

# 日立国際電気における人の動きの計測とその応用

上田 博唯

hiro-u@po.ijnet.or.jp

日立国際電気 放送・映像研究所

筆者のグループが日立国際電気(旧日立電子)で開発してきたシステムの、いくつかの実用例について紹介し、CVに技術よる身体動作の解析という研究開発分野における監視システム固有の実用上の問題点、筆者らが開発してきた具体的な解決方法、そして今後の課題などについて論じる。

## 1. はじめに

映像監視システムは、立ち入り禁止区域への侵入者、拳動不審者や故障車といった異常状態の発見や記録のために用いられ、犯罪や事故を未然に防止することに役立っている。しかし、ほとんどの場合、異常状態の発見などは人間の判断に頼っているため、膨大な人的労力を必要としている。特に大規模監視システムにおいては、監視用のカメラ・テレビモニタの数も増大しており、全てを常時見続けることによる疲労、疲れから来る見落としなども大きな問題となりつつある。これらの問題を画像認識技術を用いて(半)自動化しようという研究開発が活発化している<sup>1) 2)</sup>。

筆者らも、画像認識技術を応用した自動侵入物認識・追跡機能を有する監視システムの開発を行っており、本報告では、それらの中からいくつかの実用例について紹介し、CVに技術よる身体動作の解析という研究開発分野における監視システム固有の実用上の問題点、筆者らが開発してきた具体的な解決方法、そして今後の課題などについて論じる。

## 2. システムの概要

図1に筆者らが開発している監視システムの代表的構成の概要を表す。現場に設置された監視カメラの映像は、光ケーブルなどを介して監視場所に

転送される。監視カメラは電動雲台に取り付けられており、監視場所にあるパソコンからレンズの向き、ズーム倍率などを制御することができる。図を簡略化するために、現場側の複数のカメラの映像をまとめて伝送するための映像多重化装置などは省略して描いてあるが、8~16台の(時には100台を超える)カメラが接続される。監視場所ではスイッチャにより複数カメラのからの映像信号の選択/分配が行われ、画像処理部やVTR、監視モニタに配信されて、画像認識と記録が行われる。

筆者らの監視システムの開発においては、画像処理部にパソコンとそれに内蔵する画像認識ボード

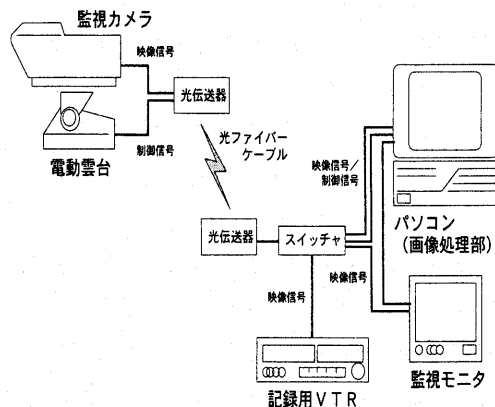


図1 代表的な監視システム構成の概要

(HITACHI IP-2000あるいはIP-5000)を主として用いている<sup>3)</sup>。画像処理部では、カメラからの映像を自動監視し、侵入物体が監視領域に侵入した場合、異常が検出された場所に対応するカメラと映像モニタを特定したうえで、監視員に警報を発して確認を促したり、侵入者に対して警告の放送(パソコンから制御される)を行う。あるいは後日の検索を容易にするためのマーキングを施して映像を記録するなどの処理を同時に実行する。

一方、よりインテリジェントな監視システムの実現をも目指しており、その開発には画像認識ボード内蔵型のカメラを利用している<sup>4)</sup>。このカメラの例を図2に示す。この画像認識ボード内蔵型のカメラは、上述のパソコン内蔵用の画像認識ボードのほぼ数分の一の能力を有し、かなり高度な画像認識処理が可能である。この能力により、カメラ自身が映像の認識結果によって雲台を制御して自律的に侵入者を追尾するというようなことも実現することができるようになった。さらには複数の画像認識ボード内蔵型カメラが互いに通信しあい、協調して侵入監視と追尾を実行することも可能となっている。また、従来は監視員が手動でカメラの雲台やズームレンズを制御して侵入物体を追跡/監視する必要があったが、最近の開発では、自律的にそれらの制御を行うことまでできるようにすることに取り組んでいる。

