監視映像におけるオブジェクト移動履歴検索*

松井 康作 [†,††], 浜田 玲子 [††], 井手 一郎 [‡], 坂井 修一 [††]

[†,††]{matsui, reiko, sakai}@mtl.t.u-tokyo.ac.jp , [‡]ide@is.nagoya-u.ac.jp

[†] 日立工業専門学院, [††] 東京大学大学院情報理工学系研究科

[‡] 名古屋大学大学院情報科学研究科／国立情報学研究所

1 はじめに

近年、犯罪数の増加と共に、監視カメラや映像レコーダといったセキュリティ用映像装置の需要が高まり、急速に普及しつつある。しかし、映像レコーダの記憶容量が増大するにつれ、目的場面を検索する作業は長時間化しており、多大な労力を要するうえに、見逃しの問題をかかえている。

佐藤ら [1] は、映像中から人物領域を検出・追跡し、出現時刻、移動経路、速度などの特徴量を抽出して、場面検索時の鍵としている。しかし、移動経路、速度などの情報は抽出が困難であるうえに、犯人と単なる通行人との区別がつきにくいという問題がある。また、長時間記録、ネットワーク経由の遠隔監視などの制約から画像データ量が制限されるため、映像の解像度が低く犯人の顔の特徴がとらえにくいという問題もある。

そこで、本研究では、画像データ量の節約と高画質化の両立のため、個人特定の上で重要となる顔領域のみを高解像度で記録する監視映像蓄積手法を提案する。また、映像中のオブジェクトの移動に着目し、それらを自動検出・分類する手法について検討する。

2 監視映像中の物体移動検出

2.1 提案手法の概要

本手法では、監視映像において被写体となる人物を特定するのに有効な情報こそが重要だと考え、特に人物の顔に注目した。そこで、最初は高解像度デジタルカメラで撮影をしておき、顔画像以外の場面情報を大幅に削ることでデータ量を圧縮する。さらに検出した顔は、索引付けの際に利用する。

また、監視映像においてとらえるべき犯罪行為を「人対人」、「人対物」に大まかに分類して、索引付けの基準とする。それらの具体例を表 1 に示す。

表 1: 犯罪行為の分類とその具体例

分類	犯罪行為の具体例
人対人	暴力、痴漢、恐喝
人対物	器物損壊、不法投棄、窃盗、落書き、有害な掲示物

「人対人」に分類される犯罪行為には明確な被害者が存在する。被害者からの情報により、ある程度の犯行時刻が分かれば目的場面の検索は容易であると考え、本研究では検出対象から外した。

一方、「人対物」に注目すると、これらの行為は共通して人気のない場所で行われ、犯行時刻が分かりにくい。犯行時刻が分からなければ、確認作業の負担は大きくなるため、本研究ではこの種の犯罪行為を検出対象とする。これらの行為の前後では、共通して背景領域に変化が生じる点に着目し、その変化をオブジェクトの「もち込み」／「もち去り」に分類して索引付けを行う。例えば、パソコンの盗難があった場合は「もち去り」、不法投棄や落書きがあった場合は「もち込み」と分類する。最後にオブジェクトの動作に最も関連の深い顔画像を対応付けし、検索時の鍵とする。

2.2 処理の流れ

提案手法は、以下のような手順で処理する。

1. 動きのあるフレームの抽出
2. 顔領域の抽出
3. オブジェクト領域の抽出
4. オブジェクトの移動判別
5. 顔とオブジェクトの対応付け

2.2.1 動きのあるフレームの抽出

本手法では、輝度情報に基づくフレーム間差分と背景差分の組み合わせて利用し、人物の写っている可能性の高い、動きのあるフレームのみを記録する。また、一連の動きのあるフレーム列をシーンと呼び、動きのないフレーム列と区別する。

2.2.2 顔領域の抽出

シーンの中から顔領域を抽出し、128 × 160 画素の大きさで記録する。顔領域抽出後のフレームは、高画質である必要はないと考え、30 万画素 (640 × 480 画素) に縮小して記録する。これにより、保存するデータ量の約 9 割を削減する。顔領域の抽出には、ニューラルネットワークによる学習を用いた face detector[2] を使用する。face detector を実際の監視映像 (18 分、1080 フレーム) に適用して顔領域を抽出したところ、97.6%の適合率と 93.2%の再現率が得られ、十分な実用性を確認した。

2.2.3 オブジェクト領域の抽出

シーンの直前のフレーム (before フレームと定義) と、シーンの終端フレーム (after フレームと定義) との間で、画素毎に輝度値の差分を求め、その値が大きい領域をオブジェクト領域として抽出する。

2.2.4 オブジェクトの移動判別

オブジェクトの「もち込み」／「もち去り」は、before フレーム、after フレームのオブジェクト領域を調査し判別する。after フレームにオブジェクトが存在してい

---

*"Indexing of surveillance video based on object relocation"
Kousaku Matsui[†,††], Reiko Hamada[††], Ichiro Ide[‡]
Shuichi Sakai[††]
[†]Hitachi Technical College
2-17-2 Nishinarusawa-cho, Hitachi-shi, Ibaraki 316-0032
[††]Graduate School of Information Science and Technology,
The Univ. of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033
[‡]Graduate School of Information Science,
Nagoya Univ. / National Institute of Informatics