

特集 「センシングウェブ」

カメラ映像におけるプライバシー対処のためのアプローチ

Approaches for Managing Privacy Information in Video Images from Cameras

角所 考
Koh Kakusho

京都大学学術情報メディアセンター
Academic Center for Computing and Media Studies, Kyoto University.
kakusho@media.kyoto-u.ac.jp, <http://www.mm.media.kyoto-u.ac.jp/members/kakusho/>

満上 育久
Ikuhisa Mitsugami

(同上)
mitsugami@mm.media.kyoto-u.ac.jp, <http://www.mm.media.kyoto-u.ac.jp/members/mitsugami/>

美濃 導彦
Michihiko Minoh

(同上)
minoh@media.kyoto-u.ac.jp, <http://www.mm.media.kyoto-u.ac.jp/>

Keywords: privacy information, camera system, surveillance, ubiquitous sensor network, information filtering, real-world information.

1. はじめに

ウェブとセンシングウェブ [美濃 09] との違いは、前者が人手で入力された編集型の情報を提供するのに対し、後者はセンサによって自動獲得された観測型の情報を提供することである。このことにより、センシングウェブは、時々刻々変化していく実世界の動的な状況を実時間で捉えた情報（これを**実世界情報**と呼ぶ）を提供可能な情報基盤となり得る。その一方、観測型の情報には被観測者のプライバシー情報が潜在し、特にカメラの映像にはプライバシー情報が含まれやすい。

上のような実世界情報を、プライバシー問題を回避しつつ社会に広く流通させ、有効活用するためには、カメラ映像に含まれるプライバシー情報への対策が必要となる。このような問題には法律的観点からの検討 [小林 07] が必須なのはいうまでもないが、最近話題の Google 社の Street View [Google] に代表されるように、法律的議論をまたずにカメラ画像が巷にあふれつつある状況を考えれば、技術的に実現可能な範囲内で、プライバシー問題が少なく、かつ実世界情報として有用と思われる情報をカメラから得るための方法論について検討することもまた重要と思われる。本稿ではこのためのアプローチについて、情報学的な観点から議論してみたい。

2. カメラとセンサ

カメラは光学信号の受容器という意味で紛れもなく

センサの一種であるが、人感センサなど、カメラ以外の一般のセンサと比べると、プライバシー問題をはらむ可能性が格段に高い。しかし一般のセンサ情報がプライバシーの問題なく有効な実世界情報を提供できているのであれば、同じセンサであるカメラについても同様の状況を実現できないものだろうか？

この発想は、監視カメラシステムのようにカメラの目的を“見る”ことそのものに置くのではなく、カメラを実世界情報を得るためのセンサの一つに過ぎないと位置づけるセンシングウェブにおいて、カメラに対するプライバシー対処のアプローチを考える際の有益な示唆を与える。そこで本稿では、まずこの点から議論を始める。

2.1 カメラと他のセンサとの違い

プライバシー問題のないセンサ情報の有効活用のイメージを明確にするために、卑近で恐縮であるが、公衆トイレに多く設置されている人感センサを例に取り上げる。このセンサの設置目的は、トイレ使用后、自動的に水を流すことであり、センサ情報はその目的のためだけに閉鎖的に利用されている。しかし、もしこの情報をインターネットなどで公開すれば、外出中などに空いているトイレを探せる情報サービスの実現などに活用できる。この例は、既設センサの情報の公開によって実世界情報基盤が創出されるというセンシングウェブの一つの実現イメージである。ちなみに、このような情報サービスの有用性は、飛行機や新幹線、高速道路のサービスエリアなどでトイレの空き状況が電光表示されていたり、公衆トイレの位置や設備の情報を提供する “Check a Toilet”

[CheckAToilet] というサイトが出現している事実からも示される。

公衆トイレの人感センサは、プライバシー的に最も問題の大きな場所に設置されているにもかかわらず、プライバシー問題は生じておらず、その情報を一般公開したとしても、プライバシー問題が生じるとは考えにくい。一方、もしこれをカメラに置き換えれば、直ちに重大なプライバシー問題が起きそうである。同じセンサなのにこのような大きな違いが生じる原因はどこにあるのだろうか？

