

確率的予測に基づく人流計測

新村貴彦[†] 新井啓之[†] 井上潮[†]

[†]NTT データ 技術開発本部

〒104-0033 東京都中央区新川 1-21-2 茅場町タワー

E-mail: shinmuratk@nttdata.co.jp

あらまし 斜め下向きに設置されたカメラ映像を利用した人流計測方法を提案する。本手法では、混雑時にも安定に動作するよう、オプティカルフローや背景差分を用いて実空間における人流の状態を計測する。斜め下向き映像では、3次元位置の不定性により、見かけの観測値を実空間での物理量に一意に対応付けることができない。そこで、実空間と画像空間でのサイズの比であるスケールファクタの出現頻度分布を用い、実空間での物理量を期待値として算出する。スケールファクタの出現頻度分布はCGシミュレーションにより事前に取得する。本稿では人流の速度計測を例に手法の提案を行うとともに基礎実験の結果を併せて示す。

Simulation-based Estimation of Human Traffic for Video Surveillance

Takahiko SHIMMURA, Hiroyuki ARAI and Ushio INOUE (NTT Data Corp.)

Abstract

In this paper, we propose a method for simulation-based estimation of people flow by a surveillance camera. To respond effectively to crowded indoor scene, we calculate actual motion vector with optical flow and a background scene model. Because you cannot specify the location of objects by cross shots, it is not possible to relate a motion vector of video sequences to that of real space. To overcome this problem, we infer the motion vector of real space by scaling that of video sequences. For that purpose, we employ “scaling factor” which is made from probability distribution of camera-distance on CG scenes. We examine effectiveness of our method by video sequences of a station surveillance camera.

1. まえがき

近年、防犯、防災意識の高まりに伴い、監視カメラの利用が急速に進んできている。施設や街頭における不審者の監視、ゲートの開閉管理、人の流れ（以下、人流と呼ぶ）の観測によるイベント会場などでの安全管理など、その用途は多岐に渡っている。通常、このような監視業務は人の目に頼っているが、長時間にわたりモニタを注視し続けることは大きな作業負担となるだけでなく、集中力の限界から見落としが多くなるという問題がある。このため、これらの監視業務を支援する技術への期待は大きい。

人が多く集まる場所においては、人流の状態を把握し、状況に応じて適切な誘導を行うことは安全管理上極めて重要である。この際、1) 集団としての移動速度、2) 集団を構成する人数（または人の密度）が人流を把握するための重要な指標となる。これら

の指標を自動的かつ定量的に計測できれば、人の殺到など局所的な状況を客観的にモニタリングできるとともに、多地点での計測結果から広域に渡る人流状況を推定することも可能となり、誘導や施設計画に役立てることができる。

移動速度や人数を計測する方法としては、頭上から撮影したカメラ映像から頭部を検出する方法[1]が提案されている。頭上から真下向きに歩行者を撮影した場合には、個人個人が良く分離された状態で映っているため、画像処理により個人を特定することは比較的容易である。このように頭上から真下向きにカメラやラインセンサなどを設置し、画像処理、信号処理により歩行者をカウントする技術は既に実用レベルに達していると言える。しかし、新たなセンサを設置するコストが問題となることも多く、既設の監視カメラを活用したいというニーズは根強い。通常の

監視カメラは、広い範囲を監視できるよう上方から斜め下向きを撮影するように設置されているため、上記技術をそのまま適用することはできない。

斜め下向きを撮影したカメラ映像を用い、人の移動速度や人数を計測する方法としては、まずトラッキング(追跡処理)の適用が考えられる[2,3]。しかし、斜め下向きに雑踏を撮影した場合には、画面上で人同士の重畳、隠蔽(オクルージョン)が頻繁に発生するためトラッキングは非常に困難となる。一方、人の集団をマクロにとらえその状態を計測する方法も考えられる。関連研究として、歩行者の空間的広がり様子をエントロピーで推定する研究[4]があるが、速度や人数を計測するものではない。集団としての移動速度は、画像上の見かけの速度場であるオプティカルフローを算出することにより、また、人数(または空間密度)は背景差分の面積を算出することにより大凡把握できる。いずれの処理も雑踏を対象としても比較的安定して動作することが期待される。

ところが、画像上での速度や背景差分面積といった空間スケールに依存する「見かけの観測値」は、「実空間の物理量」を直接表すものではないため、画像と実空間の関係に基づき、スケールを補正する必要がある。ところが、斜め下向きで撮影した場合、ある画素には手前の人の頭が映るかもしれないし、奥の人の足が映るかもしれない。このため、ある瞬間、ある画素での見かけの速度や面積を実空間での速度や面積に一意に対応付けることはできない。

