

# 非校正カメラ群を用いた室内イベントの判別

西崎 隆志<sup>†</sup> 亀田 能成<sup>†</sup> 大田 友一<sup>†</sup>

<sup>†</sup> 筑波大学大学院システム情報工学研究科, 茨城県つくば市天王台 1-1-1

あらまし 多数の監視カメラ群を運用する際, どのような出来事が起こっているを理解するために, どのカメラのどの領域をモニタリングすれば良いのかを把握することは容易な問題ではない. この問題は, カメラの数が多く, 発生しうる事象(イベント)の分類が困難な状況では, 特に顕著となる. 本研究では, 多数の非校正カメラ群を対象とし, 得られる多視点映像から観測に重要な部分(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. 本手法では, 多数のカメラを用いた観測システムによって蓄積される多視点映像データから多視点映像中の注視領域群を抽出する. そして得られた注視領域間のイベントの共起性を求めることにより, 同じイベントが観測され得る注視領域群を推定する. これにより, 多視点映像の閲覧者は, ある注視領域でイベントの発生を確認した際に他のどの領域を注目すればよいのかが分かる. また, 全映像データを処理するのではなく抽出された注視領域だけを扱えば, 大規模多視点映像に対応した自動イベント検出・認識システムを実現することが期待できる.

キーワード マッシュセンシング, 多視点映像, 注視領域, イベント検出, イベント認識

## Indoor Event Discrimination by Utilizing Uncalibrated Cameras

Takashi NISHIZAKI<sup>†</sup>, Yoshinari KAMEDA<sup>†</sup>, and Yuchi OHTA<sup>†</sup>

<sup>†</sup> Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573 Japan

**Abstract** With a large number of surveillance cameras, it is not an easy task to determine which camera should be monitored and which region of the camera images should be checked so that all the activities and/or events in a scene are examined. We present a new method to realize effective visual surveillance under an environment in which a number of non-calibrated fixed surveillance cameras are being operated. The method is based on extractions of “regions of interests (ROI)” that are important for visual surveillance, and useful for monitoring multi-view videos of the cameras and automated event detection and/or recognition. In our method, ROIs are extracted by eliminating redundant regions of multi-view images stored on the visual surveillance using multiple cameras. By calculating co-occurrence between each pair of ROIs, our method estimates ROIs where events are to be observed. This helps viewers to know which other ROIs they should check when an event is observed in a certain ROI. In addition, the method can be applied to automated event detection and event recognition on a massive number of multi-view videos.

**Key words** massive sensing, multi-view videos, regions of interests, event detection, event recognition

### 1. ま え が き

近年, 様々な環境にカメラを設置し, 日常的/非日常的な出来事を検出・観測する要求が社会的に高まってきている. そのような監視用のカメラは, 特に公共の場でのセキュリティや道路交通監視などに利用可能であり, 今後さらに設置数が増えると考えられる. しかし, 大量の監視カメラが設置・運用されるようになると, カメラがどこに配置されたか, 各カメラがシー

ン中のどの領域を撮影しているのかが, カメラ映像を閲覧した際に分かりづらいという問題が発生する. よって, 人が大量の監視カメラから得られる映像を閲覧する場合, シーン中で起こっている出来事を把握するために, どのカメラを監視し, そのカメラから得られる画像のどの部分をチェックすればよいのかを判断することが難しくなる. 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, 多数のカメラが広範囲に分散して設置された場合, 厳密なカメラ校正を適用す

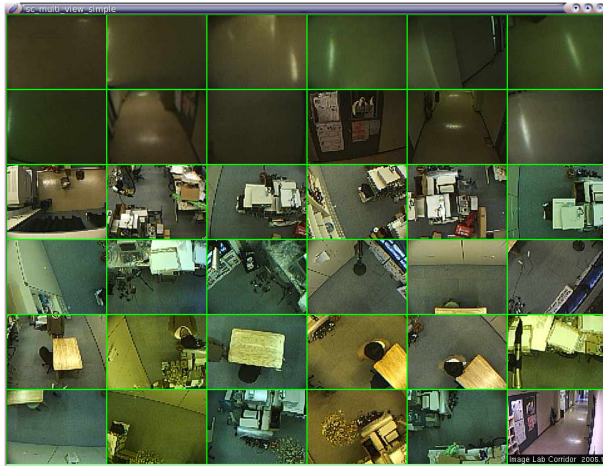


図 1 多視点映像の例 (36 台のカメラ)