筆者らが基本とする認識方式の概略の流れを図3に示す。カメラ映像中の変化を検出し、その変化状態を分類する第一階層、検出状態変化を解析する第二階層、検出物体の軌跡や面積を総合評価する第三階層の3つの階層からなる<sup>5)</sup>。

階層が高くなるにつれ、取り扱うデータの量が減り、より抽象度の高い処理を行うことになる。このデータ量が減少した分、階層が上がるごとに、処理する時間幅をより広くしていくようにしてい

る。また各階層には、それぞれに物体の形状特徴や挙動に関する知識、監視領域に関する知識を付与して、分類や解析の精度を向上させている。このようにすることによって、映像中の物体の形状変化や複雑な運動に対処することができる。

以下の章では、具体的な事例に基づいて、ここで述べたコンセプトがどのように実用化されている

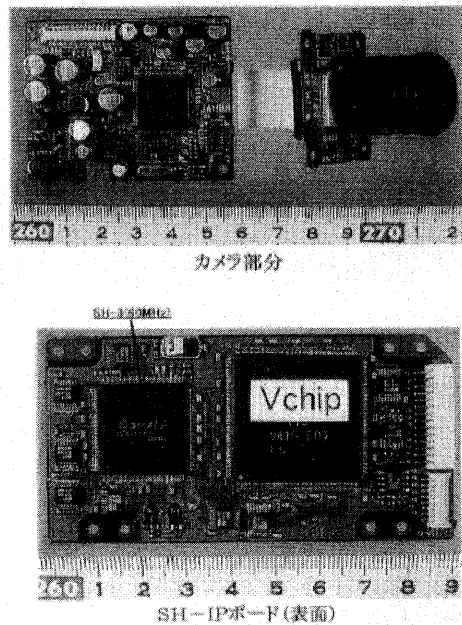


図2 画像処理ボード内蔵カメラ(分解時)

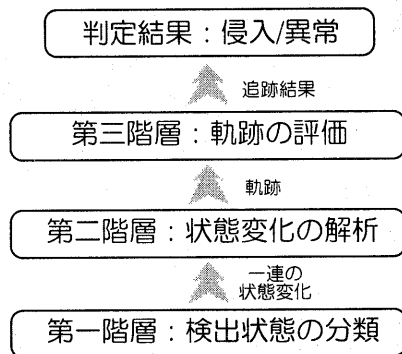


図3 基本とする認識処理の流れ

かについて説明する。

### 3. 構内監視の事例

#### (テンプレート対マッチング法)

監視エリア内に侵入する人物や自動車(以下、本稿においては人物も含めて、侵入物体と呼ぶことにする)を検知して、その移動の状況を認識する方法としては、従来から背景差分を連続的に行う方法<sup>6)</sup>やオプティカルフローを用いた方法<sup>7)</sup>などが提案されている。ところがオプティカルフローは、それ単独では、対象物の形状の判定が安定にできにくい。背景差分法には、天候などによる照明条件の変化に弱いという欠点があり、いかにして背景を更新するかという問題が生じる。

侵入物体が背景として登録されてしまうことを避けながら、この照明条件変化に追随させるためには、ある比較的短い一定時間の平均画像を背景として用いるのが有効である。しかし、このアイデアもカメラをパンしたりズームして追跡する場合に背景の更新が困難になるという問題がある。そこで筆者らは侵入物体が存在しない間は背景差分法を用い、侵入物体が検知されたら、パターンマッチング法(背景差分法での変化検知領域をテンプレートとして登録して用いる)に切り替えるというハイブリッド形式を採用している<sup>8)</sup>。