そこで本研究では、このような見かけの観測値と実空間での物理量との関係を確率的に捉え、斜め下向きのカメラ映像での見かけの速度場および背景差分から実空間での移動速度および人数の期待値を算出する人流計測方法を提案する。移動速度や人数が短時間で急激に変化しない場合であれば、算出された期待値は実用上意味のある定量指標となる。

本手法では、画素毎に、そこに現れる人物の実空間と画像空間でのサイズの比(以下、スケールファクタ)の出現頻度分布を事前に求めておき、これを利用して速度や人数の期待値を算出する。ここで、スケールファクタの出現頻度分布はCG(Computer Graphics)を用いたシミュレーションにより事前に取得しておく。なお、人の混雑度合いによってオクルージョンの状況が変化すること、つまりスケールファクタの出現頻度が影響を受けることを考慮し、スケールファクタの出現頻度は人の混雑度毎に定義される。本稿では雑踏の移動速度を例にとり手法の提案を行うとともに、基礎実験の結果を示し、その有効性、課題

について考察を行う。

以下、2章では画像空間と実空間の対応付けの問題を説明し、3章では提案手法の概要を示し、4章ではCGにより生成したスケールファクタの出現頻度分布について検証を行い、5章で実映像による提案手法の評価を行う。6章でまとめと今後の課題の整理を行う。

2. 画像空間と実空間の対応付けの問題

斜め下向きのカメラ映像において観測される見かけの速度や面積といった空間スケールに依存する観測値を実空間での物理量に対応させるためには、以下の3つの効果を考慮する必要がある。ここでは画像上での見かけの速度と実空間での速度の関係を例に説明する。

- (1)手前と奥とで画像上での見かけサイズが異なり、これに応じて速度も異なって観測される(図1)。
- (2)画像上で観測された速度が実空間のどこで発生したかを特定できない(図2)。
- (3)混雑時ほど手前の人奥の人が隠蔽する状況(オクルージョン)が頻発し、観測される見かけの速度に偏りが生じる(図3)。

これらの要因により、ある瞬間、ある画素で観測された見かけの速度を実空間の移動速度に一意に対応させることはできない。

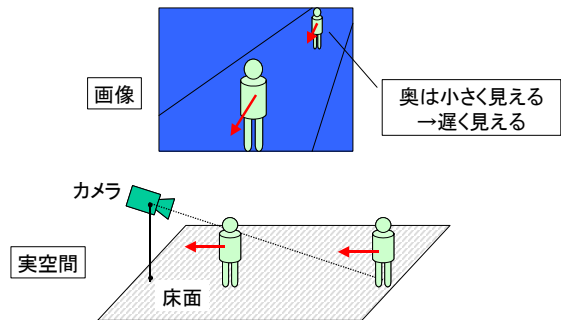


図1 奥行き方向でのスケール変化

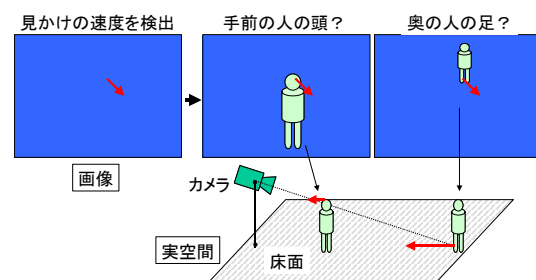


図2 実空間位置の不確定性

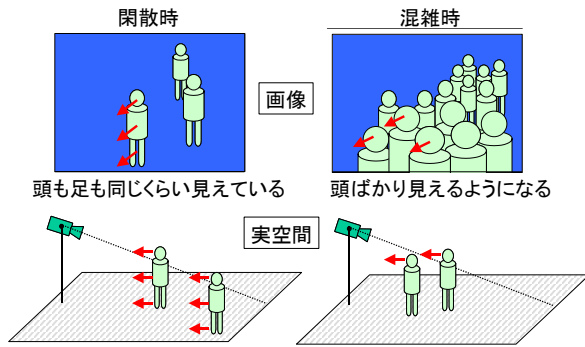


図3 オクルージョンによる観測値の偏り

3. 見かけの観測値からの人流計測方法

3.1 提案手法の概要

上述した画像空間と実空間の対応付けの問題を考慮し、本手法では、見かけの観測値と実空間での物理量との関係を確率的に捉え、斜め下向きのカメラ映像での見かけの速度場および背景差分から実空間での移動速度および人数の期待値を算出する人流計測方法を提案する。当然のことながら、ある瞬間について見れば実空間の物理量と算出された期待値とは必ずしも一致しないが、一定時間の平均として見た場合に、両者が一致していくことが期待される。計測しようとしている物理量が時間方向に急激に変化するものでなければ、算出された期待値は定量的に意味のある指標となる。

