

道路画像におけるパン・チルト・ズームカメラの状態の推定

濱村 亮¹、田中 雅博²

¹ 甲南大学大学院自然科学研究科

² 甲南大学工学部・知的情報通信研究所

(受理日 2005年4月1日)

概要

In this paper, we consider the problem of estimating the state of camera from the real-time traffic images obtained with the fixed camera with pan, tilt and zoom functions. This problem comes from the research of image processing of the road image for traffic state estimation. Since this camera can be controlled from the authorities of political surveillance, the camera is subject to move suddenly.

The state of the camera means the camera direction and the amount of zooming. This processing needs real-time processing from the nature of the problem. If we apply the simple 2-D matching method, it requires a large processing time.

Hence, in this paper, we will use two 1-D vectors for matching using two low-dimensional vectors of the vertical and horizontal average values of the pixels of certain area in the image. The feature spaces are defined from the primal eigenvectors of sample images, and the distance between the current image and the feature space is used for matching. Moreover, when we estimate the zoom magnification, eigenimage is derived from the non-zoomed images after positional matching.

Similar methods are applied both for the calibration of camera direction (pan and tilt) and zooming.

Keywords: PTZ camera, camera condition, image processing, eigenspace

1 はじめに

Intelligent Transportation System (ITS)においては、固定したビデオカメラから画像を取り込み、画像認識を行って、道路上の交通状況を把握することが1つのテーマである。画像処理については様々な研究がなされており[3,4]、画像からの車両の検出[1]、シーンの分析[2]、背景の適応推定[8,7]などが試みられている。

我々は、英国 Nottingham Trent 大学および Nottingham 交通局と共同で、交通画像処理システムの開発を手がけており ([5,6])、表1のような階層構造による画像分析を行い、携帯電話のインターネットアクセス機能を用いて、携帯電話から混雑度をチェックすることができるようにする計画である。

表1: 画像処理の層

第4層	混雑度をイラスト化してWEB上に公開、 携帯からアクセス可とする
第3層	混雑度の推定
第2層	車の検出
第1層	カメラの姿勢の同定
第0層	画像取得

本研究で扱う画像は、Nottingham警察当局が市内の道路上の防犯管理の監視のためにも用いており、そこからカメラのパン、チルト、ズーム (PTZ) により姿勢が変更されることがあるため、画像処理の前処理として、カメラ姿勢の推定が必要である。図1、2に、同じ道路を撮影したときに生じるPTの復帰が不完全な様子を示す。

この場合、これらの移動量をカメラを動かすほうのシステムから信号の供給を受ければ、より正確な移動量やズームの倍率を把握できると思われるが、本システムではカメラの制御信号が取り出せるように設計されていない。こういう状況のもとでの画像処理を研究することはある程度の応用につながると考えられる。そこで、本研究では、PTの移動量から画像の位置合わせをし、ズームを検出しズーム倍率を推定する手法を開発する。



図1: 同じ場所を撮影した画像 (1)



図2: 同じ場所を撮影した画像 (2)

2 PTによるカメラの移動量とズームの有無検出及びズーム倍率の推定

2.1 方針

マッチングの対象となる画像があらかじめ登録されたものと酷似するような状況であれば、単純差分をベースとする簡単な処理で移動や倍率の検出は可能であろう。しかし、ここで得られる画像は、さまざまな気象条件や日照条件のもとで撮影されるものであり、そのような手法ではエラーが起こる可能性が高い。マッチングさせるための手がかりは、条件により様々に変化するとはいえ、画像の中に含まれる道路や建物の大まかな類似性であ

る。そうした類似性を抽出するために、主成分分析による固有空間の抽出を行い、画像がPTZにより変化しても必ず現れる小領域を定義し、その部分の不変成分を求める。また、小領域といえども画素の数は多いため、そのまま1次元化して用いず、その小領域の縦と横方向への画素値の平均値を求め、1次元ベクトル2本として固有空間への最短距離を評価規準として用いることにする。なお、画像処理を軽くするため、768×576のカラー画像をモノクロで処理する。

2.2 モデル画像の作成

モデル画像を以下の方法で作成する。

1. 画像を1枚準備する。この画像をAとする。
2. ズームなしの同じ場所での画像について、次のステップを繰り返す。
 - (a) 画像Bを準備する。
 - (b) 適当なプログラムを使い、Aの上にBを手動でスーパーインポーズする。図3にその様子を示す。



図3: 元画像と小領域

- (c) 画素値の x 軸方向、 y 軸方向の和を求める。
 - (d) ファイル名とともに縦横の平均画素値のデータをディスクに保存する。
3. 上記の作業を画像複数枚について行い、それらの合計から x 軸方向、 y 軸方向の平均画素値を求める。