実際に撮影された多視点映像の例を示す。この図を見た際に、どのような種類のオブジェクトがあるのか、どのようなイベントが起こっているのかを人が理解することは難しいと思われる。

ることが現実的ではなくなり、この問題が顕著となってくる。

同時に撮影された 36 台のカメラのスナップショットを図 1 に示す。これらのカメラは、ある一部の天井と、その部屋に隣接する廊下の天井に設置されている。この図を見て画面上に映っている人物の位置を判断することが難しいことは明らかである。また、監視カメラの数が増えれば、それらの幾何的情報 (位置, 向き) の整合性を保持することが面倒になってくる。したがって、厳密な幾何的校正がとられていない大量のカメラを対象とした、知的カメラ監視システムが必要となってくる。

本論文では、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。ここで提案する手法の特徴は、二つの注視領域 (ROIs: regions of interests) 間でのイベントの共起性を評価することである。ここで注視領域とは、大規模多視点映像群の中で、閲覧者ないしイベント検出・認識システムが最低限注目すべき領域を指す。そして、注視領域間の共起性を求めることにより、同じイベントが観測され得る注視領域群を推定する。これにより、多視点映像の閲覧者は、ある注視領域でイベントの発生を確認した際に他のどの領域を注目すればよいか分かる。また、全映像データを処理するのではなく得られた注視領域だけを扱えば良いため、大規模多視点映像に対応した自動イベント検出・認識システムを実現することができる。

本論文の構成は以下のようになっている。まず、2. で本研究に関連のある研究を挙げ、本研究の位置付けを述べる。次に 3. において、我々が構築した多カメラ監視システムについて説明する。そして、4. において、本研究の特徴でもある注視領域の抽出手法について述べる。得られる注視領域の有用性を示すため、5. において注視領域の評価を基にした監視支援ツールの紹介を行い、6. において注視領域を用いたイベント認識の展望について述べる。そして、7. において、本論文をまとめる。

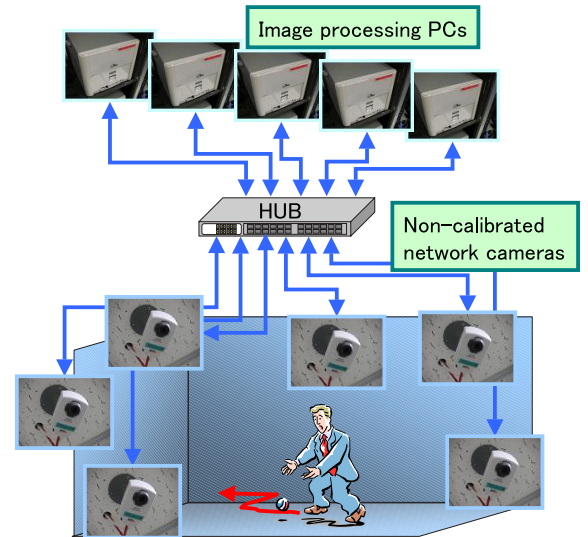


図 2 観測システムと多視点映像

## 2. 関連研究

近年、人物追跡・交通監視・非日常物体の検出などを対象とした映像監視システムの研究が盛んになってきている。広い撮影領域を撮影したり物体の複雑な運動を追跡するためには、複数のカメラを用いる必要がある。従来研究における多カメラ監視システムでは、主に手動でのカメラ校正 [1] [2] [3] [4]、もしくは複雑な自動校正手法 [5] を必要としていた。校正済みカメラを用いる監視システムは環境中におけるオブジェクトの正確な幾何情報を求めることができるが、規模の大きい監視カメラシステムに手動でのカメラ校正を適用することは非常に煩わしく、また、自動カメラ校正手法をそのようなシステムに適用することは難しい。このことから、カメラの大まかな幾何情報のみを仮定した監視手法の要求が高まってくる。

映像間の対応を求めておくことは、監視映像中の動く物体を複数のカメラ上で観察する際に有効である。したがって、動物体の位置や移動軌跡の映像間における対応を推定する多くの手法が提案されてきている [6] [7] [8] [9]。本研究が従来研究と異なる点は、本手法が大規模なカメラネットワークを想定していること、閲覧者を補助するツールも提案していることにある。

