

画像事象認識アルゴリズムの階層化

時空間 MRF モデルとパターン認識技術の連携

上條 俊介†

† 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

E-mail: †kamijo@iis.u-tokyo.ac.jp

あらまし 時空間 MRF モデルは、動きベクトル、画素、領域情報の分布の画像の時空間における相関関係を参照して領域分割を行うものであり、物体の形状モデルを参照しない。したがって、その定義から、画像上で重なりながら動きを同じくする物体同士を領域分割することはできない。そこで、本論文では、物体の形状モデルを参照したアルゴリズムと時空間 MRF モデルとの協調により、これらの移動物体同士をも精度良く領域分割可能なアルゴリズムを提案する。その際、画像認識におけるアルゴリズムの階層構造を考察し、時空間 MRF モデルと物体形状を参照したアルゴリズムとの違いを議論する。最後に、画像上での車両同士の重なりが激しい、トンネル内の低面角画像を用いた実験を行い、提案手法の有効性を検証する。

キーワード 時空間 MRF モデル、アルゴリズム階層、領域分割、トラッキング

Heirarchy of Algorithm for Image Understanding

Collaboration between S-T MRF model and Pattern Recognition algorithms

Shunsuke KAMIJO†

† Institute of Industrial Science, The University of Tokyo 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505
Japan

E-mail: †kamijo@iis.u-tokyo.ac.jp

Abstract Spatio-Temporal MRF model segments the spatio-temporal images by referring to the distribution of motion vectors, intensities, and region information of the spatio-temporal image. And it does not refer to any shape models of objects. Therefore, the S-T MRF could not segment occluded objects that move in the similar motion vectors in the images. In this paper, we then propose an algorithm to segment such the objects by the collaboration of S-T MRF model and a pattern recognition algorithm referring to shape models of objects. In order to clarify philosophy of the above algorithm, we will discuss an idea of algorithm hierarchy for image understanding and the difference between the S-T MRF model and algorithms based on object shape models. Finally, experiments of vehicle segmentation by using the low angle images of heavy occlusions in the tunnel in order to evaluate effectiveness of the proposed algorithm.

Key words Spatio-Temporal MRF model, Algorithm Hierarchy, Segmentation, Tracking

1. はじめに

トラッキングアルゴリズムに関しては、言うまでも無く従来から様々な研究が行われている。特に、車両のトラッキングに関しては、広く様々な研究がなされてきた。例えば、H.Kolling と H.Nagel [1] の研究は、その成功例としてよく知られている。Peterfreund [3] 等は、'Snakes [4]' の手法をトラッキングに適用し、車両の輪郭線の抽出を行っている。Smith [5] と Grimson [6]

等は、オプティカルフローの統計解析やオプティカルフローのクラスタリングおよびベクトル量子化を用いている。しかし、これらの研究は、いずれもオクルージョンの問題を扱っていなかった。これに対し、Leuck [7] と Gardner [8] は、車両形状の 3 次元モデルを仮定してパターンマッチングによりオクルージョンの問題を扱うことを試みている。しかし、この方法は車両形状の 3 次元を必要とすることから、様々な種類の車両や移動物体が混在する画像には適用が難しい。また、Malik [9] 等は、オ

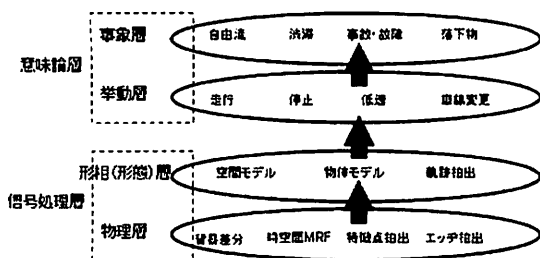


図1 アルゴリズム階層

クルージョンの問題を reasoning 手法により解決しようと試みている。これらの方法は、いずれも移動物体の個数密度が低い画像では良いトラッキング結果を示しているが、一般の交通映像のように移動物体の個数密度が高い混雑した画像でランダムに動く移動物体同士のオクルージョンの問題をロバストに扱うことは困難である。