以下、本稿では速度計測の場合に絞って説明していく。人数についても同様の考え方により期待値を算出することができるが、ここでは、その説明は行わない。

図4に速度計測のおおまかな処理の流れを示す。スケールファクタの出現頻度分布データを事前に生成するフェーズと、実映像からの速度計測を行うフェーズに大きく別けることができる。以下、3.2～3.5において、本手法のポイントとなる考え方と処理の内容を示す。

3.2 スケールファクタ出現頻度分布と物理量の関係

まず、本手法のアイデアの基本となるスケールファクタについて説明する。実空間に人物とカメラまでの撮影距離 d 、正規化定数 K とすると、

$$R = K / d \quad (1)$$

の R をスケールファクタとして定義する(図5)。スケール

ールファクタは画像上での見かけの長さを実空間での長さに変換する変換係数となる。ここで、画像上で見かけの速度を、実空間での速度に変換する場合を考える。簡単のため、カメラの正面に向かって人が歩いて来る例を考える(図6)。

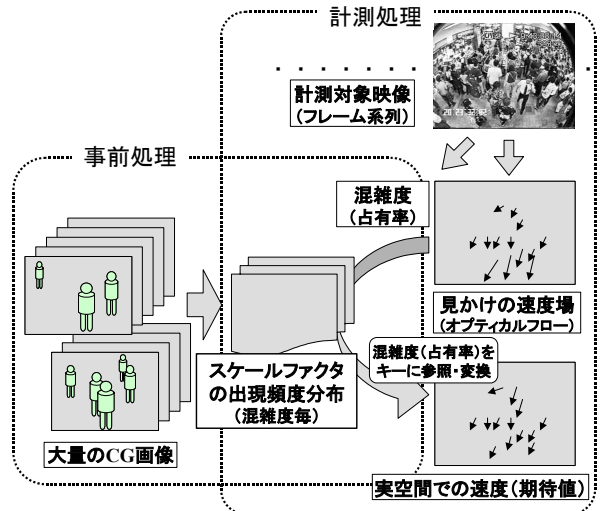


図4 速度計測のおおまかな流れ

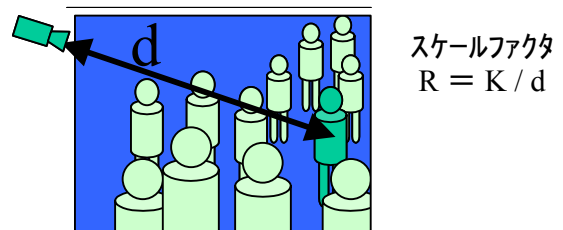


図5 スケールファクタ

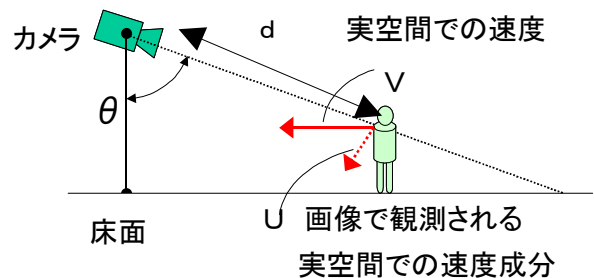


図6 カメラに向かって歩く場合の例

図5の人物の撮影距離を d とした場合、実空間での移動速度 V と画像上のある画素での見かけの速度 v とは、次式

$$U = v \cdot (K / d) = v \cdot R \quad (2)$$

$$U = V \cdot \cos \theta \quad (3)$$

を満たす。ここで、式(3)の U は実空間の移動速度 V のうち画像の中心軸に垂直な成分である。また、 R ($=K/d$) はスケールファクタである。これらの関係から、画像上で見かけの速度を発生した人物のスケールファクタがわかったとすると、実空間での移動速度が決まることになる。しかしながら、雑踏を対象とするような場合に、画像からスケールファクタを決めることは極めて困難である。

