

非較正カメラ群を用いた多視点同時観測におけるシーンの判別

西崎 隆志[†] 金成 幸司[†] 亀田 能成^{††} 大田 友一^{††}

† †† 筑波大学大学院システム情報工学研究科

〒 305-8573 茨城県つくば市天王台 1-1-1

E-mail: †{tanishi,kanari}@image.esys.tsukuba.ac.jp, ††{kameda,ohta}@iit.tsukuba.ac.jp

あらまし 近年、様々な環境に複数のネットワークカメラを設置して日常的な出来事を検出・認識し、それを生活支援や福祉に役立てることに対する社会的 requirement が高まってきている。物体や出来事の認識を行う研究に関しては、単数もしくは少数のカメラを利用した研究がこれまで報告されている。しかしながら、ネットワークカメラの低廉化に伴い、多数のカメラを配置した状況下での、カメラ較正を必要としないシーン識別法が求められている。本論文では、多数の非較正カメラ群を用いた多視点同時観測において、カメラの配置や撮影対象との位置関係に制約を設けないまま、そのシーン内で動く物体に注目して観測シーンを判別する手法について提案する。大規模なカメラシステムにおける観測シーンの判別を実現するために、本手法では、多視点映像の中で閲覧者ないしシーン判別システムが特に注目すべき重要な領域を“注視領域 (region of interest: ROI)”と定義し、抽出された注視領域群の評価のみに基づいた観測シーンの判別を行う。

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

Scene Classification in Multi-View Simultaneous Observation Using Multiple Non-Calibrated Cameras

Tanishi NISHIZAKI[†], Kouji KANARI[†], Yoshinari KAMEDA^{††}, and Yuichi OHTA^{††}

† Graduate School of Systems and Information Science, University of Tsukuba.

Tennnoudai 1-1-1, Tsukuba-shi, Ibaraki, 305-8573 Japan

E-mail: †{tanishi,kanari}@image.esys.tsukuba.ac.jp, ††{kameda,ohta}@iit.tsukuba.ac.jp

Abstract Recently, there are many demands to detect and recognize events in daily life by utilizing network cameras installed in various locations in an environment so as to support lives and improve welfare in our community. Previous works have been worked mainly with a single camera or a few cameras for event detection and recognition. However, as network cameras become less expensive and a massive number of the cameras can be utilized in a system, a new scene classification method that does not require camera calibration is desired. In this paper, we propose a scene classification method based on simultaneous multi-view observation by utilizing non-calibrated cameras. It does not assume any limitation about motion of objects and camera placements. We define “regions of interest: ROI” as image regions that a scene classifier focuses on carefully in multi-view observations. The proposed method extracts them according to statistical analysis and it can realize automated scene classification by using the ROIs.

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

1. はじめに

近年、様々な環境に複数のネットワークカメラを設置し、日常生活環境で発生する出来事を検出・観測することへの要求が社会的に高まっている。そのようなネットワークカメラは、特に公共の場でのセキュリティや道路交通監視、室内での生活

支援や老人福祉に利用可能であり、今後さらに設置数が増えると考えられる。

ここで一例として、図 1 に同時に撮影された 36 台のカメラのスナップショットを示す。これらのカメラは、ある一部屋の天井と、その部屋に隣接する廊下の天井に設置されている。