## 3. 多カメラ観測システム

### 3.1 カメラネットワーク

図 2 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、我々の提案しているマッシュセンシングシステム [10] の一部となっている。図に示すように、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。ネットワークカメラを用いたのは、以下に挙げる利点を考慮したためである。

- PC の台数とカメラの台数を一致させる必要がないため、カメラ設置の自由度・拡張性が高い。
- PC 側の接続インターフェースの数を考慮する必要がないため、一台の PC で複数 (多数も可) 台のカメラを扱える。逆に、以下のような問題も挙げられる。



図3 保存される特徴量(右)

- 撮影される映像の同期がとれない。
- 撮影される映像のフレームレートがネットワーク状態に影響され安定しない。

注視領域間のイベントの共起性を推定し、それを監視に利用するためには、上記の問題に対応する必要がある。現在の実装では、カメラからは約 10.0 ~ 30.0 fps で映像が得られ、シーン中の物体の動きはこのフレームレートに比べそれほど早くないため、カメラ間同期の時間精度は重要な問題ではない。また、ネットワークの混雑は、最大帯域を超えない通信を行うことにより回避できる。現在、本システムは 45 台のカメラと 21 台の処理用 PC で構成されており、今後カメラを約 80 台に増やす予定である。

### 3.2 画像特徴量

撮影された映像からは様々な画像特徴が抽出できるが、本研究では扱うイベントを限定しないため、基本的な特徴のみを考慮する。本研究では、カメラから得られる画像を  $R$  個の小さな矩形領域に分割し、その領域内で得られる特徴量の平均値を扱うことにより、保存されるデータ量を削減している。現在、 $R = 64$  としている。また保存する特徴量として、入力画像に背景差分法を適用することによって“前景領域”を抽出し、各矩形領域内でのその平均をとる。保存される情報の例を図3に示す。これら情報は実際には圧縮保存されており、保存データサイズは状況によって異なる。例として、4 台のカメラで約 290 時間の観測を行った場合、保存された圧縮ファイルのサイズは約 1.3 GByte となった。

## 4. 注視領域の抽出

ここでは、前述の各矩形領域から注視領域 (ROI: region of interests) を抽出する手法について述べる。

### 4.1 データ構造

カメラ台数を  $C$  とし、以下に示すような時刻  $t$  ( $1 \leq t \leq T$ ,  $T$  は観測ベクトル数) における観測ベクトル  $\mathbf{x}(t)$  を定義する。

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

このベクトルの要素数は  $N = C \times R$  となる。各観測ベクトルは、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を表現できる。各要素  $x_i(t)$  は、あるカメラのある領域  $i$  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。現在、この特徴量は 3.2 で述べた前景領域の輝度値平均をそのまま用いている。ここで、他の種類の特徴量も本手法に同様に適用できることに注意され

たい。

保存される観測ベクトルは次元数が大きく、このままでは扱いにくい。しかも、長時間観測の結果貯蓄される観測ベクトル群において有意な情報は偏って存在する。よって、次元数が大きいデータに対する冗長要素削減 (観測ベクトルに対する次元数の削減) が有効であると言える。本研究では、本節で述べる冗長要素削減手法を適用することによって得られた、観測に重要なベクトル要素  $\{x_i\}$  に対応する領域  $i$  を注視領域と定義している。

以下、冗長要素削減に基づいた注視領域の抽出手法の述べる。

#### 4.1.1 前景が観測されない領域の除去

まず、入力となる  $N$  次元の観測ベクトル  $\mathbf{x}(t)$  に関して、平均  $M$  と分散共分散行列  $V$  を計算する。領域  $i$  で観測される特徴量の平均値  $M_i$  が 0 の場合、領域  $i$  では一度も動物体が観測されなかったことを意味するため、これを冗長な領域として破棄できる。この処理を適用することによって、 $N'$  ( $N' \leq N$ ) 次元のベクトル  $\mathbf{x}'(t)$  を得る。