人感センサは“直前の障害物の有無”を検出・出力する機能をもつが、その原信号はそのような記号的な情報ではなく、アナログ信号である。特に赤外線方式の人感センサの場合、原信号は対象からの赤外線反射光強度を表す光学信号であり、性質的にはカメラの信号と同じである。それならばカメラも“直前の障害物の有無”を検出するのみにまで機能を落とせば、人感センサと等価な装置となるので、プライバシー問題はなくなるはずである。しかし、カメラの機能をそこまで落とさない限りプライバシー問題は解決しないのであろうか？ もしそうでないのであれば、どのような機能までならば、プライバシー問題を回避できるのであろうか？

これらの疑問に対する答えを得るために、人感センサがなぜプライバシー問題を生じないのか考えてみる。

2.2 記述の粒度

人感センサがプライバシー問題を生じない理由の一つは、光学信号の粒度が粗く、その中にプライバシーに関わる情報を反映できないためと考えられる。我々が実世界を記述する際の概念階層を考えた場合、個人情報とは具体性の高い階層に存在するが、光学信号の粒度を粗くすると、その中には抽象度の高い階層の情報しか残らないため、プライバシー問題が低減される。これに対してカメラ映像は、光学信号の粒度が細かいため、その中にさまざまな情報が反映される。

テレビ番組などではプライバシー保護のためにモザイクをかけることが行われているが、これはその部分の解像度を下げていることに相当する。また、モザイクをかけるべき顔の部分の部分を自動抽出して監視カメラ映像のプライバシー処理に利用する **Stealth Vision** と呼ばれる技術も提案されている [北原 04a, Kitahara 04b]。

解像度を下げることと粒度を粗くすることは、似てはいるが厳密には同義ではない。監視カメラ映像に対して、情報の記述粒度を粗くすることでプライバシー問題への対処を図った例としては、**PriSurv** と呼ばれる技術がある [馬場口 07]。前述のように、映像の記述粒度を粗くすると、含まれる情報の具体性が下がることから、**PriSurv** では、人物映像に含まれる情報の抽象化度を代表的な画像処理による記述粒度の変化と対応させて図1のように定義し、監視カメラ映像配信の際に被写体と利用者の関係に応じた適切な抽象化度を選択することを提

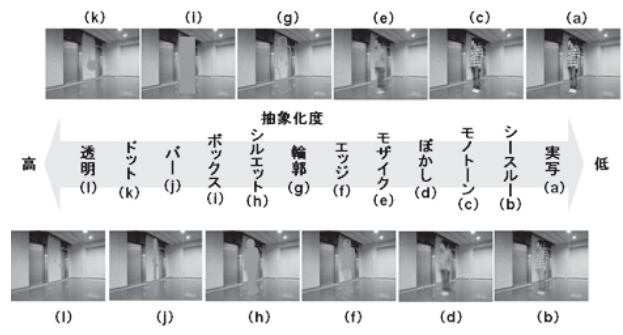


図1 カメラ映像における抽象度の階層 [馬場口 07]

案している。

2.3 情報の明示性

人感センサがプライバシー問題を生じないもう一つの理由は、原信号が、プライバシーに無関係な“直前の障害物の有無”のみを明示的に表現した記号的な情報に加工されたうえで出力されており、それ以外の情報が混入する余地がないためと考えられる。

ウェブで用いられるテキストに代表される記号表現は、記述すべき情報を、表現単位である個々の記号によって直接明示的に表現するため、情報記述の際に、用いられている記号のもつ意味以外の情報が混入する可能性は低い。これに対して、カメラが出力する映像はパターン表現である。パターン表現は、記述すべき情報を、それと意味的には直接結びつかない表現単位（画素など）の集合によって間接的かつ暗黙的に表現するため、情報記述の際に、本来意図した以外のさまざまな情報が混入することがむしろ普通である。

このようなパターン表現としての暗黙性により、カメラ映像はその中に含まれる情報をコントロールしにくく、プライバシー問題を生じやすい。人感センサでは、パターン情報である原信号を、“直前の障害物の有無”という記号的な表現に変換することで、それ以外にプライバシーなどの意図しない情報が混入することが回避されている。

