

3次元位置計測システムとステレオカメラを用いた手早い環境モデリング手法と人の行動イベント検出への応用

西田 佳史^{†,††} 北村 光司^{†††} 木村 誠^{†,††} 相澤 洋志^{††††} 堀 俊夫^{†,††}
溝口 博^{†††}

[†] 産業技術総合研究所 〒135-0064 東京都江東区青海 2-41-6

^{††} CREST, JST 〒604-0847 京都府京都市中京区烏丸押小路上ル秋野々町 535 日土地京都ビル 8F

^{†††} 東京理科大学 〒278-8510 千葉県野田市山崎 2641

^{††††} 東京電機大学 〒101-8457 東京都千代田区神田錦町 2-2

E-mail: †{y.nishida,makoto.kimura}@aist.go.jp, ††j7503618@ed.noda.tus.ac.jp

あらまし 日常環境における人の行動を計算機に入力することができれば、特定の人に対して適切に、サービスを提供したり、生活の支援を行うことができる。そのためには、人間の行動を実時間で頑健に観察する機能が必要である。本論文では、認識対象としたい行動を実時間で頑健に認識する機能を手早く実現する手法として、3次元位置計測システムとステレオカメラを用いて、対象物の物理構造や物理的機能といった対象物の性質を、計算機上でユーザーが対話的にモデル化することで、対象物自体を人間の行動を認識するための仮想的なセンサとする手法を提案する。また、提案手法の有効性を日常生活空間を模擬した環境で検証する。

キーワード 行動認識, 超音波センサ, コビキタス・コンピューティング, 3次元位置計測システム

Method for Quickly Virtualizing Real World Using 3D Location System and Stereoscopic Camera and Its Application to Detection of Human Activity Events

Yosifumi NISHIDA^{†,††}, Koji KITAMURA^{†††}, Makoto KIMURA^{†,††}, Hiroshi AIZAWA^{††††}, Toshio HORI^{†,††}, and Hiroshi MIZOGUCHI^{†††}

[†] National Institute of Advanced Industrial Science and Technology Aomi 2-41-6, Koutou-ku, Tokyo, 135-0064 Japan

^{††} CREST, JST Nittochikyouto Bldg. 8F, Akinonochi 535, Oshikouji-Agaru Krasuma Nakagyou-ku Kyoto-shi, Kyoto, 604-0847 Japan

^{†††} Tokyo University of Science Yamazaki 2641, Noda-shi, Chiba, 278-8510 Japan

^{††††} Tokyo Denki University Nishiki-cho Kanda 2-2, Chiyoda-ku, Tokyo, 101-8457 Japan

E-mail: †{y.nishida,makoto.kimura}@aist.go.jp, ††j7503618@ed.noda.tus.ac.jp

Abstract This paper proposes a method for quickly virtualizing and sensorizing real world for detecting target human activity events. The proposed method involves methods for 1) modeling a structure model for simplifying real world structure using stereoscopic camera, 2) creating a phenomenal model for binarizing real world phenomena using virtual sensors/effectors, and 3) associating binarized data from the phenomenal model with target human activity events. This paper also reports an experimental system which consists of a stereoscopic camera and an ultrasonic 3D tag system. The ultrasonic 3D tags are attached to an object to assign ID to the object and measure the location of the object. The stereoscopic camera is used for creating the simplified 3D shape model of the object. With the system, a user can interactively create a function for detecting target human activity events. The experimental results conducted in a daily living space are reported.