そこで、スケールファクタの出現頻度を考える。画素 (x,y) におけるスケールファクタ R の出現頻度分布を $P_{x,y}(R)$ が既知である場合、次式、

$$\begin{aligned} U_e &= \sum_R \{P_{x,y}(R) \cdot v \cdot R\} \\ &= W_{x,y} \cdot v \\ \text{ただし、} W_{x,y} &= \sum_R \{P_{x,y}(R) \cdot R\} \\ \sum_R \{P_{x,y}(R)\} &= 1 \quad (4) \end{aligned}$$

により、速度 U の期待値 U_e を求めることができる。ここで W_{xy} はスケールファクタ R の期待値である。

3.3 スケールファクタ出現頻度分布と混雑度の関係

2章で説明したように、人が多い状態と少ない状態とでは、オクルージョンの様子が異なる。例えば人が少ない場合には、ある画素 (x,y) では頭が見えることも、胴が見えることも、足が見えることもある。逆に人が密集している場合には、極端に言うと頭ばかりが見える状況となる。このことからわかるように、スケールファクタの頻度分布は人の混雑状況によって変化する。閑散としている状態から混雑している場合まで、定量的に一貫性のある物理量(速度など)を得るためには、この効果を考慮する必要がある。

本手法では、画像の背景差分の密度(以下、見かけの占有率と呼ぶ)をインデクスとして、スケールファクタの頻度分布を定義し、 $P_{x,y,\rho}(R)$ と表記することとする。ここで、 ρ は見かけの占有率であり、人物が占める画像中の面積の割合、

$$\rho = S1/S \quad (5)$$

として定義する。ここで S は画像全体の面積、 $S1$ は人物が存在する面積である(図7参照)。これにより、(4)式は、

$$\begin{aligned} U_e &= \sum_R \{P_{x,y,\rho}(R) \cdot v \cdot R\} \\ &= W_{x,y,\rho} \cdot v \\ \text{ただし、} W_{x,y,\rho} &= \sum_R \{P_{x,y,\rho}(R) \cdot R\} \\ \sum_R \{P_{x,y,\rho}(R)\} &= 1 \quad (6) \end{aligned}$$

と書き換えられる。

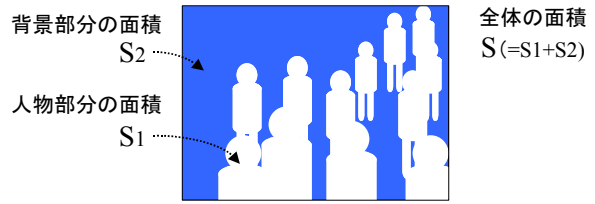


図7 見かけの占有率

3.4 CG シミュレーションによるスケールファクタの出現頻度分布の算出

スケールファクタの出現頻度分布、 $P_{x,y,\rho}(R)$ を事前取得しておく必要がある。今回は、CG (ComputerGraphics) によるシミュレーションにより $P_{x,y,\rho}(R)$ を算出する。

まず、カメラモデルやカメラ配置に関する幾何パラメータを指定し、実際のカメラ設置と同じような見えとなるようにCG空間を定義する。そして配置する人数を変えながらランダムな床位置に人物を配置し、様々な見えの擬似画像を生成していく。このとき、生成された1枚の擬似画像において、見かけの占有率 ρ を算出し、その ρ 毎に定義されたスケールファクタの出現頻度データに、擬似画像において最前面に見えている人のスケールファクタを追加していく。なお、ここで ρ の値は適当に離散化しておくものとする。これにより、 $P_{x,y,\rho}(R)$ を得ることができる。なお、人物を配置する際に、柱などの障害物や人の可動範囲などをの制約条件や人の行動モデルなどを反映させることにより、より、実体に合ったスケールファクタを算出することも可能である。

3.5 実映像における速度の算出

実映像から人の移動速度の期待値を算出する場合、各時点でのオプティカルフロー(見かけの速度場)と背景差分を常時算出しておく。ここでは、背景差分は、何らかの手段により作成した背景画像(人が映っていない画像)と現時点でのフレーム画像との差分の絶対値を取り、その値を予め設定しておいた閾値で2値化したものを指す。この処理により、図7で示した見かけの占有率を処理対象の画像に対して算出することができる。そして、算出した見かけの占有率に対応するスケールファクタの出現頻度分布 $P_{x,y,\rho}(R)$ を用いて、(6)式に従い実空間における速度の期待値を算出する(図4参照)。なお(6)式で示したように、速度の期待値の算出の際には、スケ