大量のネットワークカメラが設置・運用されるようになると、

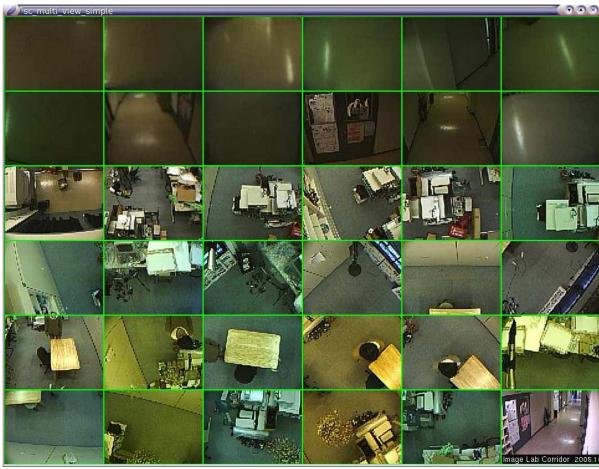


図 1 多視点映像の例（36台のカメラ）

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、といった情報が複雑化するため、従来から提案されてきている観測シーン認識手法では扱いづらくなりつつある。また、設置されるネットワークカメラの台数が増えていくと、膨大なデータ処理を行わなくてはならず、そのままでは処理に時間が掛かることが予想される。加えて、従来の映像観測システムはカメラ較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、厳密なカメラ較正を適用することは現実的ではなくなる。よって、従来行われてきた、カメラ配置が既知であることを前提にする手法を適用することは難しくなる。したがって、幾何的制約が一切与えられない大量のカメラを対象とし、得られる膨大な映像情報のみを用いて観測シーンの判別・認識ができるような知的映像観測システムが望まれる。

本論文では、多数の非較正カメラ群を用いた多視点映像観測において観測シーンを判別する手法について提案する。大規模なカメラシステムにおける観測シーンの判別を実現するために、本手法では、大規模多視点映像の中で閲覧者ないしシーン判別システムが特に注目すべき重要な領域を“注視領域（region of interest: ROI）”と定義し、統計的手法に基づいてこれを抽出する。図2に注視領域群の例を示す。図中、明るい色の枠で強調された領域が抽出された注視領域群に対応する。全映像データを処理するのではなく推定された注視領域だけを扱うようすれば、多視点映像の解像度や数に関してスケーラビリティを有した観測シーンの判別・認識が実現できると考えられる。また、注視領域群の抽出において厳密なカメラ較正を必要とせずに映像間の対応付けを行うため、注視領域群に基づくシーン判別手法は多数のカメラが無作為に分散設置された状況でも適用できる。本研究では、注視領域抽出や観測シーンの判別において撮影対象の大きさや行動、カメラの配置状況に制約を設けていないので、本手法は様々な状況・様々な目的に適用可能なものとなっている。

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。続いて、3. 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

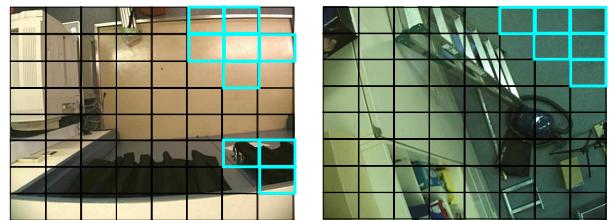


図 2 抽出された注視領域群の例

て述べる。そして、4. において注視領域群の評価に基づく観測シーンの判別について述べる。最後に、5. において、本論文をまとめる。

2. 関連研究

人物の日常生活や日常的な交通監視などを対象とした映像観測システムの研究では、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、複数のカメラを用いる事が多い。しかし、カメラの設置台数が大量になってくると、目的や条件設定に依存した映像観測手法ではスケーラビリティに問題が出てくることがある。そこで、目的や条件設定に依存しない、汎用的な観測手法が必要となってくる。

カメラ較正を前提とする従来の複数カメラ映像観測システムとして、[1][2][3] が挙げられる。これらの研究では正確な幾何情報を求めることが可能、高い精度の物体位置検出・追跡を実現できる。しかし、規模の大きいカメラシステムにおいて、カメラ較正は非常に煩わしい作業となりうる。

カメラ間の厳密な位置情報を必要としない手法として、動物体の軌跡を用いる方法が提案されている[4][5]。これらは、人が平面上を歩く、カメラの撮影領域が隣接している^(注1)、等の仮定を設けている。これらはカメラ較正を必要としないものの、軌跡を評価する手法のために最終的な追跡精度は各カメラ上での軌跡検出の精度に依存し、撮影対象がカメラに近いような状況では正しく動作しない場合がある。

一方、対象とするイベントを予め記述したテンプレートが適切に準備できていれば、正確で信頼に足るイベント認識が実現できることが報告されている[6][7]。しかし、これらの手法では、予期していないイベントや複雑なイベントを扱うことが出来ない、カメラの台数が大量になってくるとテンプレートの記述が面倒になるなど、日常的なシーンを対象とした自由度の高いイベント認識には不向きである。

また、ある程度の条件設定を用意することによって、人物や動物体のより複雑な動作を扱うイベント認識を試みる研究がなされている[8][9][10]。これらの研究は高い応用性を実現している。しかし、これらは条件設定を必要とするため、やはり様々な撮影対象・様々な撮影環境に対応するにはまだ自由度が不十分である。