これに対し、本論文の著者等が 2000 年に提唱した時空間 MRF モデル [14] [15] [16] は、時間軸方向のテクスチャおよび領域についての相関性に着目し、これを評価関数の中で定量化することで時空間画像の領域分割を行っている。これは、2次元すなわち空間軸方向のテクスチャ相関性はその存在が推定されるものの保証の限りでないのに対し、移動物体の時間軸方向へはテクスチャ相関性が必ず存在することが論拠となっている。

しかし、時空間 MRF モデルでは、車両や歩行者といった移動物体に関するモデルや学習データを用いていないことから、重なったまま同じ動きベクトルで併走する移動物体同士を分離しないといった技術的限界がある。本論文では、アルゴリズムの階層化を提案し、各階層に属するアルゴリズムの役割を論じる。さらに、時空間 MRF モデルと車両モデルによるパターン認識との協働により、オクルージョンにおける車両分離性能を向上させるための改良を行い、実験による評価を行う。

2. アルゴリズムの階層構造

本章では、交通事象認識アルゴリズムを一例にとり、その階層構造について著者の見解に基づいた定義を行い、各階層におけるアルゴリズムの役割について議論したい。したがって、必ずしも全てのアルゴリズムにそのまま当てはまるわけではないが、多くの場合を説明できるものと考えられる。

2.1 階層構造の概要

事象認識アルゴリズムの構造は、通常、大きく 2 階層に分けられる (図 1)。低次階層は信号処理層と呼ぶことができ、画像処理により移動物体 (交通画像では車両) の領域抽出・軌跡抽出・物体の分類等を行う。高次階層は意味論層と呼ぶことができ、信号処理層から抽出された移動物体の軌跡や物体分類情報を分析して事象認識を行う。さらに、信号処理層 (Signal Processing Layer) は物理層 (Physical Layer) と形相 (けいそう) 層 (Formal Layer) または形態層 (Morphological Layer) とし、意味論層 (Semantic Layer) は挙動層 (Behavior Layer) と事象層 (Event Layer) とに細分することができる。

物理層では、画像に内在する情報のみを用いた処理が行われる。即ち、車両や人の物体形状モデル、実世界・画像座標変換モデル、車線等の移動経路に関するモデル等というように、空間・物体の形相や形態に関する情報は参照されない。トラッキングの前処理としてよく使われる背景差分やエッジ抽出等の処理は、信号処理層の中の物理層に属する。これに対し、形相 (形態) 層では、物理層からの情報に対して上記の各種モデルを用いたパターン認識を行う。モデルは、学習の場合も非学習の場合もあるが、いずれにしても形相層において初めて物体や空間に関する知識が参照される。所謂トラッキングと呼ばれているのは、物理層と形相層を包含する場合と形相層のみを指す場合とがある。

意味論層の挙動層では、信号処理層から抽出した情報をもとに停止・低速・車線変更といった個々の車両の挙動を分類する。挙動層では、交通規則等の意味論的な知識が初めて参照され、形相層では「数値」としてしか抽出されていなかった移動軌跡に対して「意味 (semantics)」が付与される。さらに、事象層では、挙動層で抽出された移動物体の挙動の組み合わせにより、その事象が何であるかが判定される。即ち、検出事象の辞書としての役割を有する。例えば、ある車両が停止車両を検出した場合でも、画面全体で全ての車両が停止していれば渋滞と判定され、前方空間が大きく開いていたり他の車両が動いていれば事故または故障と判定されるのが妥当である。このことから、挙動層と事象層の役割の違いを理解することができる。

以上が、事象認識アルゴリズムの理想的な階層分類であるが、各階層が明確に分かれていないアルゴリズムもある。また、交通量や人の動線等の計測を目的とした画像解析においては、意味論層は計測規則のみを規定し、事実上意義を有しないに等しい。