これに加えて、図4に示すような、テンプレート対(ついでにマッチングと呼ぶ方式を開発した。パター

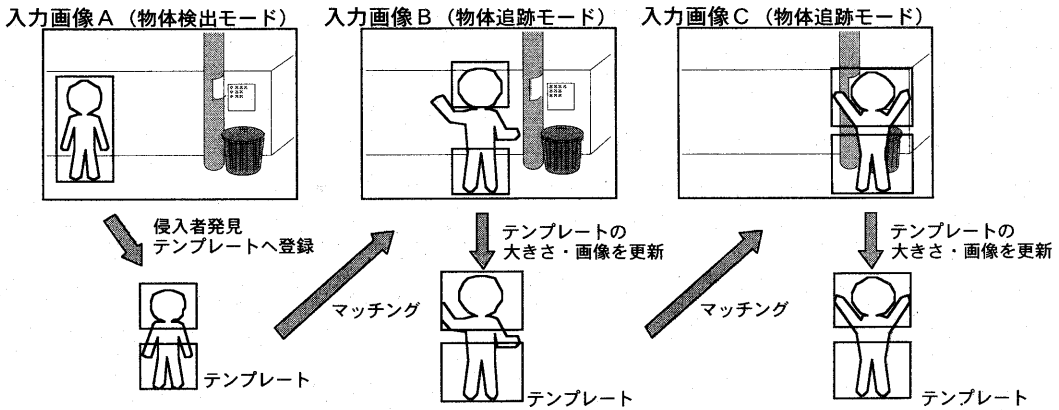


図4 テンプレート対マッチング法の原理



(a) 追跡開始時の入力画像

(b) 追跡開始4秒後の入力画像

(c) 追跡開始7秒後の入力画像

図5 構内監視の処理例

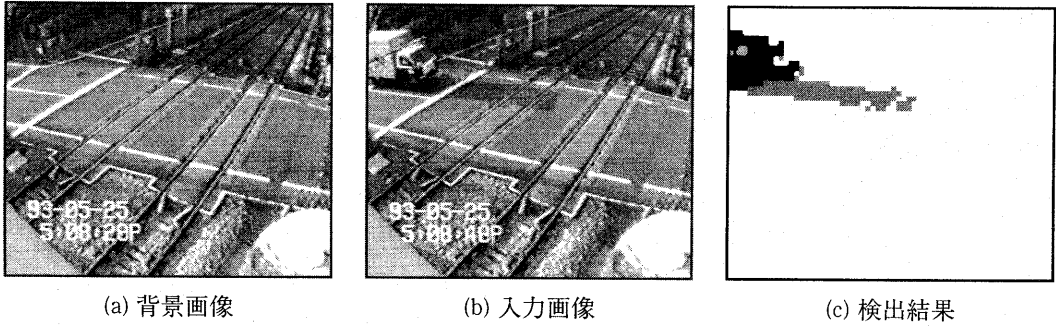


図6 踏み切り監視の処理例  
 ( ■■■ : 物体領域, ■■■■ : 光の照射部分と判定した領域, ■■■ : 影と判定した領域 )

ンマッチング法で侵入物体を追跡するためには、対象物体の姿勢の変化に追従させるために、テンプレートの更新を行う。このときカメラの光軸方向に運動する対象物体については、その見かけの大きさが変化するので、テンプレート更新の際に、その大きさ変動を吸収できるようにしたのである<sup>9)</sup>。

テンプレート対マッチング法ではカメラの視野内に侵入した物体の検出までを背景差分で行い、その後は物体の上下部分2ヶ所をテンプレートとして採取し、これらを独立にマッチングする。マッチング後の各テンプレート間の距離から対象物体