これらに対し、本論文では広く多岐にわたる状況に適応可能な観測シーンの判別手法を提案する。本手法の特徴として、以

(注1): 人物があるカメラの撮影領域から外れると、他のカメラの撮影領域に入ることになる



図 3 入力画像（左）と特徴量（右）

下の点が挙げられる。

- 撮影対象の大きさ、色、挙動などの定義は不要
- 物体の軌跡情報は不要
- カメラ配置の幾何的情報は不要
- カメラ台数に対するスケーラビリティ

3. 注視領域 (ROI) の抽出

本研究では、大規模多視点映像群の中で、観測シーンの判別・認識システムが最低限注目すべき映像領域を“注視領域 (region of interest: ROI)”と定義している。一定時間蓄積された多視点映像データを統計的に評価し、映像観測に有用であると見なされた領域のみを残すことで、注視領域群を抽出する。

3.1 画像特徴量と多次元データ構造

撮影された映像からは様々な画像特徴が抽出できるが、本研究では撮影対象の種類を限定しないため、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、保存するデータ量を小さくする。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。この際、特徴量の値域を、輝度値の値域である [0, 255] から [0, 63] に縮めている。これにより、光源の微少変動等の環境要因が注視領域抽出に与える影響を抑える。保存される情報の例を図 3 に示す。これらの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。

画像処理用 PC 群に保存された各カメラの画像特徴データを 1 台の PC に集約し、以下に述べる注視領域抽出処理を適用する。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 は観測ベクトルの総数) 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シーンを表現できる。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、この特徴量として、前景領域の輝度値平均を直接用いている。

この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用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有用要素の選出（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、本章で述べる有用要素の選出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 を注視領域と定義する。

以下、有用要素の選出に基づいた注視領域の抽出手法について述べる。

3.2 有用要素の選出

本節では、3.1 節で定義した N 次元の観測ベクトル \mathbf{x} に対する有用要素の選出について述べる。観測ベクトルに対する有用要素選出手法は、“前景が観測されない領域の除去”，“主成分分析に基づく次元数の削減”の二つの手順に分けられる。

3.2.1 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 \mathbf{M} と分散共分散行列 \mathbf{V} を計算する。本研究において観測される特徴量 x_i は 0 以上であるため、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とみなせる。したがって、この領域 i を観測に不要な領域として破棄する。この処理を適用することによって、 N' ($N' \leq N$) 次元の観測ベクトル $\mathbf{x}'(t)$ を得る。

3.2.2 主成分分析に基づく次元数の削減

次に、分散共分散行列 \mathbf{V} を用いて、 N' 次元の観測ベクトル $\mathbf{x}' = \{x'_j\}$ 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、もとになるベクトル \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、他の要素を除去する。以下、有用要素の選出手法の詳細について述べる。

STEP 1: 主成分群 $\{z_k\}$ をその寄与率 p_k が大き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 λ_k を用いて以下のように表される。

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

STEP 2: 累積寄与率

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

がある閾値 c_{th} (本稿では $c_{th} = 0.9$ 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。

STEP 3: STEP2 で寄与率が高いとして残された主成分群 z'_k の各々について次の処理を行う。ある主成分 z'_k において、 z'_k を作る $\{x'_j\}$ の係数群の絶対値 $\{|a_{jk}\|$ の平均値 \bar{a}_k を求め

る。係数の絶対値 $|a_{jk}|$ 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。

STEP 4: 観測ベクトル $\{x'_j\}$ の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。

以上の処理より、 N_v 次元の注視領域ベクトル \mathbf{v} を得る。

3.3 注視領域抽出の実験結果

実験システムとして 35 台のカメラを設置した。カメラの配置を図 4 に示す。ネットワークカメラ群は実験室の天井と実験室に面する廊下の天井に設置した。主に下方に向けて設置されており、床面の状況を撮影できるようになっている。また、いくつかのカメラに関しては部屋全体を見通すよう、床面に対し斜め方向に設置した。このうち、カメラ 4 台 (CAM 25, 26, 27, 28) を用いた注視領域抽出結果を図 6 に、カメラ 12 台 (CAM 13, 14, 15, 16, 21, 22, 23, 24, 25, 26, 27, 28) を用いた注視領域抽出結果を図 7 に示す。ここで、水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。よって、多視点映像の閲覧者や観測シーンの判別・認識システムは、この抽出された注視領域群を優先してチェックすればよい。カメラ 4 台の場合、観測ベクトルの次元数は $N = 256$ であり、得られた注視領域ベクトルの次元数は $N_v = 87$ であった。また、カメラ 12 台の場合、 $N = 768$ であったのに対し、 $N_v = 178$ が得られた。このように、対処となるカメラの台数が多数であった場合でも、注視領域を抽出することによって、より効率的な映像観測が実現できる。