2.3 モデルの生成

モデル生成のためのトリミングされた画像を、矩形領域

$$P_m = \{p_m(i, j) | 1 \leq i \leq N_1, 1 \leq j \leq N_2\} \quad (1)$$

$$P_m, m = 1, \dots, C \quad (2)$$

とし、各画像はこのように表現する。また、この矩形の横と縦の方向に値を平均化したベクトルをそれぞれ x_m と y_m とすると、各要素は、 x_m と y_m を列ベクトル

$$x_m(i) = \frac{1}{N_2} \sum_{j=1}^{N_2} P_m(i, j), \quad i = 1, \dots, N_1 \quad (3)$$

$$y_m(j) = \frac{1}{N_1} \sum_{i=1}^{N_1} P_m(i, j), \quad j = 1, \dots, N_2 \quad (4)$$

で与えられる。次の処理は、 x と y それぞれについて同じなので、以後は x のケースのみについて記載する。

$\{x_m\}$ の共分散行列を

$$M = \frac{1}{C} \sum_{m=1}^C (x_m - \bar{x})(x_m - \bar{x})^T \quad (5)$$

とする。このとき \bar{x} は平均

$$\bar{x} = \frac{1}{C} \sum_{m=1}^C x_m \quad (6)$$

である。

次に、上記の式より求められる固有値を降順に並べる。すなわち

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_C$$

となるように固有値を並べ、固有値 λ_k に対応する固有ベクトルを v_k とする。ここでは、各固有ベクトルは長さを 1 に正規化されているものとする。適当な整数 r を選ぶことにより特徴ベクトル空間

$$F = \left\{ \begin{bmatrix} v_1 & v_2 & \dots & v_r \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{bmatrix} \right\} \quad (7)$$

を得る。ここに、 $a_i (i=1, \dots, r)$ はスカラーである。

ここに得られた正規直交ベクトル v_1, v_2, \dots, v_r を特徴ベクトル、それらにより張られる空間を特徴ベクトル空間と呼ぶことにする。さらに、今後の展開のために、行列

$$V = [v_1 v_2 \dots v_r] \quad (8)$$

$$\theta = [a_1 a_2 \dots a_r]^T \quad (9)$$

を定義しておく。

この実験で使用した画像は2004年4月1日から8月9日までの間の128枚である。図4にそれらの画像の一部を示す。これらの画像には好天のものを最多にして、曇天や雨天、夜間のもの、さらには車両の多いものや少ないものなども含めた。