Key words Ubiquitous Computing, Sensorization, Distributed Sensor, Sensor Fusion

1. 緒 論

近年の情報処理技術の発展に伴い、ユービキタスコンピューターに代表されるように、情報処理技術を利用した新しいサービスの提供や生活支援を行えるシステムの開発が望まれている。そのようなシステムが、人間の行動を観察し、認識することが可能となれば、人間を中心とした人間支援が可能になると考えられる [1]。人の行動が生じる現場で、人を観察するシステムを実現するためには、観察システムが環境の多様性に対応していくことが求められる。本研究では、人の行動が生じる現場で、人間の行動を観察する機能を手早く実現するための手法について述べる。

一般に、人間の行動を観察するシステム、すなわち、人の行動を計測し、計測データから必要な情報を読み取るシステムは、大きく分けて2つの機能が必要である。人間行動を計測する機能と、それ自体意味を持っていない計測データから欲しい情報を抽出する機能である。行動計測機能に関する問題点として以下のものが考えられる。実世界で行動計測機能を実現させるためには、一般的にセンサなどを広範囲に分散して配置するため、システムの物理的に大きくなりがちになり、設置に時間を要する。また、システムがセンサがどこにあるかという情報を知っていなければならないので、キャリブレーションやキャリブレーションを不用とするための完璧な設計が必要となり手間を要する。

行動情報抽出機能に関する問題点としては以下のものがあげられる。得られた情報から人間行動を検出する機能を実現させるためには、観察対象としたい行動を記述したり、得られた情報から欲しい情報だけを抽出するためのプログラミングが必要となるため、手間がかかる。また、観察対象の行動を変更する際には、プログラムを再び作成しなければならないか、または汎用的に用いることができるプログラムが必要となるが、実際には人間行動の多義性から実現が困難である。対象領域に関する様々な知識を用いることで、専用の認識機能を実現しようとする現実的なアプローチもあるが、そのような認識機能は、環境に依存するため、再利用性が低いという問題点がある。

従来から行動観察に関する研究が数多く扱われている。カメラなどを用いて行動の観察・認識を行う技術は、日常生活環境のような複雑で多様性をもった環境では、対象物を特徴付けることが本質的に困難であるために、実用性のある頑健な観察・認識技術を実現するには至っていない。

これに対する解決法として、頑健な行動計測を実現するために、センサを空間に分散配置し、人の行動計測を行うシステムが提案されている。圧力センサを部屋中に敷き詰めることで、人間の移動経路や向き、椅子に座ったかどうか、引き出しや窓の開閉、ベット上での人間の姿勢を検知・計測できるシステム [2] や、RF タグを使用した人間行動に関するシステム [3] が提案されている。しかし、これらの研究では、行動情報抽出機能を手早く実現する手法は扱われていない。

本研究では、行動計測機能に関して、これまでに、無線技術、小型プロセッサ技術を応用し、IDを持つことができる3次元位置計測システムのタグ（超音波式3次元タグ）を対象物に取

り付けることで実時間で頑健な計測を行う手法を提案し、検証してきた [4], [5]。また、行動計測機能を手早く実現するための手法として、システム構築の手間を最小限にするためのキャリブレーション手法を開発してきた [6]。本論文の目的は、これまで扱われてこなかった行動情報抽出機能の手早い実現手法に焦点をあて、行動を観察を行いたい環境や対象としたい行動を熟知しているユーザの知識を利用することによって、対話的に対象領域に最適な人間行動認識機能を手早く実現する手法を提案し、その有効性を日常生活空間を模擬した環境で検証することにある。

2. 手早い環境モデリング手法と行動イベントの検出手法の提案

2.1 環境モデリングにおける問題点と解決手法の提案

本研究では、対象物本来の機能や性質を維持しつつ、対象物自体を人間の行動を認識するためのセンサとして振舞うことが可能な対象物モデルを計算機上に作成し、このセンサ化された対象物モデルを用いて人間の行動を観察する手法を提案する。計算機上に作成する対象物モデルが持つ機能は、特に剛体やリンク構造を持った剛体に着目すると、対象物の物理的構造や、対象物内または対象物相互の接触などの物理現象にだけ注目すれば、簡単に記述することができる。

