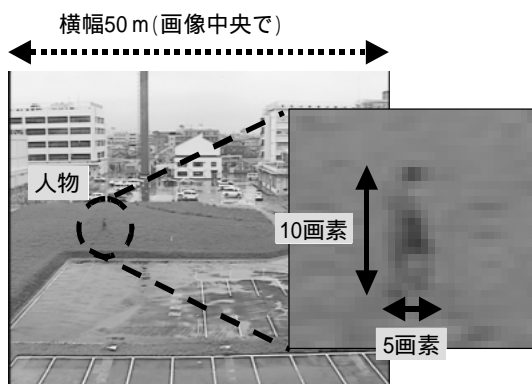


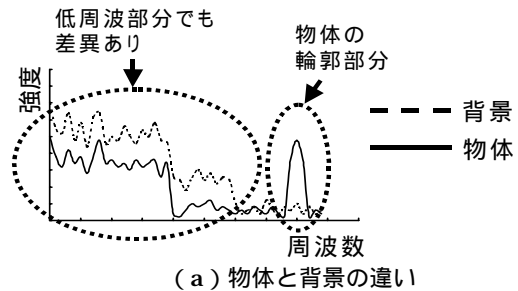
図-1 広域監視での画像(天候:雨,時間帯:昼の評価映像)
Fig.1-Outdoor image (rainy and daylight).

表-1 背景と前景の周波数特性
(差がある場合)

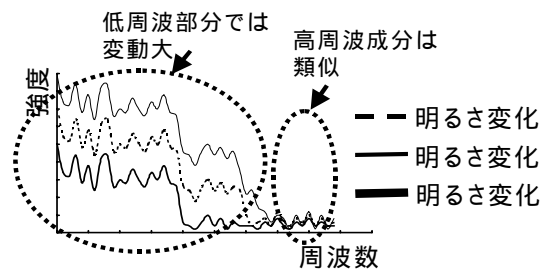
	高周波の変化	低周波の変化
物体		
日照変化	×	
ノイズ		×

輝度の違いによって、高周波から低周波にかけて広い範囲で周波数成分の変化が生じる{図-2(a)}。一方、照度変化などの環境変化の場合は、全体の輝度の増減となるので、低周波域にのみ変化を生じる{図-2(b)}。また、ノイズは画像中にスパイク状の変化となるので、高周波域にのみ変化を生じる{図-2(c)}。以上をまとめると表-1のようになり、つぎに示すことが明確になった。

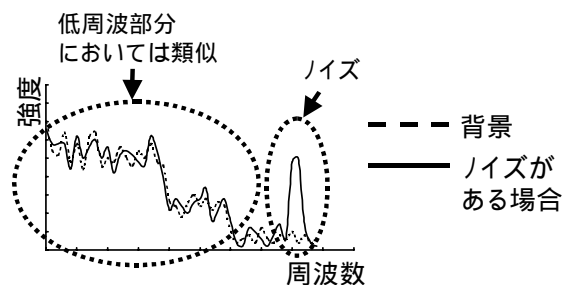
この周波数成分の変化の違いに着目すれば、高周波と低周波の両成分に変化を持つ領域を抽出することで、日照変化やノイズの影響を受けず、物体のみを抽出できる。この方式を「周波数パターン強調差分法」と名付けた。夜間でも低照度カメラを使うことで、この方式によって物体をカメラによる輝度上昇に伴うノイズと識別できるので、野球場のような大掛かりな照明装置がなくても物体を抽出することができる。



(a) 物体と背景の違い



(b) 日照変化による変化



(c) 画像ノイズの有無

図-2 広域屋外における各周波数分布
Fig.2-Spatial frequency in outdoor image.

課題(2)については、一般に監視システムでよく用いられるジャイロなどの機械補正や光学補正など映像スタビライザ(ブレ補正)により映像の揺れを補正する方法がある。しかし、画角40度以上と広角でのカメラは、レンズ歪^{ひず}みが大きく生じるため、歪みの影響でブレ補正された映像には、残留誤差と呼ぶ2画素程度のブレが残る問題がある。広域監視では、画像中での物体の大きさは5画素四方程度であるため、この残留誤差の影響は大きい。そこで、上記周波数パターン強調差分法において周波数ごとに位置補正を行うことで位置ずれを1画素以内に抑制する方式、多段周波数スタビライズ方式を開発した。次章以下で、各開発方式について詳述する。

周波数パターン強調差分法

周波数パターン強調差分法は、図-3に示すように、高周波強調処理がなされ、つぎに低周波強調処理がなされる。それぞれの周波数強調処理では、周波数強調を行った後、背景および入力画像を小領域に分割し、対応する小領域間で強調した背景、入力画像