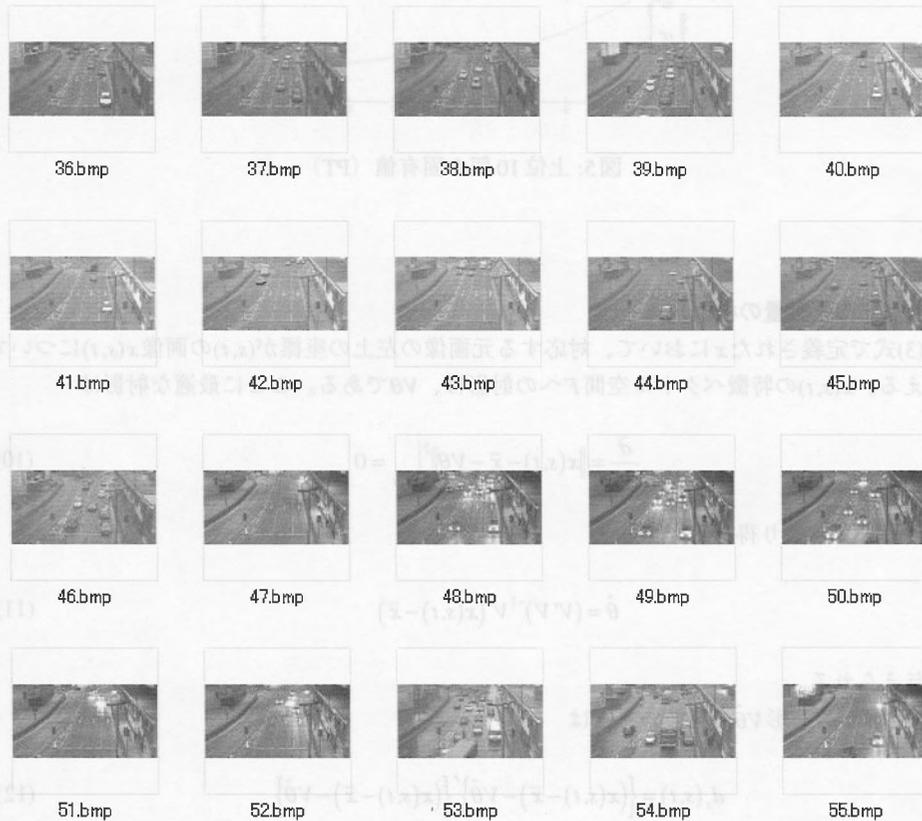


図4: モデル画像

r の決定については、実験的に行った。図5はPTの移動量検出の際に計算した、上位10個の固有値 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{10}$ を示している。下の図が横方向のベクトルに対する固有値であり、上の図が縦方向に対する固有値である。両方とも第1、第2主成分以外はほとんど寄与がなく、大きい方から2、3個の固有ベクトルを使えば十分なことがこのグラフからわかる。なお、この図の縦軸は対数軸である。

ズームにおける上位10個の固有値 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{10}$ のグラフを図5に示す。

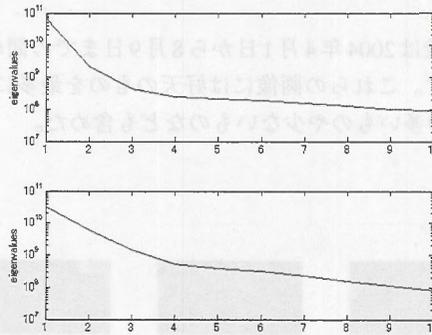


図5: 上位10個の固有値 (PT)

2.4 PTの移動量の検出

(3)式で定義された x において、対応する元画像の左上の座標が (s, t) の画像 $x(s, t)$ について考える。 $x(s, t)$ の特徴ベクトル空間 F への射影は、 $V\theta$ である。ここに最適な射影は

$$\frac{\partial}{\partial \theta} \left\| x(s, t) - \bar{x} - V\theta \right\|_{\theta=\hat{\theta}}^2 = 0 \quad (10)$$

を解くことにより得られ

$$\hat{\theta} = (V'V)^{-1}V'(x(s, t) - \bar{x}) \quad (11)$$

で与えられる。

次に x との射影 $V\hat{\theta}$ の距離の二乗は

$$d_x(s, t) = \left\{ (x(s, t) - \bar{x}) - V\hat{\theta} \right\}' \left\{ (x(s, t) - \bar{x}) - V\hat{\theta} \right\} \quad (12)$$

で与えられる。このとき $V\hat{\theta}$ は

$$V\hat{\theta} = V(V'V)^{-1}V'(x(s, t) - \bar{x}) \quad (13)$$

ここで、 $V(V'V)^{-1}V'$ の部分は、画像 x を得る前に計算可能である。

上記の手順は、同じように、 y についても実施し $d_y(s, t)$ を求めることができる。

なお、 $d_x(s, t)$ 、 $d_y(s, t)$ を使って、それぞれの値が最小となるような (s, t) を探索しなければならない。同様に、 d_x は同じ s に対して t 方向にはほとんど変化がない。これは、これらのベクトルが別の方向についてはほぼ同じ画像の平均を取っているものであるためであり、当然である。従って、別々に探索してもほぼ正確に最適な座標を求めることができる。

このようにして、モデルへの帰属誤差を求め最小となる点を求めることにより、最適な

座標を計算することができる。この節についてのアルゴリズムの構成を次に示す。

1. 適当な値 t を定め

$$s^* = \arg \min_s d_x(s, t)$$

を求める。ここで、 s^* はモデルへの帰属度が最大となる観測画像の点である。

2. 得られた s^* を用いて

$$t^* = \arg \min_t d_y(s, t)$$

を求める。 t^* も s^* 同様である。

3. 以上により求められた s^* 、 t^* より最適な座標 $x(s^*, t^*)$ が定まる。

このアルゴリズムにより位置合わせした例を図6に示す。

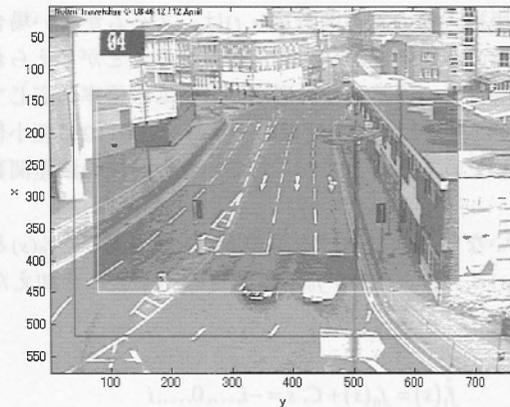


図6: 位置合わせの結果

画像の中の比較的大きな黒っぽい長方形枠で囲まれた部分が位置合わせした部分である。マッチした部分にモデル画像に含まれる画像の1枚を貼り付けている。内側にある、白っぽい枠で囲んである部分は、この固定型PTZカメラがデフォルト状態の時に赤色枠の部分で写される場所を表している。車が多数写っているにもかかわらず、うまくマッチしていることがわかる。