#### 4.1.2 主成分分析による次元縮退

次に、分散共分散行列  $V$  を用いて、残った要素  $\{x'_j\}$  に対して主成分分析を行う。求まった主成分  $z_k$  ( $1 \leq k \leq N'$ ) は、もともとなるベクトル  $\mathbf{x}'$  の各要素の線形和として以下のように表される。

$$\mathbf{z} = \mathbf{A} \mathbf{x}' \quad (2)$$

$$z_k = a_{1k}x'_1 + a_{2k}x'_2 + \dots, a_{N'k}x'_{N'} \quad (3)$$

本手法では、寄与率の高い主成分  $z_k$  を作る要素  $x'_j$  を残し、他を除去する。以下、冗長要素削減手法の詳細について述べる。  
手順 1: 主成分群  $\{z_k\}$  を寄与率  $\{p_k\}$  が高いものからソートする。寄与率  $p_k$  はその主成分がどれだけデータを表現しているかを示し、 $z_k$  の分散  $\lambda_k$  を用いて以下のように表される。

$$p_k = \frac{\lambda_k}{\sum_{l=1}^{N'} \lambda_l} \quad (4)$$

手順 2: 累積寄与率

$$c_k = \sum_{l=1}^k p_l \quad (5)$$

がある閾値  $c_{th}$  (現在 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: 寄与率が高いある主成分  $z_k$  において、 $z_k$  を作る  $\{x'_j\}$  の係数群  $\{a_{jk}\}$  の平均値  $\bar{a}_k$  を求める。係数  $a_{jk}$  がこの値よりも大きい要素  $x'_j$  に関して、要素  $x'_j$  のスコア  $s_j$  として、主成分  $z_k$  の寄与率  $p_k$  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  $\{s_j\}$  の平均値  $\bar{s}$  を求める。そして、スコア  $s_j$  が  $\bar{s}$  よりも大きい要素  $x'_j$  を残す。これに対応する領域  $j$  が注視領域となる。

### 4.2 注視領域抽出の実験結果

実験例におけるカメラの配置を図4に、注視領域抽出の結果を図5,6に示す。ここで、明るい色の枠で囲まれた領域が抽出された注視領域を示し、部屋内を歩き回る人物がその領域で頻繁に観測されたことを意味する。4 台のカメラからなるシステ

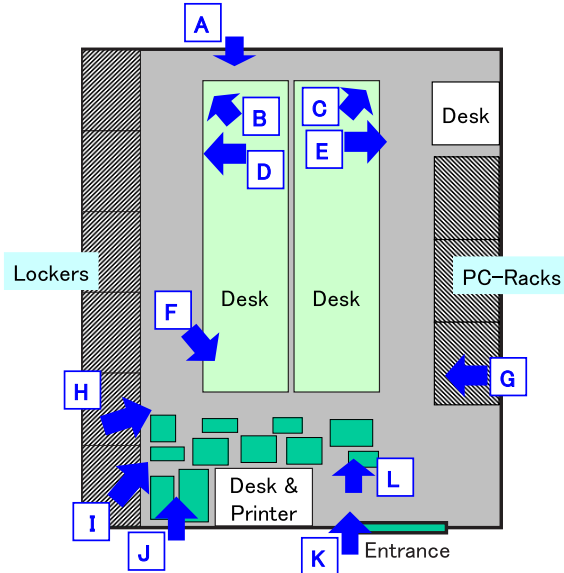


図 4 カメラ配置

アルファベットはカメラ名, 矢印はカメラの向きを示す.

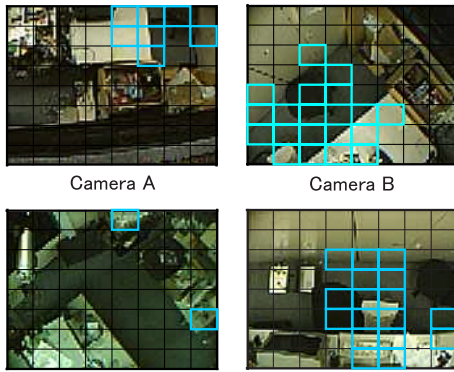


図 5 得られた注視領域 (カメラ 4 台の場合)

ムの場合に選出された注視領域 (図 5) のいくつかは, 12 台のカメラからなるシステムの場合 (図 6) において除去されている. これは, カメラ 12 台の場合において, 他の注視領域に比べてそれらの注視領域の重要度が低いためであると考えられる.