の見かけの大きさ変化量を計算し、これに基づいて、次のマッチングに利用する二つのテンプレートのサイズとそれぞれの探索領域を決定する。図5は実際に本システムを構内映像監視に適用した場合の処理結果である。対象物体の見かけの大きさ変化に追従して、安定な追跡が可能となっていることがわかる。

#### 4. 踏み切り監視の事例 (カテゴリ分離処理)

物体検出法として背景差分を用いる場合、雲の影や対象物体の影の部分の誤検出といった問題がある。図6に示したのは踏み切り監視の事例である。ここでは侵入物体の影やヘッドライトが照射された部分の存在により、実際には存在しないにも関わらず、踏み切り内に障害物があると誤認識するという問題があった。そこでカテゴリ分離と呼ぶ処理を開発した。図7の原理図のように、カテゴリ分離処理では、背景画像と入力画像の差分を計算する際、背景に対して明るく、もしくは暗くなったという符号情報と路面上の白線などの背景パターンが見えるか否かという指標(背景類似度)を利用することによって三つのカテゴリ(物体、影、ライトの反射)を区別することができる<sup>9)</sup>。このような処理を加えたことにより、図6に示した

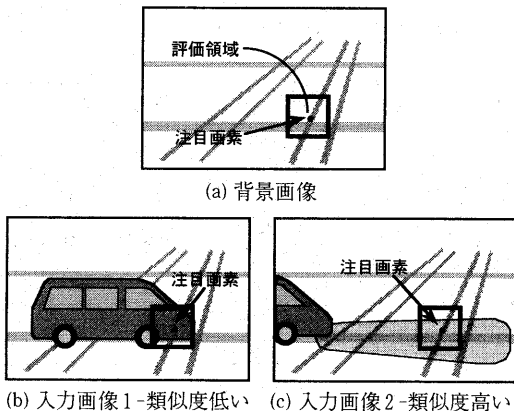


図7 カテゴリ分離処理の概要

ように、影やライトによる誤報を解決することができた。ただし、背景類似度を計算するためには、路面にかなりはっきりした模様が存在することが必要であり、また対象物体がある一定の大きさ以上に撮影されている必要もある。後者の条件については、後に述べる自動ズームアップなどと組み合わせることによって満足させることができるので、自動車のような大きなものだけでなく、人間と自動車が混在するような場合にも、この手法は有効である。

### 5. 挙動監視の事例 (三層構造の認識方式)

図8に示したのは侵入物体の挙動を監視する事例である。この事例を用いて、先に示した三層構造の

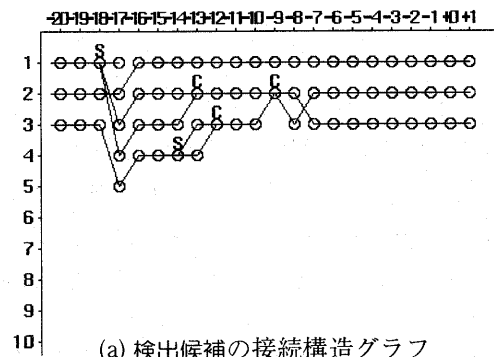


図8 路上歩行者の処理例

認識方式(図3)について詳しく説明する。三層構造の第一階層では、映像中の物体らしきものの検出状態を単純型、分離型、結合型、出現型、消滅型の5つに分類し、この結果を検出候補として第二階層と第三層に伝達する。第二階層では検出候補ごとに、これらの時間的変化を図8(a)に示すようにグラフ化して解析し第三層に伝達する。そして第三層ではこれらを総合的に判断して最終認識結果を出力する。先にも述べたように、階層が進むに従って、情報量は減少し、それに応じて取り扱う時間のスパンを広げていくようにしている<sup>5)</sup>。図8では、手前から向こうに歩いた二人の人物が、途中で(局所的な処理では)一つの物体のように見えるシーンを経過して再び二つの侵入物体として認識される様子、そして向こうから歩いてくる人物がこの二人とすれ違う様子が見られるが、これについても最終的に正しく解析されていることが分かる。