3. プライバシ情報と実世界情報

2章の議論から、記述の粒度と情報の明示性を確保することで、カメラをプライバシー的に一般のセンサ並みの装置にまでは変換できる可能性が期待できる。しかし、そのことによって逆に情報の利用価値もなくなってしまう心配はないのだろうか？ そこで次にこの点について本章で議論する。

3.1 実世界情報と変動性

センシングウェブに期待されるのは実世界情報基盤としての役割である。このとき、“実世界情報”を単に“実世界に関する情報”というだけの意味と考えるならば、そのような情報は、ウェブからも得ることができる。セ

ンシングウェブが存在価値をもち得るには、実世界に関する情報の中でも、ウェブからは得られない情報を提供できなければならない。

ウェブの情報は人手で入力されるため、人が更新しない限り変化せず、ニュース速報のようなものでも更新頻度は数十分程度である。一方、センシングウェブの情報は、センサから新たな情報が獲得される度に更新されていくため、それが通常のビデオカメラの場合には、30msごとに更新されることになる。このようにセンシングウェブはウェブに比べて情報の更新頻度が格段に高いのが大きな特徴である。

更新頻度の高い情報が価値をもつのは、実世界に関する情報の中でも時々刻々変わる動的変化を反映した情報である。すなわち、センシングウェブで提供される実世界に関する情報の中で、最も高い利用価値をもつのは、1章でも述べたように“時々刻々変化していく実世界の動的状況を実時間で捉えた情報”であり、これが本稿で“実世界情報”と呼んでいるものである。

3.2 プライバシ情報と不変性

一方、プライバシー問題を生じ得る情報とは、基本的には個人の特定に関わる情報と考えられるが、このような情報は普通、時間とともに変化しない。例えば我々の外見は、位置や姿勢、表情、服装、髪型などの変化によって頻繁に変化するが、それにもかかわらず外見から個人を同定できるのは、その中にある変わらない“その人らしさ”を捉えているためである。

パターン認識においても、あるカテゴリーの認識に用いられるのは、時間的に変化しない特徴量である。例えば、視点位置の異なる複数カメラからの映像に基づいて物体のカテゴリー（個人名を含む）を認識する処理は、図2に示すように、さまざまな視点・時刻での観測情報の中から、同一の対象に関する部分を抽出するセグメンテーション（カメラ間の対応づけ、トラッキングを含む）

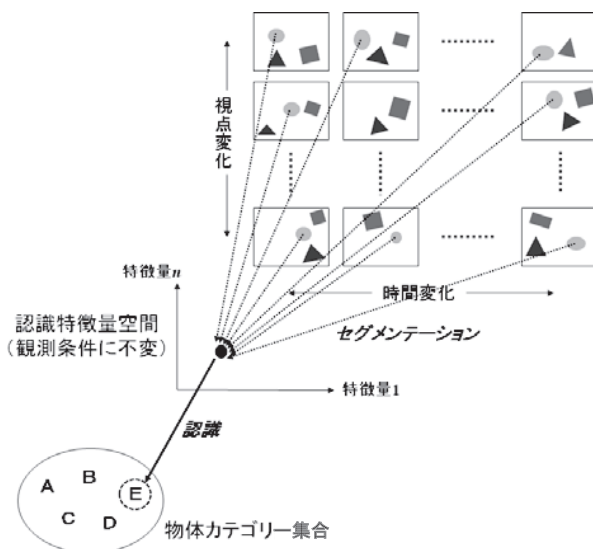


図2 物体認識処理のための特徴量

と、そこで抽出された観測情報を、その中に共通する特徴量を手掛りに物体カテゴリーに対応づける認識処理によって構成される。これらの処理を実現するうえで最も重要なのは、視点や時刻などに依存せず、同一対象に対して安定的に観測される“変わらない”特徴量を見つけることであり、それが物体認識処理問題の本質といえる。

このことは、センサ情報を蓄積することの危険性も浮かび上がらせる。情報の蓄積は情報技術の基本であるが、センサ情報を蓄積することで、不変性の高い情報の抽出が可能となるため、プライバシー問題を生じる危険性は高くなる。