注視領域抽出には, CPU: Pentium4 2.80 GHz, メモリ: 1.0 GByte の PC を用いた. 約 2 時間のシーンに対して注視領域抽出にかかる処理時間は, カメラ 4 台の場合に 181.70 秒 (平均の計算に 48.27 秒, 分散共分散行列の計算に 133.30 秒, 冗長要素削減に 0.13 秒), カメラ 12 台の場合に 1347.19 秒 (平均の計算に 147.7 秒, 分散共分散行列の計算に 1190.6 秒, 冗長要素削減に 8.89 秒) となった.

また, 手順 1, 2 で述べた, 各主成分に対する寄与率  $\{p_k\}$  の分布の例を図 7 に, 手順 3 で述べた, ある主成分  $z_k$  を作るための係数群  $\{a_{jk}\}$  の例を図 8 に, 各要素  $\{x'_j\}$  に対するスコア群  $\{s_j\}$  の例を図 9 に示す.

## 5. 注視領域間の共起性の評価に基づく監視支援

注視領域が求まると, システムは様々な映像監視支援を実現することができる. 本論文では, 以下に述べる二つの支援アプリケーションについて紹介する.

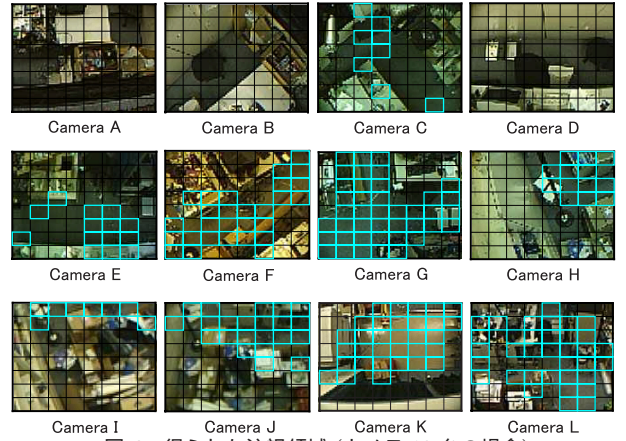


図 6 得られた注視領域 (カメラ 12 台の場合)

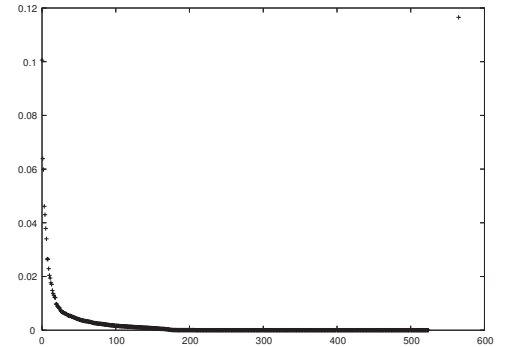


図 7 寄与率の分布 (手順 1, 2)

横軸は主成分のリスト, 縦軸は対応する寄与率を表す.

### 5.1 関連する注視領域群の提示

ある注視領域に注目すべきイベントが観測された場合, 本手法はそれに関連する他の注視領域群を選出でき, これは視聴者によるイベント把握の支援となる.

関連する注視領域を選出するために, 本手法では, 二つの注視領域  $m, n$  の間でのイベントの共起性を求める. まず, 注視領域  $m$  で観測される特徴量の値を  $y_m(t)$  ( $1 \leq t \leq T$ ) とし, 注視領域  $m$  で観測される特徴量のデータセットを以下のようなベクトルで表す.

$$\mathbf{y}_m = \{ y_m(1), \dots, y_m(t), \dots, y_m(T) \} \quad (6)$$

そして, 以下に示すような相関値  $c_{mn}$  ( $0 \leq c_{mn} \leq 1$ ) を計算し, これを共起性尺度として利用する.

$$\begin{aligned} \bullet \quad c_{1mn} &= \frac{\mathbf{y}_m \cdot \mathbf{y}_n}{|\mathbf{y}_m| |\mathbf{y}_n|} = \frac{\sum_{t=1}^T y_m(t) y_n(t)}{|\mathbf{y}_m| |\mathbf{y}_n|} \\ \bullet \quad c_{2mn} &= \begin{cases} \frac{|\mathbf{y}_m|}{|\mathbf{y}_n|} & \text{if } |\mathbf{y}_m| < |\mathbf{y}_n| \\ \frac{|\mathbf{y}_n|}{|\mathbf{y}_m|} & \text{otherwise} \end{cases} \\ \bullet \quad c_{mn} &= c_{1mn} c_{2mn} \end{aligned}$$

