

非校正カメラ群を用いた人物観測における注目領域抽出

筑波大学大学院 システム情報工学研究科 西崎 隆志, 亀田 能成, 大田 友一

ROI Extraction in Visual Surveillance

Using Multiple Non-Calibrated Cameras

Takashi NISHIZAKI, Yoshinari KAMEDA, Yuichi OHTA

Graduate School of System and Information Engineering, University of Tsukuba

Abstract : We present a method to realize efficient visual surveillance under an environment in which a number of non-calibrated fixed surveillance cameras are operated. Our method extracts regions of interest (ROIs) that are useful to understand what kind of event is occurring. A set of ROIs that are worth watching can be extracted by referring co-occurrence probabilities of pairs of ROIs.

1 はじめに

近年多数の監視カメラを設置してセキュリティ等の用途に利用する研究がなされてきている¹⁾。しかし、監視カメラネットワークの規模が大きくなると、得られる大規模多視点映像を人が見た場合に、どのカメラのどの領域で注目すべきイベントが生じているか等を理解することは難しくなる。また、従来の監視・観測システムはカメラ校正を必要とするものが多いが²⁾、カメラネットワークが大規模になると、厳密な手動/自動カメラ校正を適用することが困難となる。そこで本研究では、多数の非校正カメラ群を対象とし、得られる情報から重要な部分(注視領域)のみを抽出する手法について提案し、それが監視カメラ映像の閲覧や自動イベント検出・認識システムに利用できることを示す。

2 観測システム

図1に本観測システムの概要と、観測される多視点映像の一部を示す。本システムはネットワークカメラ群と画像処理用PC群で構成される。本構成では、PCとカメラとを組で設置する必要がないため、設置の自由度・拡張性が高い、一台のPCで複数台のカメラを扱える、等の利点がある¹⁾。

本研究では大規模データを扱うため、基本的な特徴量のみを保存する。まず、得られる映像に対して背景差分を取ることでより前景領域を求める。得られた前景領域をR個の領域に分割し、各領域内で平均輝度値を計算して、これを保存している。現在、 $R = 64$ に設定している。保存される情報例を図2に示す。

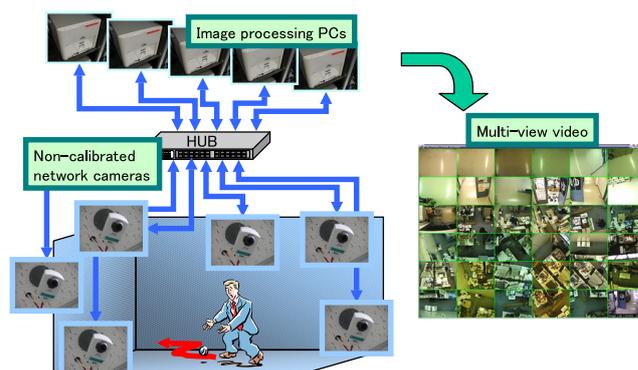


Fig. 1: 観測システムと多視点映像



Fig. 2: 保存される特徴量

3 注視領域抽出とイベントの共起性

3.1 多次元データ構造

カメラ台数を C 、画面上での領域分割数を R とし、以下に示すような“時刻 t におけるイベントベクトル $\mathbf{x}(t)$ ”を定義する。

$$\mathbf{x}(t) = \{ x_1(t), \dots, x_i(t), \dots, x_N(t) \} \quad (1)$$

このベクトルサイズは $N = C \times R$ となる。各イベントベクトルは、どのカメラのどの領域に動物体が映っているかというイベントに対応する。各要素 $x_i(t)$ は、あるカメラのある区画に時刻 t において検出された物体の特徴量で、どのような物体がそこに映っているのかを表す。

¹⁾逆に、カメラ間で撮影映像の同期がとれない、映像のフレームレートがネットワークトラフィックに影響されて安定しない、等の欠点があり、撮影モデルの推定時にこれらを補償する必要がある。

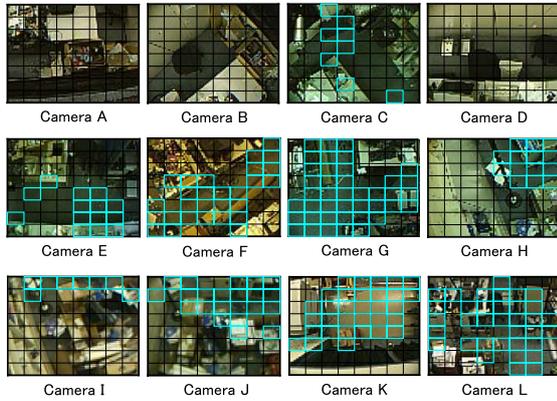


Fig. 3: 抽出された注視領域（明るい色の枠）

3.2 注視領域の抽出

保存されるイベントベクトルは大規模多次元データ構造であり、これを扱いやすくするために、冗長要素削減手法を適用する。この結果得られる、観測上重要なベクトル要素 $\{x_i\}$ を“注視領域”と定義する。