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。まず約 3 時間の多視点映像に対して観測ベクトルを事前にオンラインで求め、137,351 個の観測ベクトルを得た。この観測ベクトルに対し注視領域抽出にかかる処理時間は、カメラ 4 台の場合に 1467.57 秒（平均の計算に 142.16 秒、分散共分散行列の計算に 1324.68 秒、冗長要素削減に 0.73 秒）、カメラ 12 台の場合に 2157.85 秒（平均の計算に 474.38 秒、分散共分散行列の計算に 1677.75 秒、冗長要素削減に 5.72 秒）となった。この処理時間は撮影時間、カメラの台数以外に、蓄積データ中の有意な情報の偏り方の影響を受けると考えられる。

4. 観測シーンの判別

本手法によって抽出された注視領域群のみを評価することによって、従来手法に比べて効率的な観測シーンの認識が実現できると考えられる。本章では、観測シーンを認識するために重要であると考えられる、観測シーンの判別について述べる。本研究では、オンラインでシステムに入力される映像データに関して、類似した注視領域ベクトル $\mathbf{v}(t)$ のクラスタ群を形成することを、観測シーンの判別としている。あるクラスタのあるメンバ $\mathbf{v}(t)$ は、時刻 t において注視領域群で観測される人の動作の一例を示し、生成されたクラスタは注視領域群上でみた類似した人物の動作の集まりを示す。

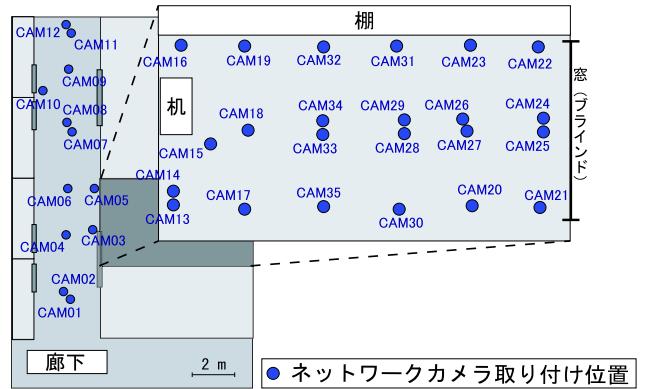


図 4 カメラ配置



図 5 天井に配置されたカメラの様子（左図は廊下、右図は実験室内の様子）

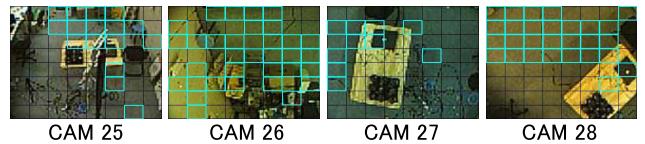


図 6 得られた注視領域 (カメラ 4 台の場合)

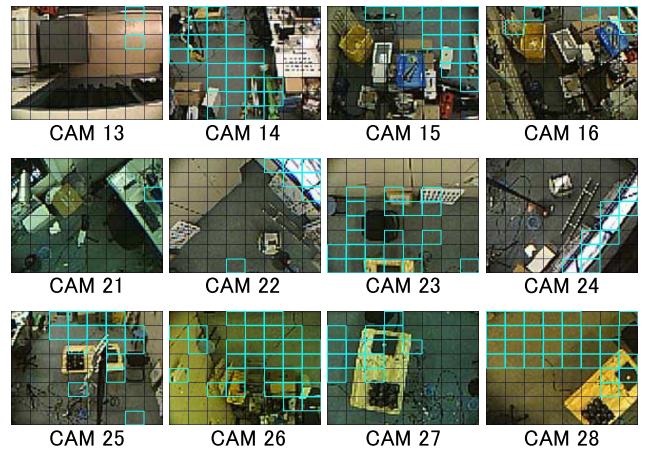


図 7 得られた注視領域 (カメラ 12 台の場合)