このように本方式は、時間的に解析範囲の異なる複数の階層を用いて物体の検出状態変化を解析することにより、従来のように一つの映像フレーム内の局所的な処理では判断できないような、物体の「すれ違い・追い越し」や「故障車・事故の検出」といった複雑な事象に対する認識性能の向上を図ることに成功している。

### 6. 親子カメラ協調型監視システム

図9に示すのは、開発中の親子カメラ協調型監視システムの実験風景である。差分法を用いて広域を監視する広域監視用カメラ(親カメラ)と、カメラ雲台を制御しながら視野内に侵入物体を捉えるように追尾する追尾監視用カメラ(子カメラ)という役割分担にしている<sup>10)</sup>。このような親子カメラの構成にすることで、親カメラによって検出した「侵入物体の移動経路」と子カメラ映像による「侵入物体の動作・表情」を同時に監視員に提

供することができる。ダムなどの広い範囲の監視を必要とする場所では、監視員が数多くあるカメラの雲台やズームレンズを操作して侵入物体を追跡・監視する必要があり、その操作が面倒なものであったが、提案する監視システムによれば、この監視員の負担を大幅に減少することができる。実験で使用している模型は、一周約 190cm のレーン上を 2 体の人形が約 30 s で周回するようになっている。この速度は、実寸に換算すると人間が早歩きするときの速さに相当する。画像処理サイズは親カメラが 512×440 画素であり、子カメラは 256×220 画素としている。各カメラの画像処理速度は、親カメラが約 300 ms/処理であり、子カメラ

は約 200 ms/処理である(親カメラは子カメラの 4 倍の画素数を処理しているが、画像認識処理ボードの性能が高いために、処理時間は 4 倍にはなっていないことに注意されたい)。

図10は、親カメラによる監視処理結果を表す画面の写真である。太い方の白枠で囲んだ侵入物体を子カメラで追尾していることを表している。白い破線の矩形は監視領域内で警戒すべき領域(第1次警戒領域)、実線は最も警戒すべき領域(第2次警戒領域)を表す。図11は、子カメラによる監視処理結果である。テンプレート領域が白い矩形で示されている。また、画面左下には親カメラからの侵入物体位置情報をもとにその位置変化も重ねて表示している。子カメラと親カメラが位置情報を交換しつつ協調して侵入物体を追尾しており、監視員が容易に侵入者の表情や動作を把握できるということが分かる。

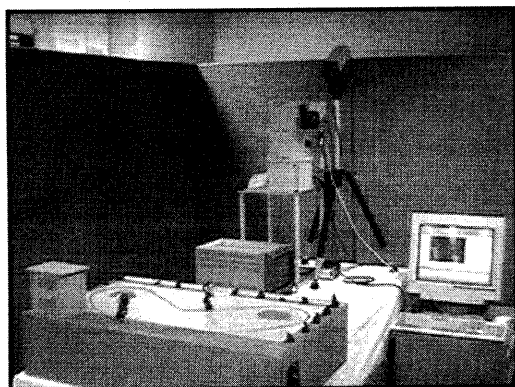


図9 親子型追尾システム実験風景

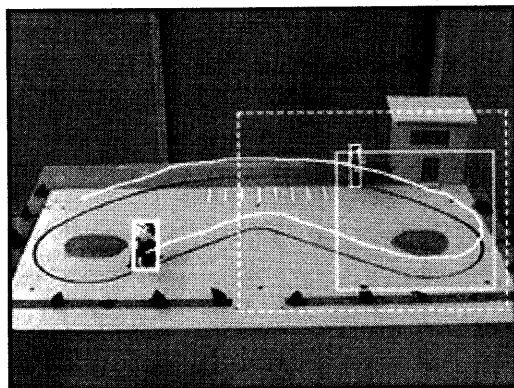


図10 親カメラの監視処理結果  
(白い線は侵入物体の移動経路を表す)