図7、8はx軸方向、y軸方向の単一方向についての誤差関数である。PTの移動量を推定する場合において、参照する画像の大きさは縦(x方向)300ピクセル、横(y方向)600ピクセルのモノクロ画像であり、探索領域は観測画像の左上の点 $(s, t) = (0, 0)$ として、 $40 \leq s \leq 220$ 、 $40 \leq t \leq 168$ とした。168という半端な値は、画像の右端の位置から出てくるものである。

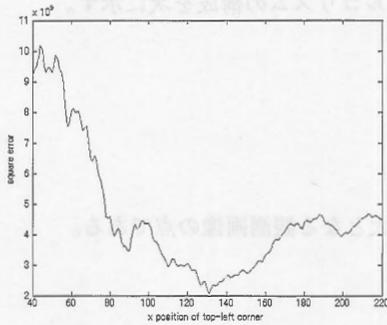


図7: x軸方向の誤差関数

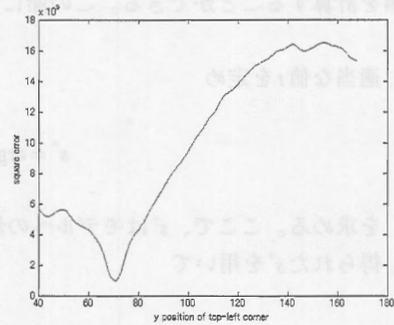


図8: y軸方向の誤差関数

2.5 ズームの検出

前節でPTによる移動量を検出する際に求めた、特徴空間との距離 $d_x(s, t)$ からズームの検出を行う。最適となる座標近傍の誤差関数 $d_x(s, t)$ は、ズームがない場合、最適解付近の誤差の変化の仕方は、画像の非周期性から、形は類似することが考えられる。一方、ズームしている画像の、最適解付近の誤差関数の形は、ズームの倍率に応じて、凹の形が広がることが考えられる。図9に探索した誤差関数のグラフを、図10に極小値付近の拡大図を示す。これより、ズームしている場合とズームしていない画像の誤差関数のグラフの形の違いが読み取れる。

そこで、ズームしていない画像における誤差関数を1つ選び、 $f_0(s)$ とおく。テスト画像における誤差関数を $f_1(s)$ とおき、 $f_1(s)$ が $f_0(s)$ に固定のバイアスを加えたものであると仮定すると、 $f_1(s)$ の推定値 $\hat{f}_1(s)$ は

$$\hat{f}_1(s) = f_0(s) + C, s = -i, \dots, 0, \dots, i \quad (14)$$

と書ける。ここに、位置 s は、極小値の位置が同じ0になるよう、あらかじめ調整されているものとする。このとき、マッチング領域 $s = -i, \dots, 0, \dots, i$ における C の最小2乗推定値は