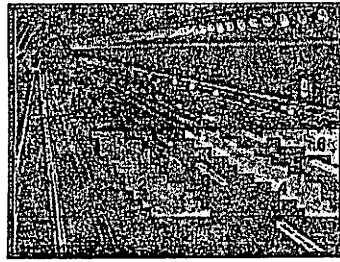
2.2 階層における時空間 MRF モデルの位置づけ

トラッキングアルゴリズムの議論において、物理層と形相層の違いが必ずしも常に明確に意識されているとは言えない。しかし、様々なトラッキング技術における時空間 MRF 技術の位置づけを理解するためには、両階層の区別を議論することが重要である。時空間 MRF モデルは、物理層で定義されている時空間画像の領域分割技術であり、空間モデルや物体モデルに関する知識は用いられない。時空間 MRF モデルからの出力は、あくまで領域分割結果と領域を構成する要素の動きベクトルの分布であり、これらの情報をもとに形相層において実世界軌跡の獲得、物体の分類等を行うことを前提としている。

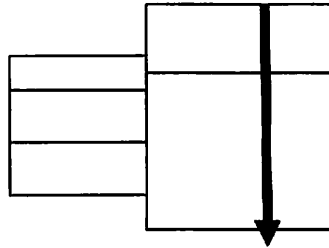
したがって、時空間 MRF モデルでは物体モデルに関する知識を用いないため、画像内へ進入してから退出するまで重なったまま並走する複数物体を分割することはできない。これは、時空間 MRF モデルに汎用性を持たせるため、意図的に物理層の範囲で定義されているからである。このため、重なったまま並走する複数車両は形相層で車両モデルを用いたパターン認識により分割することが要請される。

2.3 物理層と形相層の協働

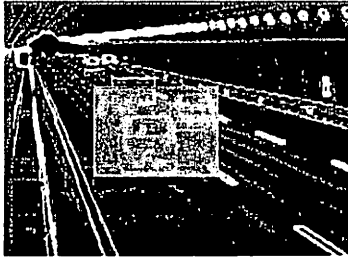
物理層のアルゴリズムのみでは分割できなかった複数移動物体を含む領域は、その上位に位置する形相層において車両モデ



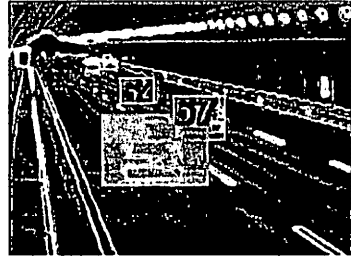
(a) 領域幅判定



(b) エッジパターン分類



(c) 時空間 MRF モデルにより複数車両が単一領域と判定された場合 (地点 1)



(d) エッジパターン分類による領域分割 (地点 1)



(e) 時空間 MRF モデルにより複数車両が単一領域と判定された場合 (地点 2)



(f) エッジパターン分類による領域分割 (地点 2)

図 3 エッジパターン分類による併走車両の領域分割

複数物体を含む領域は、上位の形相層において物体や空間に関するモデル（知識）を用いて領域分割を行うのが適切である。この場合、移動物体モデルとは車両の3次元形状や2次元画像への投影形状に関するモデルをいい、空間モデルとは、実空間座標と画像座標との変換や車線といった規則に関するモデルをいう。ここで、ある画像フレームにおいて形相層に属するアルゴリズムによる領域分割が行われた場合でも、時空間 MRF によって後続の画像フレームに対する領域分割が可能となる点に留意したい。このように、形相層のアルゴリズムを領域分割が高い信頼度で可能と判定される場合のみ適用し、一般には時空間 MRF を適用することで、高精度な領域分割が可能となる。このため、一般に、形相層のアルゴリズムのみでオクルージョンの問題を解決しようとする場合に比べて、形相層のアルゴリズムを単純化することが可能となる。