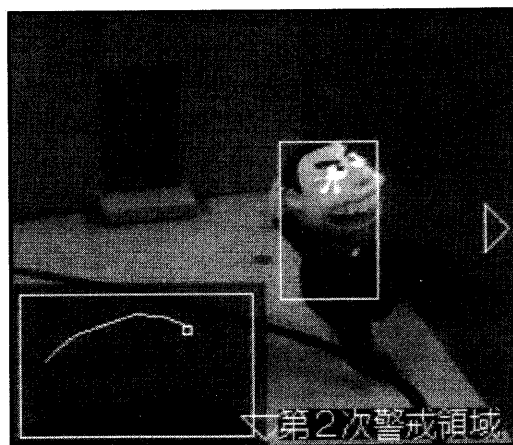


図11 子カメラの監視処理結果  
(追尾開始から 6 s 後)

物体を検知して追尾に移行した後、ズームレンズをワイド状態からズームアップ状態へと自動的に変化させることができるようなシステムも開発している。図12は、実際に屋外に設置したドーム型監視カメラの様子であり、図13に追尾処理の結果の画面例を示す。(a)はズームアップ前、(b)はズームアップ後である。

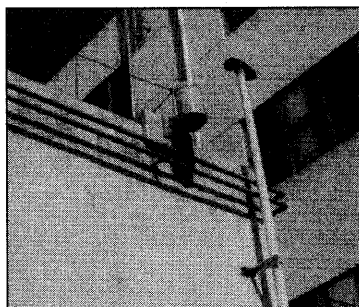


図12 屋外に設置した監視カメラ

### 8. ハードウェア構成例

図14に試作した親子型協調監視システムの構成例を示す。親カメラは、カメラがHITACHI KP-D51 (1/2型CCD、最低被写体照度: 2lx) であり、この映像をパソコン (画像認識処理ボード HITACHI IP-5005を内蔵) により処理している。子カメラは、画像処理ハードウェア内蔵カメラ HITACHI IPcam4010SHを、高速電動雲台 HITACHI HC-210 (最大回転速度: パン 180°/sec、チルト60°/sec) に搭載している。画像認識処理ボードと画像処理ハードウェア内蔵カメラとは、RS485 を使って通信することができるので、これを利用して親子間の通信を実現しており、子カメラは複数台増設可能である。