3.3 プライバシ情報と実世界情報の排反性

以上のことから、センシングウェブにおいて有用性をもつ実世界情報と、プライバシー問題を生じ得る情報は、時間変化という点では正反対の性質をもつことがわかる。この排反性により、プライバシー情報を削除しても、センシングウェブとしての情報の有用性は失われないことが期待される。例えば道路を撮影したカメラ映像において、車のナンバプレートや車体の形状・色といった情報は時間とともに変化せず、かつプライバシー問題を生じやすい。また各車種の車体形状や色のバリエーションといった情報は、必要ならばウェブからも容易に得られるため、わざわざセンシングウェブから入手する必要はない。逆にセンシングウェブから得る価値のある情報として考えられるのは、道路の混雑状況や車の流れといったものであり、これは各時刻の車の位置や向きといった時々刻々変化する情報に基づいている。したがってこのように時間変動性の高い情報を中心にセンサ情報の活用を検討すれば、プライバシー問題は生じにくいと考えられる。

4. プライバシ処理のためのアプローチ

2～3章での議論の内容を踏まえて、センシングウェブのカメラ映像から、プライバシー問題を生じにくく、かつ有用性のある実世界情報を抽出するための具体的なアプローチを列挙してみると、以下ようになる。

4.1 プライバシ情報が含まれない配置をする

プライバシー情報と実世界情報の間に3.3節で述べたような排反性が存在することを前提として、カメラ映像へのプライバシー問題に対処することを考えた場合、最も理想的な解決法は、カメラ映像の中にそもそも実世界情報しか含まれないようにすることである。これは、プライバシー問題のある映像が映らないようにカメラの設置位置や方向を工夫することに相当する。実際、テレビのニュース番組などでも群衆の後ろ姿や足元のみを撮影した映像が用いられることがある。ただし街角に固定カメラを設置することを考えた場合には、常に後ろ姿のみが映るように撮影することは難しいし、足元のみが映る位置では歩行の邪魔になりやすいことから、現実的な設置方法

は、頭上から見下ろすように設置することである。この方法であれば、顔が映らないのでプライバシー問題が生じにくいだけでなく、実世界情報として人数変化や動線といった人に関する動的情報を獲得するために人物追跡などを行う場合、隠れが生じにくいという副次効果も期待できる。

4.2 問題のあるプライバシー情報を削除する

4.1 節で述べた方法はシンプルで効果的ではあるが、カメラの設置位置に関する物理的制約などから常に利用できるとは限らない。したがって、カメラ映像の中にプライバシー情報が含まれてしまった場合のアプローチも当然必要となる。

このためのオーソドックスなアプローチは、プライバシー問題のある部分を削除するというものである。具体的には、人物領域や顔領域を削除する、その部分の解像度を下げる、あるいは点や線などで抽象化して表示する、といった方法が考えられる。Google社のStreetViewや2.2 節で紹介したStealth VisionやPriSurvなどはまさにこのアプローチを採用したものといえる。

このアプローチで問題となるのは、画像処理技術の精度の問題である。センシングウェブでは既設センサの利用を前提とすることから、カメラ映像の撮影条件や撮影シーンにはさまざまなものが考えられ、これを特定のクラスに限定することはできない。しかしそのような場合、人物や顔などの領域を100%の精度で抽出することは非常に難しいため、本来削除すべきプライバシー情報を実際に確実に削除することは困難となる。

4.3 有用な実世界情報のみを抽出する

プライバシー情報への対処のための第三のアプローチは、プライバシー情報を削除するのとは逆に、むしろ実世界情報のみを抽出するというものである。このアプローチは、図3に示すように4.2 節のアプローチとはちょうど裏表の関係となる。

センシングウェブの目的は、プライバシー情報を含まない映像データを得ることではなく、実世界情報を獲得することである。この目的を直接達成とすることを考えれば、3.3 節で述べたようにプライバシー情報と実世界情報の排反性により、自ずとプライバシー問題の少ない情報が得られるものと期待される。