4.2 隣接車線を併走する車両の分離

まず、図 3(c) のように、時空間 MRF モデルからの領域分割結果が得られた場合、図 3(a) のように空間モデルとして定義

される車線情報と物体モデルとして定義される車幅情報を参照する。その結果、図 3(c) の場合には、単一の領域に複数車両が含まれている状態と判定されることになる。

車両の3次元モデルを用いるものとして、ステレオ画像を用いたアルゴリズム [13] があるが、ステレオ画像も画像座標から実世界座標への変換を行っているという意味では、座標モデルを用いた形相層のアルゴリズムという見方が可能である。しかし、本論文では、時空間 MRF モデルと組み合わせる目的から、単眼画像のパターン認識に基づくアルゴリズムを開発した。3(b) では、車両には水平エッジが多く見られることに着目し、車両の概形を水平エッジパターンの集合としてモデル化した。単一領域に複数車両が存在する場合には、図 3(b) の矢印のように、鉛直または車両の進行方向に沿ったスキャンライン上でのエッジパターンを計測し、スキャンラインを走査することによって、その変化点を検出することにより、車両間の領域分割が可能となる。

図 3(c)(d) に、地点 1 における画像に当該アルゴリズムを適

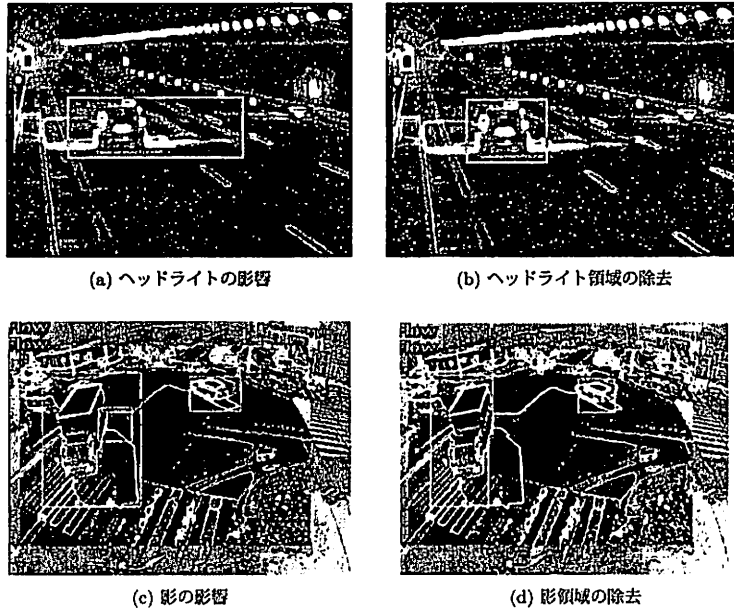


図4 エッジパターン分類によるヘッドライト・除去

用した領域分割の実験結果を示す。図3(c)において、時空間MRFモデルのみを用いた場合には2台の車両が単一領域として判定されているが、図3(a)(b)のような形相層におけるモデルおよびアルゴリズムを組み合わせることにより、図3(d)のような2台の車両の境界が正しく判定され、領域分割に成功している。図3(c)(d)の例では、画像中の第2車線を走行している車両が第1車線を走行している車両より画面奥に位置するため、図3(c)で抽出された領域の下辺の幅を判定するのみでは、1台の車両しか存在しないものと誤って判定される。そこで、図3(b)に対してオフセットを考慮した変形モデルを仮定することにより、図3(d)では単一領域に2台の車両が存在すると判定され、領域境界も正しく判定されている。

また、図3(e)(f)は、地点2における画像に当該アルゴリズムを適用した領域分割の実験結果を示している。地点2においては、地点1より車線幅が狭くカメラが画角が低い条件であるが、図3(f)のように精度良く分離することが可能となった。