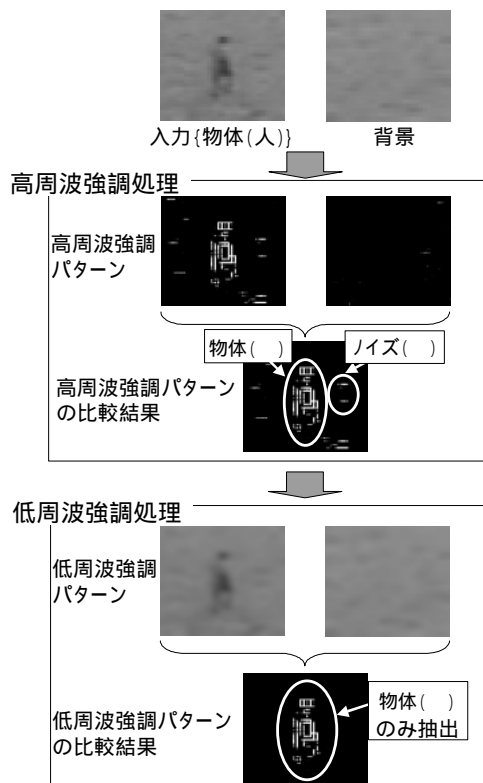


図-3 周波数パターン強調差分法の処理例

Fig.3-Flow of difference between spatial frequency of the background and objects.

のパターンを比較する。

このように小領域に分割することで、環境変動により部分的に画像の明るさが変化した場合でも、小領域ごとに見れば一様変化であり、その場合、先に述べた低周波成分の増減でモデル化できるため悪影響を受けない。

高周波強調処理では、最初に3×3のSobelオペレータを用いて周波数強調を行い、背景と入力画像とで対応する小領域間で強調したパターンを正規化相関により比較して、背景と類似していない小領域を抽出する。この処理により図中の物体()候補を抽出し、ノイズ()が抽出される。明るさ変化の影響を受けにくい高周波成分による物体候補の絞込みと小領域単位の局所的な比較により、環境変動の影響はこの段階で除外できる。

つぎに、抽出された物体候補について、3×3の平均化フィルタを用いて低周波強調処理を行い、高周波強調処理と同様に、強調したパターンで背景と入力画像とを比較する。これにより、ノイズ()は低周波成分では背景と類似しているため除去し、物体のみを抽出できる。

多段周波数スタビライズ方式

従来のスタビライザ(ブレ補正)でも残る±2画素の残留誤差を更に補正するために、周波数パターン強調差分の各周波数強調処理において位置補正を行う。

残留誤差がある場合、図-4(a)のように、背景と入力画像で対応する位置がずれているため、差分処理によって、本来背景の模様であるにもかかわらず、わずかであるが差が生じる。画像上で5画素四方程度と小さい抽出対象とこのわずかな差を区別できず物体と誤抽出する。そこで、小領域ごとの差分処理において、図-4(b)に示すように、背景と入力画像が最も一致するように位置を補正してから差分を取る。このように小領域ごとに個別に補正するので、カメラの揺れがあっても安定に物体を抽出できる。

性能評価

監視範囲50m(画像中央の横幅)における抽出性能を評価した。また、本技術で抽出可能な物体と監視可能な範囲を明確にするために、背景と物体と

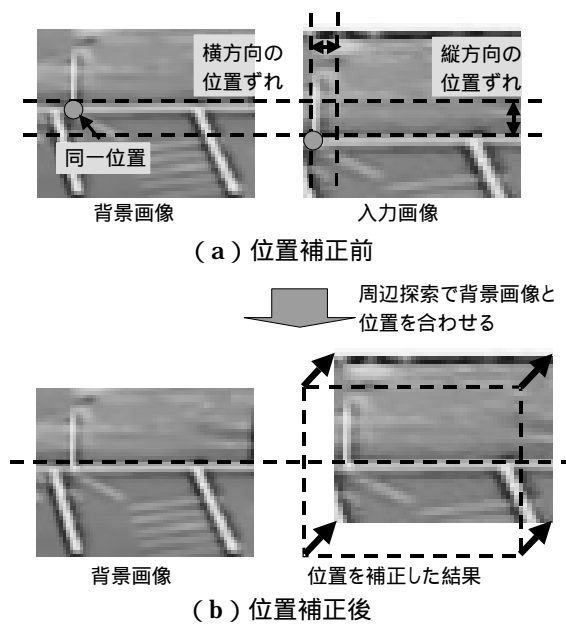


図-4 小領域ごとの位置補正
Fig.4-Adjustment of each area's position.

表-2 区間長を変化させた場合の抽出性能

区間長 (フレーム数)	抽出率 (%)	誤抽出率 (%)
10	93.4	60.9
20	95.9	30.7
30	96.9	7.1
40	97.3	3.7
50	97.5	0.6

網掛け部は、抽出率95%以上、誤抽出率5%以下

表-3 処理時間と計測PCの仕様

処理時間 (平均)	303 ms/frame
プロセッサ	Pentium4 2.2 GHz
メモリ	512 Mバイト
OS	Linux kernel 2.4.9

の輝度差と監視範囲を変化させて撮影した画像を用いた限界性能も評価した。

監視範囲50 mでの評価

評価データは、図-1に示すような監視範囲50 m (画像中央で撮影可能な横幅)で撮影した環境変化のある8時間分の実映像である。

なお、この実映像における主な環境変化を以下に示す。

- ・天候変化 (2種類): 晴, 雨
- ・時間変化 (2種類): 昼 (3万lx), 夜 (0.3 lx)
- ・カメラの揺れ (水平, 垂直, 回転方向の各組合