(a) プライバシー情報の削除 (b) 実世界情報の抽出

図3 プライバシー処理のための相補的アプローチ

4.4 情報をシンボル化する

4.3 節のアプローチを用いる場合、実世界情報の記述に用いる意味表現にはさまざまな選択肢があり得る。しかし画像のようなパターン表現を利用すると、2.4 節で述べた暗黙性の問題から、意図しないプライバシー情報が混入する可能性がある。特に、実世界情報の粒度を細かくすると、プライバシー情報を含む危険性が高まる。このような状況では、4.2 節のアプローチと4.3 節のようなアプローチの境界は曖昧となり、両者は相補的ではあるにせよ、あまり本質的な相違はなくなる。

4.3 節のアプローチの特徴を生かすには、実世界情報の記述を明示的なものにする必要がある。これは、実世界情報をテキストや図形のような記号的な意味表現によって記述することを意味する。例えば3 章で述べた人感センサは、この語彙集合として、直前に障害物が、“ある”、“ない”の二つを用いていることになる。このときセンサによって提供される実世界情報のクラスは、記述の単位となる語彙集合によって決まることから、これをどのように選ぶかが問題となる。これはセンサ情報の利用目的に依存するため、次章で具体的な事例を紹介する。

5. 道路状況把握のための映像処理

4.4 節で述べたアプローチにおいて、センサが記述すべき実世界情報のクラスは、センサ情報の利用目的や利用場面に依存する。本節では、センシングウェブに利用できる既設カメラの設置事例が多く、かつそこから得られる実世界情報の有用性が高いと思われる例として、道路交通管制や人流調査等の目的で市街地の屋内外に設置された既設カメラを利用して道路状況を把握する場合を取り上げる。

5.1 実世界情報の記述語彙と映像処理過程

本節でいう道路状況把握とは、道路の混雑度や人通りの多さなどを把握することをいう。このような用途は例えば、勤務先からの帰宅時に、“帰りが遅くなったので、人通りが少ないようならタクシーを呼ぼうと思うが、人通りはどの程度か？”と考えているような状況で生じる。このための実世界情報を記述する語彙として考えられるのは、対象物体の種類（車、人、自転車など）、数、移動速度であることから、これらを語彙とした実世界情報を抽出することを考える。

これらの語彙をセンサの出力とするには、それぞれに対応する実世界情報をカメラ映像から抽出する処理が存在する必要があるが、これは、理論的には図4のような映像処理の流れの中で順次抽出可能である。そこで、通常のカメラを利用して上記のような語彙を出力とする実世界情報センサを実現することを考える。

5.2 長時間観測による観測条件のモデル化

5.1 節で述べた映像処理は、コンピュータビジョンに

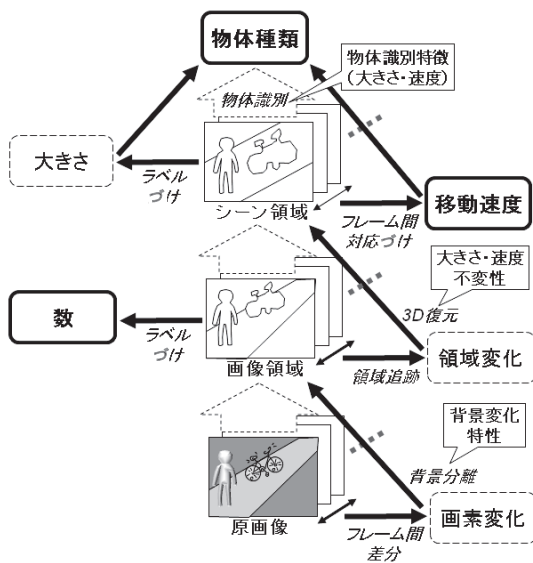


図4 映像処理の流れと実世界情報の記述語彙

における伝統的な物体追跡や物体認識の問題を含む。このような処理をセンシングウェブのような既設カメラの映像を対象とした多様な環境で安定的に実現することは一般には難しい。ここでこの問題の解決の鍵となるのは、長期間観測である。