4.3 ヘッドライトおよび影の除去

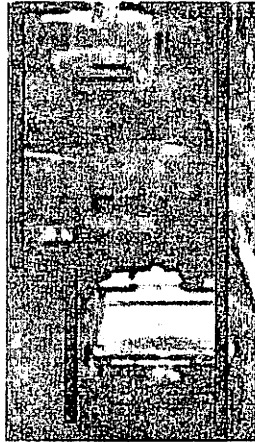
一方、物理層の処理のみでは、夜間のヘッドライトや影等の車両と異なる領域も車両と同一の領域として検出されてしまう。車両軌跡を正確に抽出するためには、これらの事物を車両から分離することが重要である。その際、当該単一領域に複数の車両が含まれている場合と、車両と影等が含まれている場合とを区別することが必要である。

先述と同様に、ステレオ画像を用いたアルゴリズムにより車両と影とを区別することが可能である[13]。しかし、ここでは時空間MRFモデルと組み合わせる目的から、先述のエッジパターン分類による領域分割手法を拡張することにより、ヘッドライトおよび影の領域を車両領域から分離することを考える。

図4(a)はヘッドライト領域を、図4(c)は影領域を、それぞれ車両と同一の領域として認識した例である。単一の領域に複数の車両またはヘッドライトや影が存在するか否かの判定は、先述と同様に車幅情報を参照して行われ、続いてエッジパターンの変化点の検出、領域分割が行われる。しかし、車両とヘッドライト・影とでは、それぞれにエッジパターンが異なる。車両においては、複数の水平エッジラインが領域全体に渡って分布するという特徴がある。ヘッドライトおよび影は領域の輪郭線にのみエッジラインが現れ、領域内にはエッジラインは現れない。これらの違いにより、車両として追跡対象とすべきかノイズとして除去すべきかの判定が可能となる。

著者は、過去の論文[16]において、画素値の空間微分を用いた照度変化にロバストな手法を提案した。この手法は、入力画像を隣接画素値との絶対差分和の関数として表される空間MRFエネルギーを画素値とする画像に変換する。これにより、画素の相対変化のみが抽出され、照度変化に対してロバストな画像処理が可能となり、ヘッドライトや影の除去に一定の成功を収めている。しかし、図4(a)(c)のように、ヘッドライトや影の領域境界のエッジが明確な場合には、領域として抽出してしまう。物理層のアルゴリズムの動作としては、このような結果は妥当であり、当該領域を除去するためには、形相層のアルゴリズムを導入することが必須である。

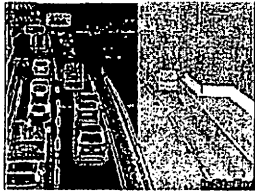
また、ヘッドライトや影の領域をノイズと考えて除去の対象とするか否かは、意味論層において判断される問題である。即ち、形相層においては、あくまでモデルに基づいて、分割・抽出された領域がどのような事物を表しているかが認識されるのみである。このように認識された事物に対してどのような意味を付与するかに関して、形相層のアルゴリズムが関与すること



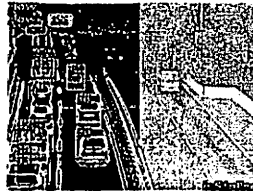
(a) 車幅および中心軸のずれ



(b) エッジパターン密度変化



(c) 時空間 MRF モデルにより複数車両が単一領域と判定された場合 (地点 3)



(d) エッジパターン分類による領域分割 (地点 3)

図 5 エッジパターン分類による追従車両の領域分割

は妥当ではない。

4.4 同一車線を追従走行する車両の分離

同一車線を追従走行する車両の分離を行うためには、図 5 にあるような 2 つの条件を定義することにより分離判定を行う。

まず、図 5(a) で表される条件は、大型車と小型車のように車幅が異なる車両同士が追従走行している場合に、その車線に沿った車両側面のエッジの位置が異なることに着目している。また、図 5(a) の条件は、同等の車幅を持つ車両同士が追従走行している場合であって、車両の中心軸位置が異なる場合についても適用することが可能である。