まず、観測中全ての時刻において物体が観測されなかったベクトル要素 x_i を、センシングにおいて意味をなさないと思われ、除去する。次に、主成分分析を施した後、以下の手順で冗長要素削減を行う。

手順 1: 主成分群 $\{z_j\}$ を寄与率 $\{p_j\}$ が高いものからソート。

手順 2: 累積寄与率がある閾値 (0.9 に設定) になる主成分までを“寄与率が高い主成分”として残す。

手順 3: ある寄与率が高い主成分 z_j において、 z_j を表す $\{x_i\}$ の係数群 $\{a_{ij}\}$ の平均値 \bar{a}_j を求め、係数 a_{ij} がこの値よりも大きい要素 x_i に関して、要素 x_i のスコア s_i として、主成分 z_j の寄与率 p_j を加算する。

手順 4: 寄与率が高い主成分の全てに対して手順 3 を適用し、各要素に対するスコア群 $\{s_i\}$ の平均値 \bar{s} を求める。そして、スコア s_i が \bar{s} よりも大きい要素 x_i を残す。

図 3 に、カメラ 12 台、領域分割数 64 の場合の冗長要素削減結果を示す。なお現在、得られた注視領域群のみを用いた観測イベントのクラスタリング処理部を実装しており、全ての観測情報を用いた場合に比べ、高速・効率的なイベントの分類が行えることを確認している。

3.3 注視領域間におけるイベントの共起性

求められた注視領域の情報から、二つの注視領域間におけるイベントの共起性を求めることで、“関連注視領域群の提示”と“注目カメラ群の選出”を実現することができる。

注視領域ペア (x_m, x_n) の間で、イベントの共起性を示す相関値 c_{mn} を計算する。

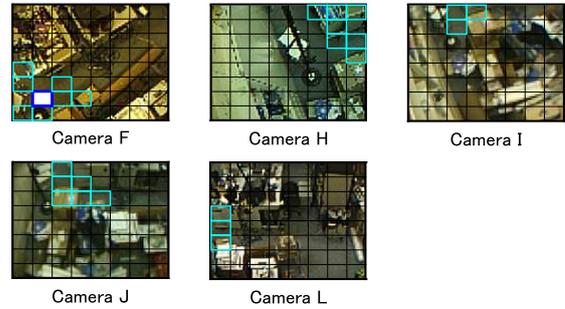


Fig. 4: 共起率の高い注視領域群

$$\begin{aligned}
 \bullet c_{1mn} &= \frac{\mathbf{x}_m \cdot \mathbf{x}_n}{|\mathbf{x}_m| |\mathbf{x}_n|} = \frac{\sum_{t=0}^T x_m(t) x_n(t)}{|\mathbf{x}_m| |\mathbf{x}_n|} \\
 \bullet c_{2mn} &= \begin{cases} \frac{|\mathbf{x}_m|}{|\mathbf{x}_n|} & \text{if } |\mathbf{x}_m| < |\mathbf{x}_n| \\ \frac{|\mathbf{x}_n|}{|\mathbf{x}_m|} & \text{otherwise} \end{cases} \\
 \bullet c_{mn} &= c_{1mn} c_{2mn}
 \end{aligned}$$

関連注視領域群の提示の例を図 4 に示す。ある特定の注視領域（白く塗りつぶされた領域）に注目し、これと共起性が高い注視領域を明るい色の枠で示した。このような関連注視領域群の提示を実現することにより、大規模多視点映像の閲覧者は同じイベントが観測されているカメラやカメラ映像上の領域を迅速に確認することができる。

また、共起性の高い注視領域ペアが属するカメラをスコア付けすることによって、注目カメラ群（特に注目する必要のあるカメラ群）の選出を実現している。実際に、図 3 のカメラ群に対して本処理を適用した結果、12 台中 5 台のカメラ (F, G, J, K, L) が選出された。

4 まとめ

本稿では、多数の非校正カメラ群を対象とし、映像監視に重要と思われる注視領域を自動的に抽出し、多視点映像の閲覧や自動イベント検出・認識システムに利用する手法を提案した。多視点映像閲覧者に対するサポートツールの例として、“関連注視領域群の提示”と“注目カメラの選出”を実装し、その有効性を示した。

今後は、本手法を基にしたイベントの検出・認識システムを実装し、現在撮影シーンにおいて起こっているイベントの分類と映像閲覧者に対する注目喚起などを実現していくことを予定している。

参考文献

- 1) C. Stauffer and K. Tieu, ‘Automated multi-camera planar tracking correspondence modeling,’ CVPR 2003, pp. 259–266.
- 2) M. D. Beynon, D. J. Van Hook, and M. Seibert and A. Peacock, ‘Detecting abandoned packages in a multi-camera video surveillance system,’ IEEE Int. Conf. on AVSS 2003, pp. 221–228.