このような幾何モデルや物理的な現象のモデルを作成する際に、あらゆる状況を想定した汎用的なモデルを作成しようとすると、作成が困難であったり、ユーザにとって必ずオーバースペック、またはアンダースペックとなってしまう、実際には汎用的に使用しづらいという問題があるため、ユーザが必要とする情報を過不足なく含んだモデルを作成しなければならない。そこで、本研究では、モデルを用意しておくのではなく、ユーザが対話的に必要な対象物機能のモデリングを行うことで解決し、そのモデリング手順が簡便で手早く行える手法を提案する。本研究で提案する手法は、次の手法から構成される。

- 物理構造のモデル化手法
- 物理現象のモデル化手法
- 物理現象モデルの出力と行動イベントの対応付け手法
- 3次元位置計測システムとセンサ化された対象物を用いた行動認識手法

詳細については、以下で述べる。

2.2 物理構造のモデル化手法

対象物の物理構造は、対象物の性質や機能を記述する上で重要である。しかし、全ての対象物の幾何形状データを用意することは、不可能ではないが、極めて手間のかかる作業である。また、仮に、対象物のCADデータを入手できたり、3次元スキャナで対象物の3次元形状モデルを作成できたとしても、情報が詳細過ぎ、計算機上で対象物の干渉を調べるのに多大な計算コストが必要となる。観察対象としたい人の行動が明確である場合には、CADデータのように詳細なデータは不要であることが多い。そこで本研究では、ステレオカメラを用いたフォトモデル機能を利用することで、対象物の物理構造を単純な形状、例えば線形や四角形や多角形に抽象化することで、必要な物理構造のみを記述する手法を提案する。対象物構造を単純化する

という考え方はロボティクスの分野でも提案されており、マニピュレータにネジやナットを締める作業させる場合に、ネジやナットの情報は位置・形状・向きという必要最小限で記述する試みがなされている [7]。本研究で扱う人間の行動に関して例を挙げれば、一般的な机は大きく分けて、面部分と脚部分があると考えられる。人が物を机に置くことを検出した場合には、面部分のみの物理構造をモデル化すれば十分であり、脚部分の物理構造のモデル化は不用である。

2.3 物理現象のモデル化手法

対象物の物理現象をモデル化することで、対象物の機能を単純化してモデル化することが可能である。特に「接触」という物理現象に着目すると、机に物を置いたり、机から物を持ち上げたこと、ドアの開閉、引き出しの開閉、引き戸や窓の開閉、またこれらの類似した動きをするものなどといった、対象物の機能を記述することができると考えられる。そこで、本研究では、対象物内や対象物間で起こる物理現象を単純化させるために、接触状態、すなわち、接触しているか、接触していないかの二値を用いて記述する手法を提案する。具体的には、物理構造のモデル化によって作成された構造モデルに、あらかじめ用意された仮想センサ・効果器を計算機上で対話的に取り付けていくことで、手早く処理系を実現する手法である。このような物理現象の単純化により、多大な計算コストを必要とせずに、対象としたい物理現象が記述可能になる。

2.4 物理現象モデルの出力と行動イベントの対応付け手法

人の行動は多義的であり、認識対象としたい行動やそこから得たい情報は人によって大きく異なる。そこで、本研究では、上述した物理現象の二値情報にもとづいて、ユーザが必要なルックアップテーブルを作成し、行動情報を登録可能にする手法を用いる。例えば、コップと机の接触という物理現象が、オン状態になったときに "put a cup on the table" と得られたり、オフ状態になったときには "hold the cup" と得られるようになる。

2.5 3次元位置計測システムとセンサ化された対象物を用いた行動認識手法

上述した、物理構造モデル、物理現象モデル、行動イベントの対応付けの手順によって行動検出ソフトウェアが、対話的に手早く実現できる。本研究では、このように作成された対象物の仮想モデルが人間の行動を認識するためのセンサとして振舞うことができると考え、これを「対象物センサ」と呼んでいる。この対象物センサに、実際の対象物に取り付けられた3次元位置計測システムの3次元位置情報を入力することにより、対象物センサは実世界で動かされた対象物と同じ動きを計算機上で行うことができる。