4.1 クラスタ間距離

クラスタ間のユークリッド距離の閾値を決定するために、オフラインクラスタリングを行う。本稿では実際には、K-Means 法がある一定時間のデータセット $\{\mathbf{v}(t)\}, (1 \leq t \leq T)$ に適用した。クラスタ数 K は、観測シーンの分類においてクラスタの粒度をどのように設定するのかに応じて設定される数である。クラスタ数 K の設定方針は様々なものが考えられるが、ある時間 T のシーンにおいて見つけ出したい人の動作の数を設定す

ることが指針の一つであると考えらる。K-Means 法によって K 個のクラスタが形成された後、クラスタ間の最短距離をオンラインクラスタリングにおけるクラスタ間距離閾値 D_{th} として用いる。なおこのクラスタ間距離の算出は、一般には、本稿で行ったようにオンライン時の観測データを用いる必要はもちろんな。しかし、同一環境、同一システムのデータを用いた方が、その閾値はより有用であるとは考えられる。

4.2 オンラインクラスタリング

次元数 N_v の注視領域ベクトル群 $\{v(t)\}$ をオンラインクラスタリングへの入力とする。本手法では、特微量空間中に初期クラスタを設置しない状態でオンラインクラスタリングを開始する。我々が用いたオンラインクラスタリング手法を以下に述べる。

STEP 1: ゼロベクトルの無視 もし入力ベクトル $v(t_l)$ がゼロベクトルであった場合、非ゼロベクトルが入力されるまで何もせず待つ。

STEP 2: クラスタ判別 入力ベクトル $v(t_l)$ に関して、既存の各クラスタの中心とのユークリッド距離を計算する。そして、その中から最短となる距離を選出する。

STEP 3: クラスタ更新／追加 STEP 2 で得られた最小距離がクラスタ間距離閾値 D_{th} よりも小さい場合、入力ベクトル $v(t_l)$ を最短距離となったクラスタに登録する。大きい場合は、入力ベクトル $v(t_l)$ を中心とする新しいクラスタを追加作成する。

STEP 4: クラスタ再構成 クラスタに新しい入力ベクトルが追加登録された場合、クラスタ中心はそれに応じて変動する。従って、そのクラスタに登録されているメンバ群と、そのクラスタに隣接するクラスタのメンバ群に関しては、どちらのクラスタに属するかを再度判別する必要がある。したがって、これらのメンバに関しては、クラスタ中心との距離を再計算し、再度最近傍のクラスタに登録し直す。

なお STEP 4 で、両方のクラスタから離れてしまったメンバに関しては、再度オンラインクラスタへの入力とするか、破棄するという二つの対処が考えられるが、本手法では高速性を踏まえ、後者を採用している。

4.3 実験結果

本節ではオンラインクラスタリングの実験結果を示す。撮影環境となった実験室には、壁付近に複数個のロッカーがあり、中央付近に作業用の机と椅子が置かれていた。撮影中、人物は実験室に靴を脱いで入室し、机の前に座り机上作業を行った。また、実験室への入退出を何度も行ったり、部屋内に置かれている物体の移動なども行った。

まずクラスタ間距離閾値 D_{th} を求めるための K-Means 法に基づくオフラインクラスタリングでは、 $K = 20$ にして処理を行った。本実験では、注視領域抽出と観測シーンのオンラインクラスタリングについて同じ条件で検証できるよう、同じ 3 時間の映像データセットをハードディスクから読み出して入力として用いた。ただし、我々が実装したオンラインクラスタリングはリアルタイムでも実行可能である。本実験において、入力となった注視領域ベクトル数は 137,351 であり、その内非ゼ

ロベクトル数は 44,122 であった。

今回の実験では、クラスタ間距離閾値 D_{th} のためのオフラインクラスタリングと、観測シーンの分類のためのオンラインクラスタリングと同じ PC を用いて行った。クラスタ間距離閾値 D_{th} の計算には 1341.77 秒かかり、映像データセットに対するオンラインクラスタリングには 2457.49 秒かかった。観測時間が約 10800 秒 (3 時間) であることから推測すると、このオンラインクラスタリングにかかった処理時間から、オンラインクラスタリングがリアルタイムで実行できることが分かる。