3.2節で述べたように、物体認識処理の本質は、同一対象に対して安定的に観測される不変的な特徴量を見つけ出すことである。しかし、任意の観測条件に普遍的に有効な特徴量を見つけるのは困難なため、いきおい利用環境を限定し、その環境に有効な特徴量を設計するに留まることとなる。ただしその場合、ほかの環境にも適用しようとする、特徴量やパラメータの再チューニングが必要となるため、汎用性が低いと批判されることになる。しかし特定の利用環境ならば曲がりなりに処理が実現できるということは、裏を返せば任意の環境で任意の観測条件が一様に生じるわけではなく、観測条件の生じ方には利用環境に依存した偏りがあることを意味する。

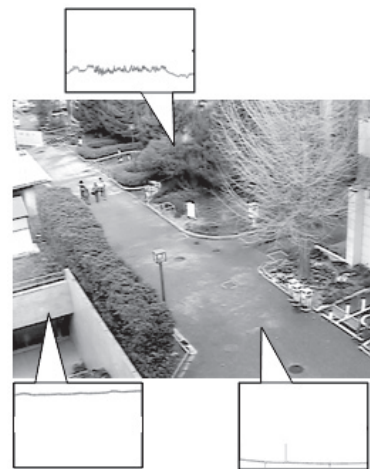
センシングウェブで利用される既設カメラは、通常、環境内に非常に長期間にわたって固定されていると考えられる。したがって、この過程で得られるさまざまな状況での環境の観測データを分析すれば、その環境での観測条件の生じ方に関する情報が得られ、その環境に応じた特徴量やパラメータの設計に活用できると考えられる。この考え方は、図4の処理では、背景分離のための背景変化特性のモデル化、3D復元のための対象の大きさ・速度の不変性のモデル化、物体識別のための識別特徴や識別関数のモデル化など、各処理に適用可能であるが、以下ではその一例として、背景分離処理の実現例を紹介する。

5.3 長時間観測に基づく背景変化のモデル化

固定カメラからの映像を対象とした背景分離処理には通常、背景差分やフレーム間差分が用いられる。これは、

“背景”を特徴づける不変的な特徴量として画素値の不変性を採用し、画素値の変わらない領域を背景として分離しようとするアプローチといえる。しかし、実際には背景であっても画素値が全く変わらないことはあり得ないので、どのような変化までを事実上変化していないものとみなすかという背景変化特性のモデル化が必要となる。実際、従来研究では、背景自体あるいは背景変化の原因となる照明条件の変化などを、しきい値や分布関数などを用いてモデル化することが行われている [松山01, 長屋96, 島田07, 島井03]。

したがって、個々のカメラについて、その設置環境における背景を長期観測したデータを入手し、それをモデル作成に活用すれば、利用環境に依らず、その環境に依存した精密なモデルが作成できるため、背景分離処理が



(a) 屋外画像の背景各点での輝度値変化



(b) テスト画像



(c) 背景差分結果(白色ほど差分が大)

図5 長時間観測に基づく背景分離処理の例 [八代08]

容易となることが期待される。このことを実際の観測画像に基づいて確認してみた例が図5である。長期観測により、背景各点に対して図中(a)のような輝度値変化が観測されるため、これに基づいて背景変化特性をモデル化し、(b)のような画像に対する背景差分を求めた結果が(c)である。斜線で示す領域は、長期観測の結果、背景以外の物体が存在し得ないと判断された領域を示し、それ以外の領域では、濃淡値で背景差分の大きさを表している。人物の部分が白色として抽出できていることがわかる。

6. ま と め

本稿では、センシングウェブにおいて、プライバシー問題を避けつつ有用性の高い実世界情報を公開・利用するためのアプローチについて、情報学的な観点から議論した。本稿で述べた内容は、まだ概念レベルの議論に留まっている部分が多く、5章で紹介したような実現例を含め、実際の利用状況への適用を通じた有効性の検証は十分ではない。このため、現在、ショッピングモールなどのような公共環境での実証実験の準備を進めているところであり、その結果については、機会が与えられればまた報告させていただきたい。