ここで, 相関値  $c_{mn}$  の値が大きい場合に, 二つの注視領域  $m, n$  の間でのイベント観測の共起性が高いと見なせる. このうち,  $c_{1mn}$  がベクトルの傾きの類似度を評価し,  $c_{2mn}$  がベクトルの大きさの類似度を評価する.

関連する注視領域群を提示した例を図 10, 11 に示す. 図中の白塗りされた領域と共起性の高い注視領域を明るい色の枠で強調した. このような関連する注視領域群の提示を実現することにより, 多視点映像の閲覧者は同じイベントが観測されている

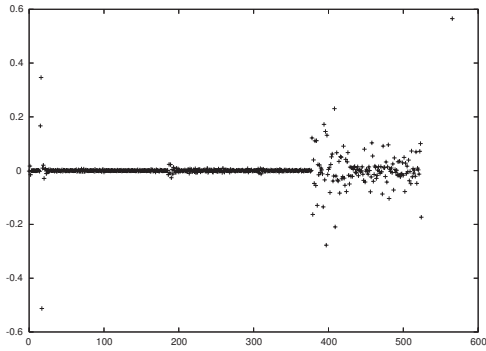


図 8 主成分を作るための係数の分布 (手順 3)

横軸はベクトル (データ) の要素  $x'_j$  のリスト, 縦軸は対応する係数を表す.

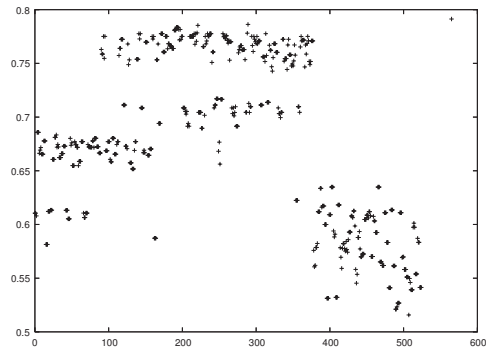


図 9  $x'_j$  のスコアの分布 (手順 3)

横軸はベクトル (データ) の要素  $x'_j$  のリスト, 縦軸は対応するスコアを表す.

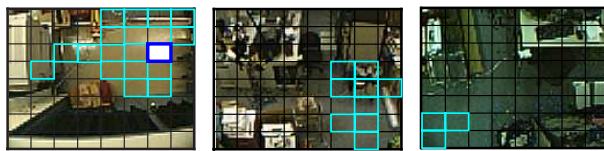


図 10 関連する注視領域 (1)

太い枠で囲われた領域は白塗りされた領域と共起性が高い領域を示す.

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.

## 5.2 注目すべきカメラの選出

本手法では, 多数の監視カメラの中から特に注目する必要があるカメラ群を自動的に選出できる. この処理は, カメラ台数  $C$  が多くなると有効性が増す.

以下, その手順を示す. まず最初に, 注視領域の組をその間でのイベントの共起性についてソートする. そして, 高い共起性を持つ上位の組を選び出し, それに属する注視領域を持つカメラ  $c$  ( $1 \leq c \leq C$ ) の評価スコア  $u_c$  を加算する. 最終的に, スコアが平均  $\bar{u}$  よりも大きいカメラが, “注目すべきカメラ” となる.

図 12, 13 に実験結果を示す. 図中, 太い枠で囲まれた映像が注目すべきと見なされたカメラによって撮影されたものである. カメラ 4 台の (A, B, C, D) の場合では, イベントを監視するために注目すべきカメラ 2 台 (B, D) が選出・提示されている (図 12). 一方, カメラ 12 台の場合では, 注目すべきカメラ 5 台 (F, G, J, K, L) が選出・提示されている (図 13).