本実験において、3 時間の映像データ全てが入力された後には、762 個のクラスタが生成された。図 8, 9, 10 に生成されたクラスタの例を 3 種類示す。また、図 11 に各クラスタのメンバ数 (各クラスタに登録された注視領域ベクトルの数) を示す。図 8 は 665 個のメンバを持つクラスタに対応し、以後これをクラスタ α と呼ぶ。また、図 9 は 2,422 個のメンバを持つクラスタに対応し、以後これをクラスタ β と呼ぶ。そして、図 10 は 2 個のメンバを持つクラスタに対応し、以後これをクラスタ γ と呼ぶ。図中、明るい色の枠で囲まれた領域は、前景領域が度々観測された注視領域を示す。

求まったクラスタについて解説を試みると、クラスタ α は人物が机上作業を行っている様子を示し、クラスタ β は無人の実験室における光源環境の変化を示している。クラスタ α では、木製の机の前に人物が座っており、机上にあるノート型 PC を用いて作業を行っている様子が確認できる。クラスタ β では、窓領域で前景領域が確認されているが、調査の結果、実際にクラスタ α の同じ領域に比べ輝度値が変化していることが分かった。実際に撮影を行ったのは夕方の比較的日照変動の多い時間帯であったことから、実験結果は妥当であると言える。このように、本手法に環境の変動をシーンの違いとして捉える性質があることが確認できた。また、クラスタ α のメンバ数はその他多くのクラスタに比べて多かったが、クラスタ β に比べると少ない結果となった。クラスタ β のメンバ数は他のクラスタに比べて突出して多かったことから、部屋内が無人でありながら光源変化だけが生じていた状態が頻繁に観測されたことを意味する。実際、人物は実験室への入退出を幾度と行っており、本実験室以外でも作業を行っていたため、3 時間に渡って常に在室していたわけではない。一方、クラスタ γ はメンバ数が 2 個と極端に少ない。このクラスタは、人物が部屋内の机付近を移動している状態を示す。このように、本手法では物体が移動する観測シーンに対応するクラスタのメンバ数が少なくなってしまう傾向がある。これは本手法が時空間的継続性を考慮に入れていないため、“移動” というイベントに対するクラスタを一つにまとめることが出来ないからである。図 11 の内、メンバ数が非常に少ないクラスタ群が多数確認できるが、これらが物体が移動する観測シーンに対応していると考えられる。

本節で示した実験結果から、我々の手法がカメラの配置や物体の挙動等に関する先見知識を必要とせずにシーン中の物体の挙動のクラスタリングを実現できていることが示せた。しかし、本論文での議論はまだ移動に対応できていない等改良の余地があり、今後より詳しい解析と実験を行っていく必要がある。