$$\hat{C} = \frac{1}{2i+1} \sum_{s=-i}^i (f_1(s) - f_0(s)) \quad (15)$$

である。

マッチング誤差によりあらかじめ決めておく閾値との2乗誤差の大小によりズームの有無を決定する。

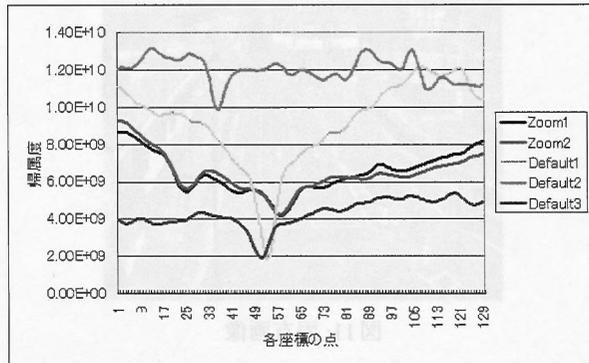


図9: 誤差関数

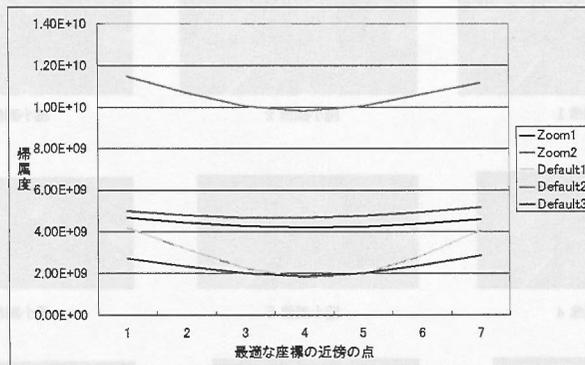


図10: 極小値付近の拡大

2.6 ズーム倍率の推定

ここでは、前節でズームと判断された画像のズーム倍率を推定する。ズーム倍率を推定する場合も2.4節で提案した、画像を縦横に和をとった2次元ベクトルを用いてマッチングを行う。まず、あらかじめズームした部分を含む位置合わせされたズームなしの画像複数枚から、それらの固有画像を作成する。固有画像とは、画像の共分散行列の大きな固有値に対応した数個の固有ベクトルを、画像の形に並べなおしたものである。第一主成分で作成した固有画像を図11に示す。

テスト画像については、補間を用いて各種倍率の縮小画像を生成する(図12参照)。この例では元画像を0.25倍から0.005倍間隔で0.20倍まで縮小している。

縮小した画像のx軸方向、y軸方向の和を計算し、それぞれの固有ベクトルを求め、2.4節同様、2次元ベクトルを用いてそれぞれの方向のみについてマッチングを行う。



図11: 固有画像

差異検出結果



縮小画像 1



縮小画像 2



縮小画像 3



縮小画像 4



縮小画像 5



縮小画像 6



縮小画像 7



縮小画像 8



縮小画像 9



縮小画像 10



縮小画像 11

図12: 縮小画像の例

このアルゴリズムにより、ズーム倍率を推定した結果を図13に示す。外側の大きな長方形領域が探索範囲であり、内側の長方形枠で囲まれた部分がズームされた範囲である。図13では、最適な縮小画像を固有画像に貼り付けている。ズームされている範囲の固有画像と縮小画像の差分による誤差が最小となるものをズーム倍率の推定値としており、ここでは、観測された画像はデフォルトの状態から4.5倍ズームされたものであるという結果を得たことになる。図14、15は最適な倍率での誤差関数をグラフで示している。

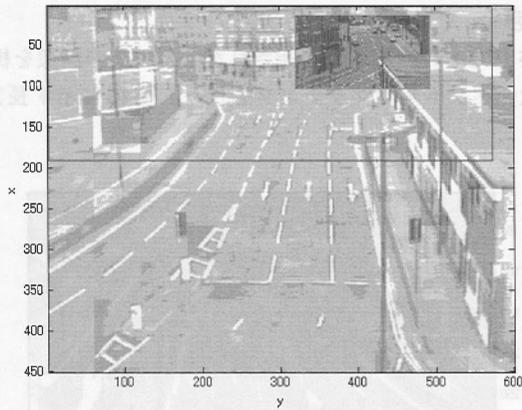


図13: ズーム倍率推定の結果

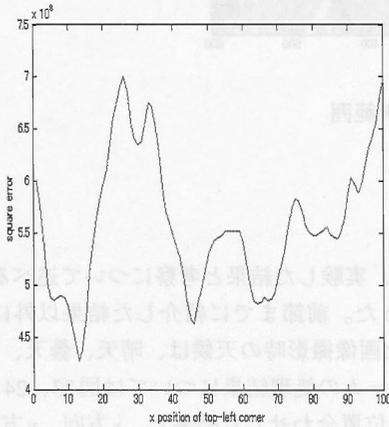


図14: x方向の誤差関数の変化

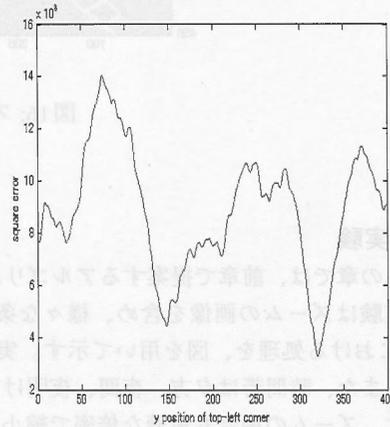


図15: y方向の誤差関数の変化

2.7 ズーム検出の際の窓の大きさ

ズームを検出を行うには、2.5節で述べたように、

1. 誤差関数の最小値を探索する
2. 最小値を与える点の近傍における誤差関数の形でズームの有無を判定する

の手順で行う。窓の大きさとしては、片側3点ずつの両側とした。

2.8 ズームの探索範囲

観測画像がズームされたものであると判断されたとき、その画像を様々な倍率で縮小し、縦方向、横方向の2次元ベクトルでマッチングを行う。図16に示す長方形枠が探索範囲である。