次に、図 5(b) で表される条件は、車線に垂直な車両のエッジ (車両の水平エッジ) が車線に沿った方向の密度分布を検証し、高密度にエッジが出現しているところを車両同士の境界として認識する方法を示している。しかし、当該条件は、あくまでも経験則として位置づけられ、論理的な根拠を主張することは困難である。

また、本手法は車両上部のエッジが画像中に見えていることが条件となるため、本手法の適応可能な画角に一定の限界があることを断っておく。

図 5(c)(d) は、地点 3 における画像に当該アルゴリズムを適用した領域分割の実験結果を示している。図 5(c) は、時空間 MRF モデルのみでは分離できなかった場合を示している。この画像は渋滞中の画像であり、車両同士の距離が短い状態で追従走行している。このため、第 1 走行車線を走行する 5 台の車

両が単一の領域として検出されている。これに対し、図 5(d) は、時空間 MRF モデルに図 5(b) で示したアルゴリズムを組み合わせて領域分割を行った結果である。画面の上から 2 台目と 3 台目の車両が単一領域のまま分割できなかったことを除けば分割できていることが分かる。

4.5 実験結果

実験は、3 つの地点に設置されている ITV カメラ画像を用いて行った。先述のようにまた、図 3(c)(d) は地点 1 における実験結果を、図 3(e)(f) は地点 2 における実験結果を、図 5(c)(d) は、地点 3 における実験結果を、それぞれ示している。

横方向分離に関する実験を、地点 1, 2 の画像を用いて行った。両画像の合計で約 18,000 台の通過車両について領域抽出を行った結果、1827 領域が複数車両が存在しているにもかかわらず、時空間 MRF によって単一領域として判定された。これらのうち、1372 領域がエッジパターン認識と時空間 MRF モデルの連携アルゴリズムによって正しく領域分割されたため、recall rate は $1372/1827 = 83.2\%$ であった。主な未分割の原因は、大型車両の陰に小型車両が隠されていて、画像上で小型車両が現れている領域面積が小さい場合等に領域分割がされていないことである。また、エッジパターン認識と時空間 MRF モデルの連携アルゴリズムが領域分割処理を行ったものが 1487 領域であったため、precision rate は $1372/1487 = 92.3\%$ であった。主な過分割の原因は、大型車両が 1 台である場合であっても、画角によって 2 台分が存在するものと誤認識される

場合があることが挙げられる。

次に、縦方向の分離に関する実験を、地点3の画像を用いて行った。約10,000台の通過車両について領域抽出を行った結果、418領域が複数車両が存在しているにもかかわらず、時空間MRFによって単一領域として判定された。これらのうち、314領域がエッジパターン認識と時空間MRFモデルの連携アルゴリズムによって正しく領域分割されたため、recall rateは $314/418 = 75.1\%$ であった。また、エッジパターン認識と時空間MRFモデルの連携アルゴリズムが領域分割処理を行ったものが390領域であったため、precision rateは $314/390 = 80.5\%$ であった。エッジパターン密度変化による分割は、経験則として提案されており、その有効性および改良方法については、今後継続して検討していきたい。

5. ま と め

本論文では、画像認識のアルゴリズムの階層を提案し、それぞれの階層の意義を明確に定義した。アルゴリズムの各階層には固有の役割があり、それぞれを明確に意識することで安定したアルゴリズムの開発が可能となり、かつ開発の負担が軽減されると考えられる。

時空間MRFは、最下層である物理層に位置づけられる技術であるが、上位の形相層における車両モデルに基づく簡易なアルゴリズムと協調させることにより、さらに高精度なトラッキングが可能となった。このように、時空間MRFモデルを用いて本来の物理層の役割を十分に果たすことにより、形相層のアルゴリズムを簡素化しつつ、アルゴリズム性能を向上させることが可能であることが確認された。