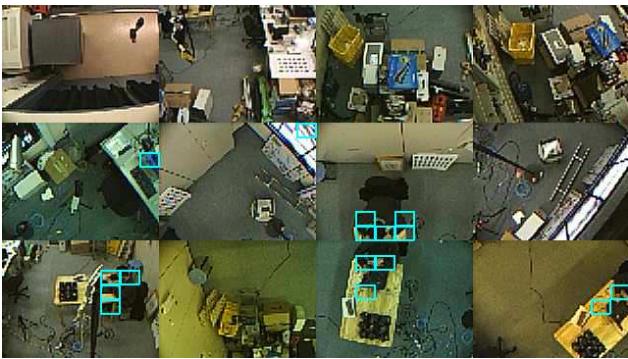


図 8 クラスター α の注視領域群



図 9 クラスター β の注視領域群

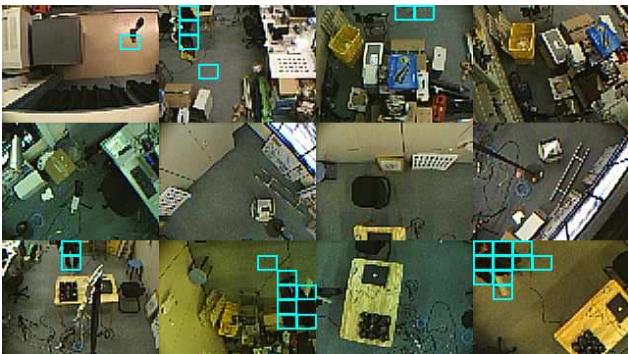


図 10 クラスター γ の注視領域群

5. まとめ

本論文では、多数の非較正カメラ群を用いた多視点同時観測において、撮影対象の大きさや行動、カメラの配置状況に制約を設げずに観測シーンを判別する手法について提案した。大規模なカメラシステムにおける観測シーンの判別を実現するため、本手法では、大規模多視点映像の中で閲覧者ないしシーン判別システムが特に注目すべき重要な領域を“注視領域 (region of interest: ROI)”と定義した。この注視領域の抽出方法の詳細について述べ、注視領域の抽出結果を示した。そして、抽出された注視領域群の評価に基づいた観測シーンの判別を実現し、その結果を提示・考察した。

今後の課題として、観測シーンの判別 (クラスタリング) 手法の改良と観測シーンの認識体系の構築が挙げられる。例えば、現在のクラスタリング手法では、各クラスタ内での観測ベクトル

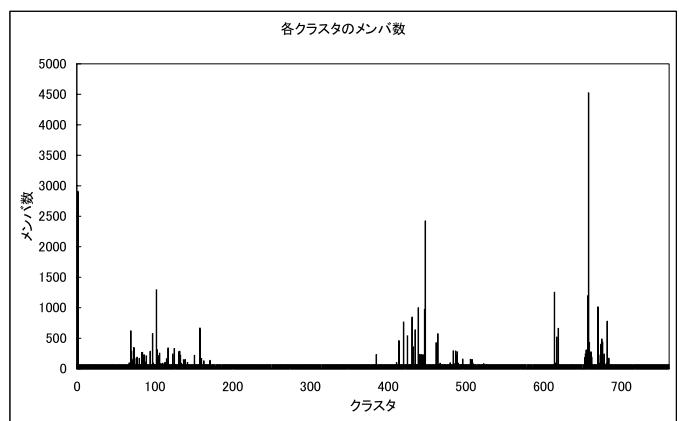


図 11 各クラスタのメンバ数

ル群の分布モデルを考慮していない。将来的には、クラスタ間距離尺度にマハラノビス距離を用いるなど、確率空間におけるクラスタ形状の記述を行っていくことも必要と考えられる。また、分類された観測シーンに意味付けを行い、注視領域抽出を実際の状況認識へ応用していくことも検討課題である。

文 献

- [1] 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,” Int. Conf. on Pattern Recognition, vol.1, pp. 13-16, 2002.
- [2] M. D. Beynon, D. J. Van Hook, and M. Seibert abd A. Peacock, “Detecting abandoned packages in a multi-camera video surveillance system,” IEEE Int. Conf. on Advanced Video and Signal Based Surveillance, pp. 221-228, 2003.
- [3] 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,” 11th ACM Int. Conf. on Multimedia, pp. 528-538, 2003.
- [4] C. Stauffer and K. Tieu, “Automated multi-camera planar tracking correspondence modeling,” IEEE Computer Vision and Pattern Recognition, pp. 259-266, 2003.
- [5] K. Tieu, G. Dalley, and W. Eric. L. Grimson, “Inference of non-overlapping camera network topology by measuring statistical dependence,” Int. Conf. on Computer Vision, pp.1842-1849, 2005.
- [6] G. L. Foresti, C. Micheloni and L. Snidaro, “Event classification for automatic visual-based surveillance of parking lots,” Int. Conf. on Pattern Recognition, vol.3, pp.314-317, 2004.
- [7] S. J. McKenna and H. Nait-Charif, “Activity summarisation and fall detection in a supportive home environment,” Int. Conf. on Pattern Recognition, vol.4, pp.323-326, 2004.
- [8] M. Higuchi, S. Aoki, A. Kojima, and K. Fukunaga, “Scene recognition based on relationship between human actions and objects,” Int. Conf. on Pattern Recognition, vol.3, pp.73-78, 2004.
- [9] A. H. Kam, K. A. Toh, H. L. Eng, W. Y. Yau and J. Wang, “Automated recognition of highly complex human behavior,” Int. Conf. on Pattern Recognition, vol.4, pp.327-330, 2004.
- [10] M. Leo, T. D’Orazio, I. Gnoni, P. Spagnolo, A. Distante, “Complex human activity recognition for monitoring wide outdoor environments,” Int. Conf. on Pattern Recognition, vol.4, pp.913-916, 2004.