プライバシーの問題は根本的には人の主観に依存するものであるため、非常に奥が深く複雑である。本稿で述べたようなアプローチは、カメラ映像に含まれる実世界情報とプライバシー情報の間の技術的観点からの線引きをするうえではある程度有効とは考えているものの、これのみによって、本来カメラ映像の捉えている実世界のありのままの様子を、プライバシー情報を全く混入させることなく余すところなく捉える、という究極目標を達成することは難しい。願わくは、現時点でプライバシー問題を危惧して閉鎖的にのみ利用されているさまざまな場所の既設カメラからの情報を、完全に埋もれさせてしまうことなくある程度有用に活用するための技術的枠組みを構築していくうえで、本稿での拙い議論が多少とも示唆を与えるものとなれば、望外の喜びと考える次第である。

謝 辞

本研究は、文部科学省の科学技術振興調整費（科学技術連携施策群の効果的・効率的な推進）による「センサ情報の社会利用のためのコンテンツ化」の一環として実施したものである。

◇ 参 考 文 献 ◇

- [馬場口 07] 馬場口登：プライバシーを考慮した映像サーベイランス、情報処理、特集 安全と安心のための画像処理技術、Vol. 48, No. 1, pp. 30-36 (2007-01)
- [CheckAToilet] <http://www.checkatoilet.com/>
- [Google] <http://maps.google.com/help/maps/streetview/>
- [北原 04a] 北原 格, 小暮 潔, 萩田紀博: Stealth Vision: 被写体のプライバシーを保護する映像獲得方式, 信学技報, PRMU2003-299, pp. 89-94 (2004)
- [Kitahara 04b] Kitahara, I., Kogure, K. and Hagita, N.: Stealth vision for protecting privacy, Proc. 17th Int. Conf. on Pattern Recognition (ICPR 2004), Vol. 4, pp. 404-407 (2004)
- [小林 07] 小林正啓：人物を認識することの法的問題点、情報処理、特集 安全と安心のための画像処理技術、Vol. 48, No. 1, pp. 37-42 (2007-01)
- [松山 01] 松山隆司, 和田俊和, 波部 齊, 棚橋和也：照明変化に頑健な背景差分, 信学論 (D), Vol. 84-D2, No. 10, pp. 2201-2211 (2001)
- [美濃 09] 美濃導彦：センシングウェブ概念と課題一、人工知能学会誌, Vol. 24, No. 2, pp. 179-184 (2009)
- [長屋 96] 長屋茂喜, 宮武孝文, 藤田武洋, 伊藤 渡, 上田博唯：時間相関型背景判定法による移動物体検出, 信学論 (D), Vol. 79-D2, No. 4, pp. 568-576 (1996)
- [島田 07] 島田敬士, 有田大作, 谷口倫一郎：適応的な分布数の増減法を利用した金剛ガウス分布による高速な動的背景モデル構築, 信学論 (D), Vol. 90-D, No. 9, pp. 2606-2614 (2007)
- [島井 03] 島井博行, 栗田多喜夫, 梅山伸二, 田中 勝, 三島健稔：ロバスト統計に基づいた適応的な背景推定法, 信学論 (D), Vol. 84-D2, No. 6, pp. 796-806 (2003)
- [八代 08] 八代武大, 満上育久, 角所 考, 美濃導彦：画素値の時間変化相関に基づく前景領域抽出, 画像の認識・理解シンポジウム (MIRU) 2008 論文集 (2008)

2009年1月9日 受理

著 者 紹 介



角所 考 (正会員)

1988年名古屋大学工学部電気学科卒業。1993年大阪大学大学院工学研究科通信工学専攻博士課程修了。1992～94年日本学術振興会特別研究員。1993～94年スタンフォード大学ロボティクス研究所客員研究員。1994年大阪大学産業科学研究所助手。1997年京都大学総合情報メディアセンター助教授(現 学術情報メディアセンター准教授)。視覚情報メディア、コミュニケーション、インタラクションに関する研究に従事。博士(工学)。IEEE, ACM, 電子情報通信学会, 情報処理学会などの各会員。



満上 育久

2001年京都大学工学部電気電子工学科卒業。2007年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了。同年、京都大学学術情報メディアセンター研究員(科学技術振興)。コンピュータビジョン、対象抽出・追跡、プロジェクトによる実環境視覚情報提示に関する研究に従事。博士(工学)。

美濃 導彦は前掲 (Vol. 24, No. 2, p. 184) 参照。