領域分割精度は、いずれも90%未満と低く留まったが、今後も改良を続けて行いたい。しかし、別の見方をすれば、地点1, 2, 3のいずれにおいても時空間MRFモデルによって85 - 95%の車両が既に分離されており、全通過車両に対する領域分割の精度という意味ではprecision rate, recall rateともに、90 - 97%程度を達成しているといえる。

謝 辞

本研究は、独立行政法人新エネルギー・産業技術総合開発機構の産業技術研究助成事業のテーマとして行われた。

文 献

- [1] H.Kolling and H.Nagel, "3D Pose Estimation by Directly Matching Polyhedral Models to Gray Value Gradients", IJCV, Vol.23, No.3, pp283-302, 1997.
- [2] H.Kolling, H.Nagel, and M.Otte, "Association of motion verbs with vehicle movements extracted from dense optical flow fields.", Proc. of ECCV'94, Vol.II, Lecture Notes in Computer Science 801(Springer-Verlag), pp388-347.
- [3] Natan Peterfreund, "Robust Tracking of Position and Velocity With Kalman Snakes", IEEE Trans. Pattern Analysis and Machine Intelligence(PAMI), Vol.21 No.6, 1999, pp.564-569.
- [4] M.Kass, A.Witkin, and D.Terzopoulos, "Snakes: Active contour models", Int'l J.Computer Vision, Vol.1, 1988, pp.321-331.
- [5] S.M.Smith and J.M.Brady, "ASSET-2: Real-Time Motion Segmentation and Shape Tracking", IEEE Trans. Pattern

Analysis and Machine Intelligence(PAMI), Vol.17 No.8, 1995, pp.814-820.

- [6] C. Stauffer and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking", Proc. of CVPR 1999, Vol.II, Jun 1999, pp246-252.
- [7] Holger Leuck and Hans-Hellmut Nagel, "Automatic Differentiation Facilitates OF-Integration into Steering-Angle-Based Road Vehicle Tracking", IEEE CVPR'99, pp.360-365.
- [8] Warren F.Gardner and Daryl T.Lawton "Interactive Model-Based Vehicle Tracking", IEEE Trans. PAMI, Vol.18 No.11, 1996, pp.1115-1121.
- [9] K.Weber, J.Malik, "Robust Multiple Car Tracking with Occlusion Reasoning", Proc. of ECCV'94, Vol.I, Lecture Notes in Computer Science 800(Springer-Verlag), pp189-196.
- [10] 羽下哲司, 鷲見和彦, "時間平均シルエットを用いた非剛体移動対象の追跡", 第9回画像センシングシンポジウム講演論文集, pp265-270.
- [11] S.Geman and D.Geman, "Stochastic Relaxation, Gibbs Distribution, and the Bayesian Restoration of images", IEEE trans. PAMI, Vol.6, No.6, pp721-741, 1984.
- [12] B.S.Majunath and R.Chellappa, "Unsupervised Texture Segmentation Using Markov Random Field Models", IEEE Trans. PAMI, vol.13, no.5, pp478-482, May 1991.
- [13] 阪坊 正樹, 東海 雅俊, 緒方 司郎, "環境や対象の変動に強い交通流計測のための画像センサ Silhouette Vision", 第5回画像センシングシンポジウム講演論文集, pp379-382, 1999年6月, 横浜.
- [14] 上條俊介, 松下康之, 池内克史, 坂内正夫, "時空間 Markov Random Filed 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", 電子情報通信学会論文誌 D-II, Vol.J83-D-II No.12, pp2597-2609, 2000年12月.
- [15] S.Kamijo, Y.Matsushita, K.Ikeuchi, M.Sakauchi, "Traffic Monitoring and Accident Detection at Intersections", IEEE trans. ITS, Vol.1 No.2, June. 2000, pp.108-118.
- [16] S.Kamijo, T.Nishida, M.Sakauchi, "Occlusion Robust and Illumination Invariant Vehicle Tracking for Acquiring Detailed Statistics from Traffic Images", to appear in IEICE TRANSACTIONS on Information and Systems, Vol.E85-D, No.11(2002).