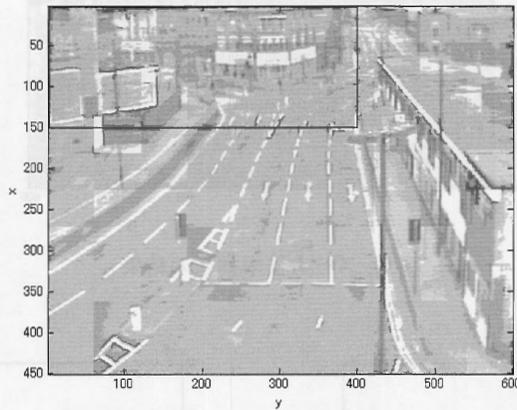


図16: ズームの探索範囲

3 実験

この章では、前章で提案するアルゴリズムを使い、実験した結果と考察について述べる。

実験はズームの画像を含め、様々な条件下で行った。前節までに紹介した結果以外に、PTにおける処理を、図を用いて示す。実験を行った画像撮影時の天候は、晴天、曇天、雨天、また、時間帯は夕方、夜間、夜明けである。ズームの処理結果については図23、24に示す。ズームの場合も最適な倍率で縮小した画像を位置合わせした結果と、 x 方向、 y 方向の帰属度の変化、平均化した差分をそれぞれ示す。なお、これらでは、いずれもモデルに含まれない画像で実験を行っている。