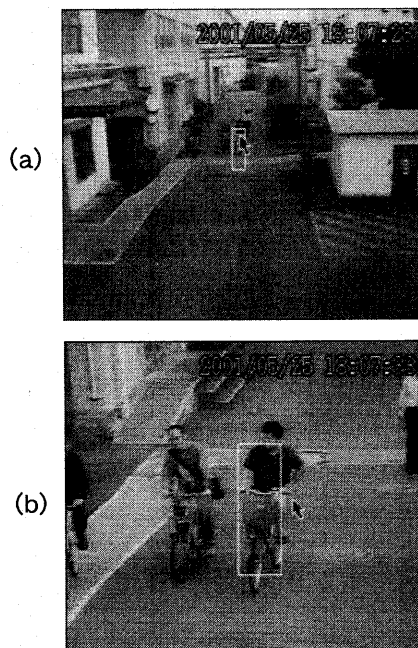


図13 自動ズームアップ追尾処理結果

### 9. まとめ

筆者のグループが日立国際電気(旧日立電子)で開発してきたシステムの、基本的なコンセプトにつ

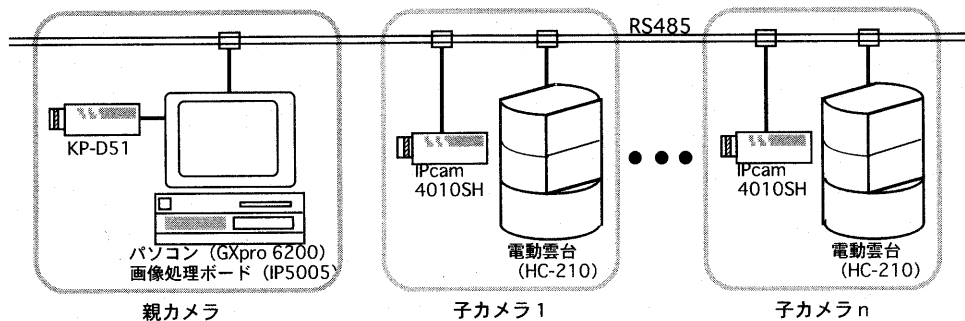


図14 親子型協調監視システムの構成

いて述べ、いくつかの実用例を紹介し、CVに技術よる身体動作の解析という研究開発分野における実時間(屋外)監視システム固有の実用上の問題点、筆者らが開発してきた具体的な解決方法について述べた。今後の課題としては、何と云ってもロバスト性の向上が上げられる。まだまだユーザーのニーズを100%満足させることはできていない。屋外の複雑背景下、あらゆる天候の変化、時々刻々変化する照明条件、そのような状況での任意の服装を身に付けた人物の検知能力という観点で見たときに、人間の目の能力の何分の一であるか。まだ、そういうことを数値化して言うことすらおこがましいというのが現状である。我々が開発してきた各種の方式は、それぞれが独立性の高いアルゴリズムであり、組み合わせることで互いに相補的に動作することができるようにすることを強く意識している。近年のパソコンのCPUを含む画像認識ハードウェアの性能向上は目覚ましいものがある。従って、組み合わせることができる持ち駒をたくさん持っていることは、そのままロバスト性の向上につながると言える面がある。実際、我々の実用システムにおいても、年々向上するハードウェアの処理速度の向上に合わせて、かつては速度重視のために面積を小さくしたり、間引いたりしていた処理画素数を徐々に増やすことで識別能力を上げ、さらには複数のアルゴリズムを並行して走らせて、その処理結果を

時間的・空間的に統合的な判定を行うということが可能となり、ロバスト性は大きく向上している。特に後者が大切であり、今後もハードウェアのみに頼ることなく、本当にインテリジェントなアルゴリズムの開発に邁進したいと考えている。

#### 参考文献

- 1) 元木 他：『小特集 モニタリング技術』, TV学会誌, Vol.140, No.3, pp.260-283 (1995)
- 2) 伊藤 他：『パソコンによる監視カメラのエージェント化を指向した実時間認識・追跡システムの試作』, 第2回知能情報メディアシンポジウム予稿論文集, pp167-174 (1996)
- 3) 上田 他：『IP-2000/5000シリーズの監視システムへの実用例について』, 第3回動画画像処理実利用化研究報告会, pp6-11 (1999/03)
- 4) 伊藤 他："画像認識ハードウェア内蔵カメラを用いた自律動物体追尾監視カメラ", SII'99 講演論文集, pp.13~18 (1999)
- 5) 伊藤 他：『物体検出状態の階層的解析による複数物体追跡方式の提案』, 信学技報, PRMU-97-8, pp57-64 (1997)
- 6) 谷口 他：『スキーヤの動きの画像処理的解析』, TV学会誌, Vol.43, No.12, pp.1370-1374 (1989)
- 7) 太田直哉：『信頼情報をもったオプティカルフローからの形状復元とその移動物体検出への応用』, 信学会論文誌, Vol.J76-D-II, No.8, pp.1562-1572 (1993)
- 8) 伊藤 他：『物体の見かけの大きさ変動に追従できるテンプレートマッチング法』, 信学技報, PRMU96-43, pp.45-50 (1996)
- 9) 伊藤 他：『光の反射と影に対してロバストな物体抽出法』, 1995年TV学会年次大会予稿, pp.195-196 (1995)
- 10) 伊藤渡, 上田博唯：『親子カメラによる自律型侵入者認識追尾システム』, 情処研報, Vol.2000, No.50, 2000-CVIM-122, pp109-114, (2000)