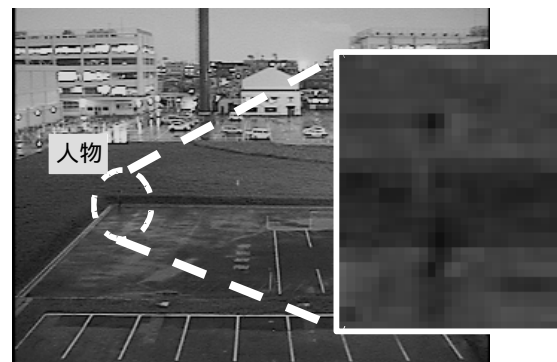


図-5 性能低下の主要原因 (低コントラスト)
Fig.5-Case of misdetection by low contrast of images.

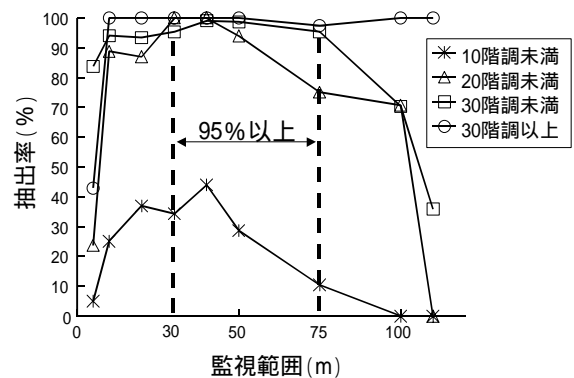


図-6 監視範囲と輝度差および抽出率の関係
Fig.6-Relationship between sizes of surveillance area and accuracy of detection.

せ): 最大 ± 20画素

また、性能評価の項目は以下の2点である。

(1) 抽出性能

抽出率: 物体を抽出した割合

誤抽出率: 物体以外を抽出した割合

(2) 処理時間: 1フレームあたりの処理時間

時間安定性のパラメタである、区間長を変化させた場合の抽出率と誤抽出率の関係を表-2に、処理時間の計測結果を表-3に示す。表-2の結果より、区間長を40フレーム以上に設定することで、誤抽出率5%以下で抽出率95%以上の性能が得られた。

なお、性能低下の原因を分析したところ、図-5に示すような、目視でも判断が難しい、背景との輝度差が10階調以下の低コントラストで、物体を安定に抽出できないことである。

限界性能評価

監視範囲や背景と物体との輝度差を変えた10時間分の映像を用いて両変化に対する抽出性能 (抽出

率のみ)の関係を調査した。評価に用いたデータは、監視範囲は5～110 m、背景と物体との輝度差は5～100階調程度の変化を持つ。監視範囲と背景と物体との輝度差および抽出率の関係を図-6に示す。図中の横軸は監視範囲の広さ、縦軸は抽出率を示し、観測点の系列は背景と物体の輝度差の違いを示している。

図-6の結果より、前節での評価結果と同様に、背景との輝度差が10階調未満の物体で大きく性能が低下するが、目視で確認可能な輝度差20階調以上の物体であれば、30～75 mの監視範囲で抽出率95%以上の性能が得られることが分かった。

む す び

物体が小さく写り、かつ安定した映像が得られないため、従来実用が困難であった広域屋外での物体抽出について、複数の周波数強調により背景と物体のパターンの差を強調し、統合する背景差分法である、周波数パターン強調差分法による新方式を提案した。提案方式を監視範囲50 mの実映像で評価した結果、撮影範囲を通過する物体について通過時間の95%以上で抽出でき、かつ誤抽出は全撮影時間の5%以下であることを確認した。さらに、目視で存在を確認可能な、背景との輝度差20階調以上の

物体であれば、30～75 mの範囲を抽出率95%以上で監視可能な見込みを得た。本技術は、画像監視システムのコア技術として、事業展開を始めており、複数地点での監視映像だけでなく、侵入者情報も集中管理できるソフトウェアも開発した。これにより、河川や港湾などの屋外の広域監視だけでなく、空港などの屋内の大規模施設での侵入物体監視システムなど、幅広い安心・安全な社会作りへの貢献が期待できる。

参 考 文 献

- (1) 川西隆仁ほか：動的アクティブ探索法と複数の能動カメラを用いた物体の位置推定．PRMU2001-81，2001，p.125-132．
- (2) 小林芳樹ほか：時空間情報を用いた屋外移動人物追跡．第2回画像センシングシンポジウム講演論文集，1996，p.173-176．
- (3) 長井 敦ほか：複雑背景下における移動物体の抽出．電子情報通信学会論文誌 Vol.J80-D- ，No.5，p.1086-1095 (1997)．
- (4) 福井真二ほか：輝度変化に対する頑健な背景差分法．画像の認識・理解シンポジウム，2004，p.305-310．