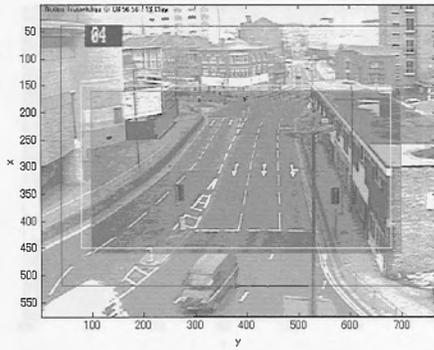


図 17: 晴天時の処理結果

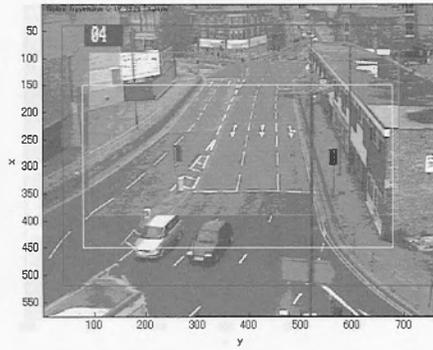


図 18: 曇天時の処理結果 1

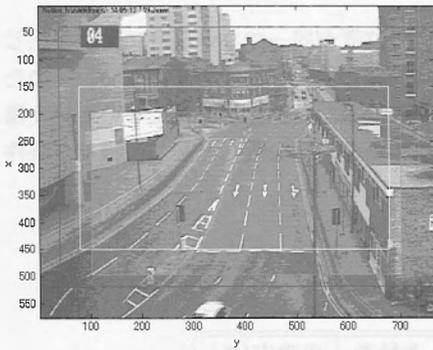


図 19: 曇天時の処理結果 2

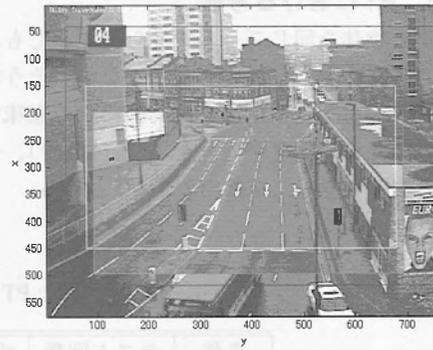


図 20: 雨天時の処理結果 1

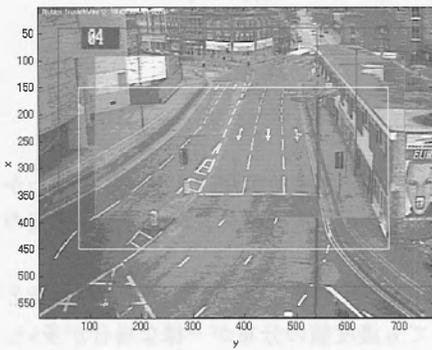


図 21: 雨天時の処理結果 2

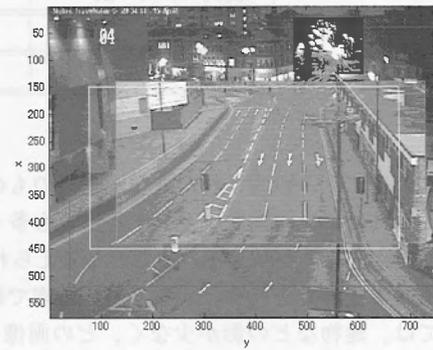


図 22: 夜明けの処理結果

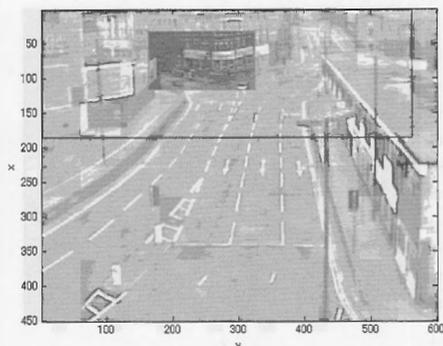


図23:ズームの処理結果1

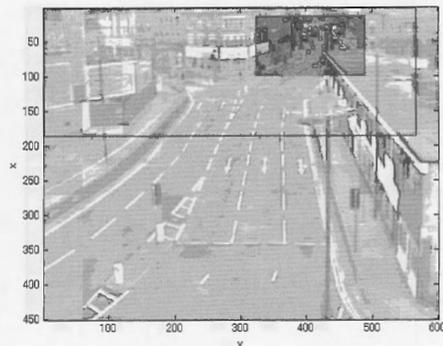


図24:ズームの処理結果2

3.1 PTにおける考察

同じ物体を同じズームで捕らえていても、画面上の位置が同一でない場合には歪が生じるため、完全に一致することはない。そういう意味では、完全一致はあり得ない。そのため、ターゲットの位置が、単純移動の意味で、これ以上適した位置が存在しないという場所を発見することができた場合、「成功」、そうでない場合を「失敗」として、様々な条件下で実験を行った結果を表2にまとめる。

表2: PTの推定結果

条件	テスト回数	成功数	失敗数	成功率(%)
晴天	101	95	6	94.1
曇天	55	50	5	90.9
雨天	25	22	3	88.0
夜間	20	14	6	70.0
その他	20	11	9	55.0
全体	221	192	29	86.9

この表より、観測された画像が晴天のものの位置合わせの成功率が最も高いことがわかる。これは晴天の画像が、モデルに最も多く含まれており、特徴ベクトル空間が、好天の画像をうまくカバーしているためと考えられる。

曇天の場合も約90%とかなり良い確率で結果を得ることができている。データの特徴としては、建物などの影が少なく、どの画像をとっても濃度値の分布が一様な場合が多い。位置合わせに失敗した画像の多くは、参照するモデルと同じ部分に、大型バスや、大型トラックで大幅に平均画素値が変わるような画像である。その例を図25、26に示す。

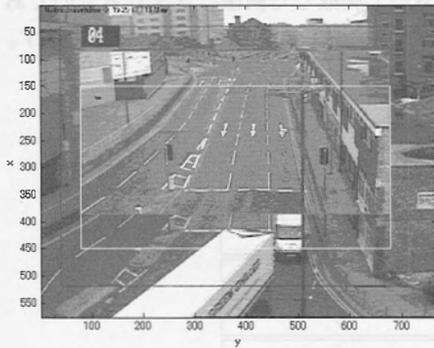


図25: 位置合わせに失敗した画像1

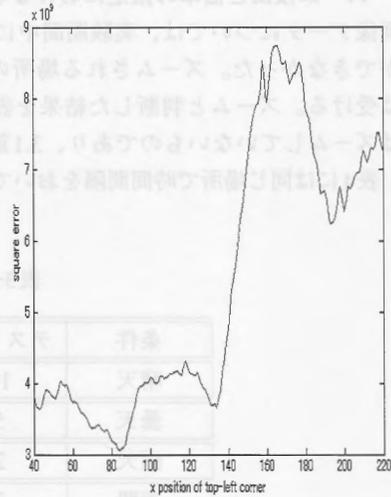


図26: 誤認識したときのx軸方向の帰属度の変化1

雨天の場合も晴天、曇天の場合同様に良い結果を得ることができた。位置合わせに失敗した画像は曇天同様に大型車の影響によるものである。

夜間の画像についてもある程度の対応はできているが、夜間は車のヘッドライトおよびその路面への反射の影響で、成功率は少し下がる。

「その他」の画像には、夕方、夜明けの画像が含まれている。これらの時間帯の画像はモデルに含めなかったが、50%以上の確率で位置合わせが可能であった。その結果を図28に示す。図27は夕方における元画像である。