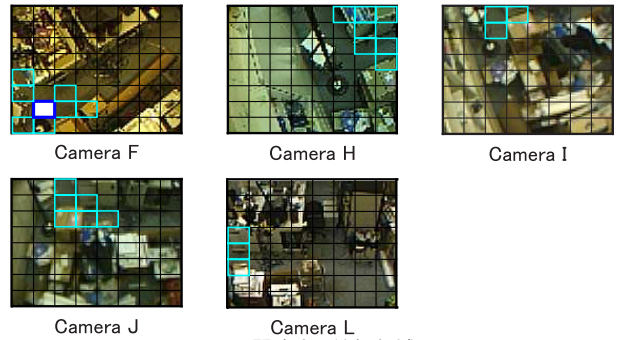


図 11 関連する注視領域 (2)

太い枠で囲われた領域は白塗りされた領域と共起性が高い領域を示す.



図 12 注目すべきカメラ (4 台中 2 台を選出)

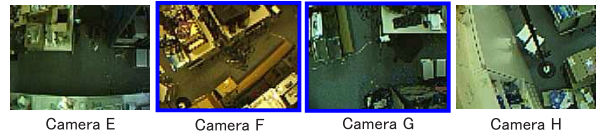
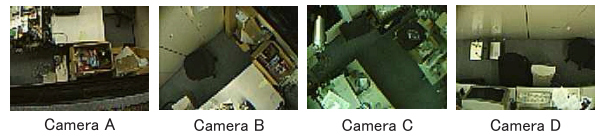


図 13 注目すべきカメラ (12 台中 5 台を選出)

## 6. イベント認識への応用

カメラの台数が多い場合に観測される情報をそのままイベント検出・認識に適用しようとすると, 特徴空間の次元数が大きいために, 処理に時間がかかることが予想される. 本研究では, そのような状況においても適用可能なイベント認識体系を提案することを最終的な目標としている. 現在, 抽出された注視領域群のみを用いた観測イベントのオンラインクラスタリング処理部を実装しており, より効率的なイベントの分類を試みている.

ここで述べるオンラインクラスタリング手法の特徴は,

- クラスタ数が予め指定されていない (シーンに応じてクラスタ数を決定)
  - 複雑な処理を行わないため, 処理時間が短い
- ことである.

手順 1: クラスタの判別 観測イベントと既存のイベントクラスタ群の各中心との距離を求める. 現在はユークリッド距離を用いており, マハラノビス距離など他の距離尺度による判別は今後の課題となっている.

手順 2: クラスタの更新 観測イベントとクラスタ中心との距

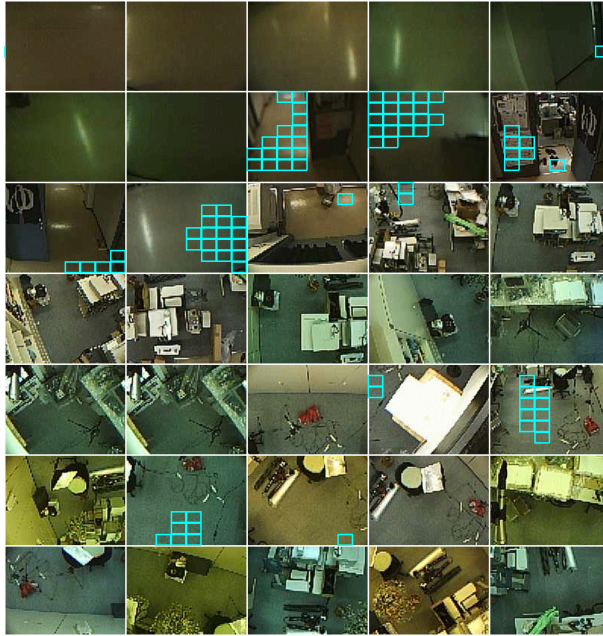


図 14 クラスタの例

離が閾値未満のクラスタがあれば、観測イベントをそのクラスタに追加する。それに伴い、クラスタの中心を更新する。

手順 3: クラスタの追加 観測イベントがどのクラスタとも閾値以上距離が離れている場合には、そのイベントを中心とした新しいクラスタを作成する。

オンラインクラスタリング処理に先んじて、予め蓄積したデータセットにオフラインクラスタリング処理 (K-means 法) を適用し、クラスタ間の距離を求めておく。これを基に手順 2 で必要となる距離の閾値を決定している。