ールファクタ R の各画素での期待値 $W_{x,y,\rho}$ を予め計算しておくことにより毎フレーム毎の計算を簡単にできる。

4. CG によるスケールファクタの例

CG により取得したスケールファクタの例を示す。図8は生成したCGによる擬似画像の1つである。今回は、人の形状は簡単な円筒で近似している。図9は画像の中央付近のスケールファクタの様子であり、ここでは、見かけの占有率を増加させた場合の円柱重心までの撮影距離 d の平均値を近似曲線で示す。これからわかるように見かけの占有率が増えるほど、カメラに近い人の頭部で後ろの人を隠すため、撮影距離が短いほうへシフトしている。これは、人が増えることにより人の頭ばかりが見えるようになることに対応している。

また、図 10 に画像全体について算出した重み係数のスケールファクタ R の期待値 $W_{x,y,\rho}$ の例を示す。ここでカメラは x 軸側にある。この例から遠方ほどスケールファクタが大きくなっていることがわかる。これは、遠近による拡大縮小の効果が反映されていることによる。

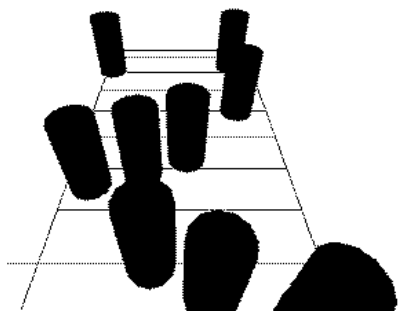


図 8 CG により生成した擬似画像 (640x480) の例

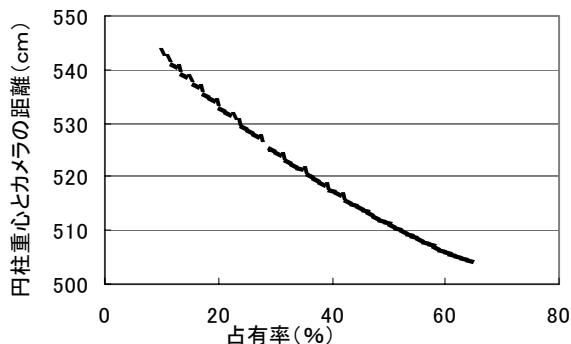


図 9 見かけの占有率と画図中央付近のスケールファクタ平均値

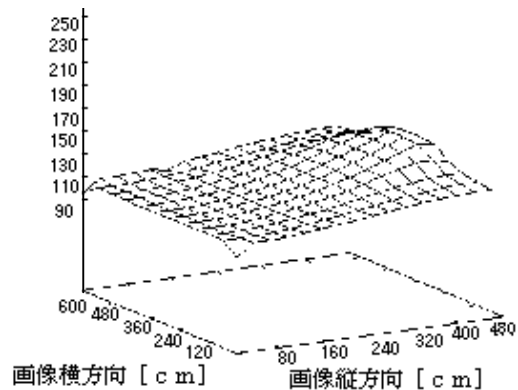


図 10 画像全体について算出したスケールファクタ R の期待値 $W_{x,y,\rho}$ の例 ($\rho=60\%$)

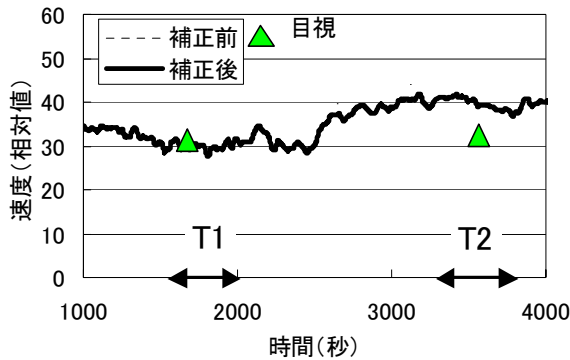
5. 実験結果2: 実映像での速度計測への適用

提案手法の有効性を検証するため、実映像を用いた速度計測の実験を行った。実験に用いたのは図12に示したカメラ映像であり、処理の対象範囲は図12の矩形内である。図4に示した手順に従い、画素毎に速度の期待値を求め、それを矩形内で平均した値、つまり集団としての移動速度について評価を行った。ここでは画像奥(上)方向への速度を計測した。図11は、矩形内での