図27: 夜明けの道路画像

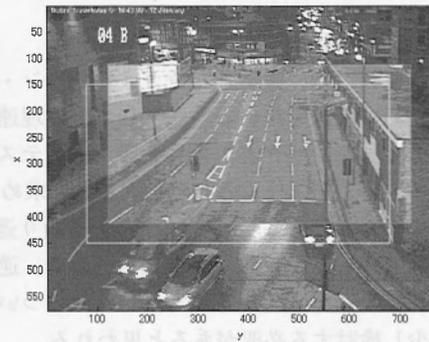


図28: 夜明けの画像の処理結果

3.2 ズーム検出と倍率の推定における考察

画像データについては、実験期間中にはほぼ決まった場所をズームしたものしか得ることができなかった。ズームされる場所のPTの変化はほとんどないが、車や天候などの影響は受ける。ズームと判断した結果を表3に示す。晴天、曇天、雨天、夜間、その他の画像はズームしていないものであり、3.1節の表2において失敗した画像に含まれるものである。表4には同じ場所で時間間隔をおいて得ることができたズーム画像の倍率の分布を示す。

表3: ズームの判定

条件	テスト回数	ズームの判断 (有)
晴天	101	2
曇天	55	0
雨天	25	3
夜間	20	2
その他	20	3
ズーム画像	21	19

表4: 倍率の分布

倍率	4.1	4.4	4.5	4.6	4.7	4.8	4.9
枚数	2	3	3	7	1	2	1

また、表4からは、同じ場所を一定時間写した画像からのおおよそのズーム倍率の推定を行うことができています。

最後に、対象としていない画像が観測されたときの結果について述べる。図29、30は対象としている部分が途中で切れている場合の、元画像、位置合わせ結果を示している。

4 おわりに

本研究では、固定されたカメラのパン・チルトの移動量、ズーム検出と倍率の推定を行うために画像を用いる方法として、処理速度を向上させるために、規準サンプル画像数十枚より、特徴ベクトル空間を構成し、テスト画像と特徴空間との距離を基にして、パン・チルトおよびズームの有無と大きさを求める手法を提案した。

ズームの倍率については、実験を繰り返してみると、表3より、ズームしていない場合でもズームと判断してしまうケースや、逆にズーム画像をテストしたとき、ズームでないと判断するケースがあった。これらについては、判定規準の閾値や、窓の大きさなどをもう少し検討する必要があると思われる。

実際への適用までには、まだ検討すべき問題も多数残っている。今後の課題としたい。



図29: 対象としていない画像

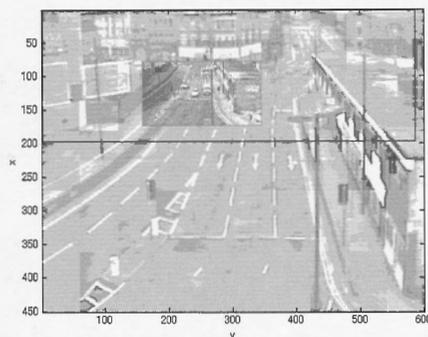


図30: 対象としていない画像の処理結果

謝辞

本研究を行うに際して、画像のご提供をいただいた、英国 Nottingham Traffic Centre に感謝致します。また、本研究の一部は、文部科学省のオープンリサーチセンター整備事業による私学助成を得て行われた。

参考文献

- [1] R. Cucchiara, and M. Piccardi, "Vehicle detection under day and night illumination," *Proc. of 3rd Int'l ICSC Symp. on Intelligent Industrial Automation*, pp. 618–623, 1999.
- [2] D. Koller *et al.*, "Towards robust automatic traffic scene analysis in real time," *Proc. Int'l Conf. Pattern Recognition*, pp. 126–131, 1994.
- [3] C. S. Regazzoni, G. Fabri and G. Vernazza (eds), *Advanced Video-Based Surveillance Systems*, Kluwer Academic Publishers, Boston, 1999.
- [4] P. Remagnino *et al.*, *Video-Based Surveillance Systems –Computer Vision and Distributed Proc.*, Kluwer Academic Publishers, Boston, 2002.
- [5] M. Tanaka, A. Bargiela, J. Coggan and S. Adachi, "Development of traffic image analysing system for mobile terminals using Internet," *The Fifth Asia-Pacific Industrial Engineering and Management Systems Conference (APIEM)*, pp. 17.7.1–17.7.12, 2004.
- [6] M. Tanaka, R. Hamamura and A. Bargiela, "Information extraction from traffic images," *The 35th International Symposium on Stochastic Systems, Theory and Its Applications*, pp. 41–46, 2004.
- [7] Y. Weiss, "Deriving intrinsic images from image sequences," *Proc. Int'l. Conf. Computer Vision*, pp. 68–75, 2001.
- [8] 島井博行ほか, ロバスト統計に基づいた適応的な背景推定法, 電子情報通信学会論文誌, J86-D-II, (6), pp. 796–806, 2003.