ここで、クラスタの更新・追加時にクラスタ間で重なりが生じてしまったり、クラスタ更新時に登録イベントがそのクラスタから外れてしまう問題が生じる可能性がある。これを回避するために、以下に述べる登録イベントの再配置処理を施す。

- 2 クラスタ間の重複領域に存在するイベントに関しては、再度 2 クラスタの中心との距離を求め、近い方のクラスタに登録し直す。

- クラスタ更新の際にクラスタ領域から外れてしまった場合は、次のいずれかの対策が考えられる。

- 1: 破棄する
- 2: 新しい観測イベントとして入力する

現在は、前者を採用している。

本来ならばクラスタの更新・追加処理を行う必要があるにも関わらず破棄処理を行っているのは、クラスタ再配置手法を多次元データ構造に適用した場合に処理時間が膨大になることが予想されるためである。

本手法をカメラ 35 台のデータに対して適用した。ここで、観測ベクトルの次元数は  $35 \times 64 = 2240$  で、そこから抽出された注視領域の数は 599 となっている。オンラインクラスタリング処理によって得られたイベントクラスタの一例を図 14 に示す。ここで注意されたいのは、図 14 で強調された領域全てが一つのイベントクラスタに対応することである。このクラスタは、実験室中に一人の人物が同じ場所に立ち続けていて、同

時に廊下に複数の人物が立っているイベントと見なせる。このように得られたイベントクラスタは、イベントの認識・分類に利用できると考えられる。今後は、その他のクラスタリング手法の評価と、実際のイベント認識に取り組む予定である。

## 7. 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支援ツールの例として、“関連する注視領域群の提示”と“注目すべきカメラの選出”を実装し、その有効性を示した。また、得られた注視領域を用いたイベント認識への手がかりとして、注視領域を基にした観測イベントのクラスタリングを実装し、注視領域の抽出がイベント認識において有効である可能性を示した。

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等を実現していく予定である。

## 文 献

- [1] M. D. Beynon, D. J. Van Hook, and M. Seibert abd A. Peacock, “Detecting abandoned packages in a multi-camera video surveillance system,” in *IEEE Int. Conf. on Advanced Video and Signal Based Surveillance*, 2003, pp. 221–228.
- [2] T. Chang, S. Gong, and E. Ong, “Tracking multiple people under occlusion using multiple cameras,” in *11th British Machine Vision Conference*, 2000.
- [3] N. T. Nguyen, S. Venkatesh, G. A. W. West, and H. H. Bui, “Hierarchical monitoring of people’s behaviors in complex environments using multiple cameras,” in *16 th Int. Conf. on Pattern Recognition*, 2002, Volume I pp. 13–16.
- [4] G. Wu, Y. Wu, L. Jiao, Y. Wang, and E. Y. Chang, “Multi-camera spatio-temporal fusion and biased sequence-data learning for security surveillance,” in *11th ACM Int. Conf. on Multimedia*, 2003, pp. 528–538.
- [5] G. Stein, R. Romano, and L. Lee, “Monitoring activities from multiple video streams: Establishing a common coordinate frame,” in *IEEE Transactions on Pattern Analysis and Machine Intelligence August 2000 (Vol. 22, No. 8)*, 2000, pp. 258–267.
- [6] O. Javed, K. Shafique, and M. Shah, “Appearance modeling for tracking in multiple non-overlapping cameras,” in *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, 2005, pp. 26–33.
- [7] P. KaewTraKulPong and R. Bowden, “A real-time adaptive visual surveillance system for tracking low resolution colour targets in dynamically changing scenes,” in *Image and Vision Computing, Volume 21, Number 10*, 2003, pp. 913–929.
- [8] V. Kettner and R. Zabih, “Bayesian multi-camera surveillance,” in *IEEE Conf. on Computer Vision and Pattern Recognition*, June 1999, pp. 2253–2261.
- [9] C. Stauffer and K. Tieu, “Automated multi-camera planar tracking correspondence modeling,” in *Computer Vision and Pattern Recognition*, 2003, pp. 259–266.
- [10] 服部傑, 亀田能成, 大田友一, “非日常性の認識に向けた多数センサの関連付け”, 電子情報通信学会 技術研究報告 MVE, Vol.104, No.489, ISSN 0913-5685, MVE2004-43, Dec 2004, pp.19-25.