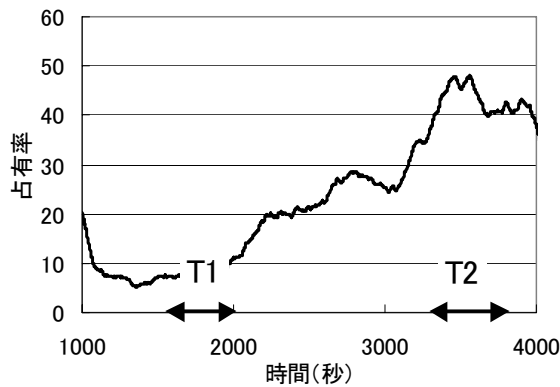
- ・ オプティカルフローを単純平均したもの(破線)
- ・ 本手法を適用して算出した移動速度(実線)
- ・ 目視により算出した移動速度(Δ)

である。ここでは、比較を容易にするために時間帯 T1 で規格化する形でグラフのスケールを調整している。なお、平均化の際にはフローの値が0の画素は処理の対象外としている。つまり人がいない背景部や止まっている人は計算の対象外としている。

図 11 のオプティカルフローの平均値(補正前:破線)を見ると、人が比較的まばらな時刻 T1 付近と比較し、混雑状態にある時刻 T2 付近では大きな値となっている。しかし、目視で計測した速度は T1, T2 でほぼ同じ値であることから、見かけのフローの偏りにより、混雑時のフローの値が相対的に大きく算出されていることが伺える。一方、本手法を用いたケース(補正後:実線)では、補正前に比べ目視に近い結果となっている。期待していたようにオクルージョンの効果を反映した補正が行われていると思われるが、目視結果との乖離が残っている。これは、実映像において背景差分処理により算出した見かけの占有率が実際よりも小さい値になっており、占有率の小さい場合のスケールファクタの期待値を用いて補正を行っていることによると考えられる。



(a)速度の相対値 (時間帯 T1 で規格化)



(b)映像から算出した見かけの占有率

図 11 速度の計測結果



図 12 混雑時 (時間帯 T2) の様子 (矩形内が処理範囲)

まとめ

斜め下向きのカメラ映像での見かけの速度場および背景差分から実空間での移動速度の期待値を算出する人流計測方法を提案した. 本手法では, 画素毎に, そこに現れる人物の実空間と画像空間でのサイズの比として定義されるスケールファクタの出現頻度分布を事前に求めておき, これを利用して速度や人数の期待値を算出する. スケールファクタの出現頻度分布は CG (Computer Graphics) を用いたシミュレーションにより事前に取得することができる.

本手法により, 人がまばらな状態から, 雑踏となっているような状態まで一貫した人流状態の評価指標を算出できる可能性があり, 実用上の意義は大きいと思われる. なお, 今回は速度の計測を中心に説明したが, 画像上の差分面積から人数を推定する場合にも, 同様の考え方を拡張することができる.

また, 雑踏の移動速度の計測を例にとり, 基本的な評価実験を行い, 一定の効果を確認した. しかし, その定量的な評価は不十分であり, また, 様々な誤差要因についての検討がなされていない. 今後詳細な実験を行う必要がある.

謝辞: 映像データ取得において御支援いただきました東京急行電鉄株式会社, 並びに研究に関する御協力, アドバイスをいただいた株式会社東急総合研究所の皆様へ深く感謝致します.

参考文献

- [1] 寺田賢治, 山口順一, “カラーカメラを用いた人数計測,” 電気学会論文誌, Vol. 118-D, No. 3, 322-328頁, 1998年.
- [2] 上条俊介, 松下康之, 池内克史, 坂内正夫, “時空間Markov Random Fieldモデルによる隠れにロバストなトラッキングアルゴリズム,” 電子通信学会論文誌, Vol. J83-D-II, No. 12, pp. 2597-2609, 2000.
- [3] 内田恭輔, 三浦純, 白井良明, “雑踏における歩行者の追跡,” 電子情報通信学会技術研究報告書, PRMU99 - 209, Vol. 99, pp39 - 44, 2000.
- [4] 岩崎洋一郎, 福田靖久, “画像処理とエントロピーを用いた歩行者交通流評価方法,” 電子情報通信学会論文誌, Vol. J81-A No. 4, pp610 - 612, 1998.
- [5] 黄秉元, 高羽禎雄, “ITV 画像による人の流れの実時間計測,” 電子通信学会論文誌, Vol. J66-D, No. 8, pp. 917-924, 1983.