3. 手早い環境モデリングと行動イベントの検出システムの構築

3.1 手早い環境モデリングと行動イベントの検出システムのハードウェア

3.1.1 超音波3次元タグシステム

本研究では、3次元位置計測システムとして、超音波を利用した、誤差が20~80mm程度で低コストであるモーションキャ

プチャシステムを開発してきた [4], [5]。このシステムは、天井や壁といった環境側に超音波の受信器を埋め込み、超音波発信器を含んだ小型装置である3次元タグ(図1)から、発信された超音波が受信器に到達する時間を複数計測することによって、3次元位置計測が行えるものである。また、超音波の発信は時分割で行っているため、超音波タグにIDを持たせることが可能で、それぞれを識別できる。本研究では、図2に示すように、このシステムを実装したセンサルームを構築した。

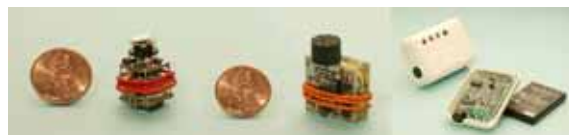


図1 無線型超音波3次元タグ
Fig.1 Wireless Ultrasonic 3D Tag



図2 超音波3次元タグシステムを実装したセンサルーム
Fig.2 Sensor Room of Ultrasonic 3D Tag System

3.1.2 ウルトラビジョンシステム

カメラのステレオシステムを用いて、ステレオ画像における対象物の特徴点をマウスにより対話的に指定していくことで、対象物の重要な性質である3次元形状を簡易な形状にモデル化を行い、「物理構造のモデル化手法」を実現した。

しかし、ステレオカメラシステムを用いて対象物の3次元モデルを作成する際、両眼の視野内に対象物が入っていないと、広い部屋をモデル化することは非常に困難である、という問題点があった。そこで、本研究では、超音波3次元タグとステレオカメラを組み合わせることで、ステレオカメラを移動させて、3次元モデルを作成しても、一つの座標系上にモデルを作成することができるウルトラビジョンシステムを開発した。図3に開発したウルトラビジョンシステムを示す。二眼のステレオカメラと6個の超音波タグから構成されている。



図3 ウルトラビジョンシステム
Fig. 3 UltraVision System

実際の手順は、1) ステレオカメラのキャリブレーション、2) キャリブレーションした位置でウルトラビジョンに取り付けられた最低3個の超音波タグの3次元位置を登録、3) ウルトラビジョンを移動させた場合、その位置における超音波タグのそれぞれの3次元位置を取得し、前回の位置から現在の位置への移動を回転行列と並進ベクトルの形で算出する(図4)、4) 求められた回転行列と並進ベクトルを用いて、3次元モデルを一つの座標系に統一させる。

具体的な変換手順を以下に述べる(記号は図4中のものである)。地点2に置かれたカメラ座標系(camera2)から地点1に置かれたカメラ座標系(camera1)へ変換する手順について述べる。式で書くと次のようになる。

$$P_{c1} = M_{c1c2} \cdot P_{c2}$$

また、

$$M_{c1c2} = M_{c2u2}^{-1} \cdot M_{u2u1} \cdot M_{c1u1}$$

という関係がある。まず、ウルトラビジョン上の超音波タグ最低3個で座標系を作る(ultrasonic1, ultrasonic2)。この座標系とステレオカメラの座標系の関係は、キャリブレーション後は変化しないので、

$$M_{c1u1} = M_{c2u2} = M_{cu} \text{ と書け, } M_{cu} \text{ とおく。また, } M_{cu} \text{ は}$$

$$M_{cu} = M_{wu1} \cdot M_{c1w} = M_{wu2} \cdot M_{c2w} \text{ と書ける。}$$

次に、ultrasonic2 から ultrasonic1 へ変換する行列は次のように書くことができる。

$$M_{u2u1} = M_{wu1} \cdot M_{wu2}^{-1}$$

以上より、

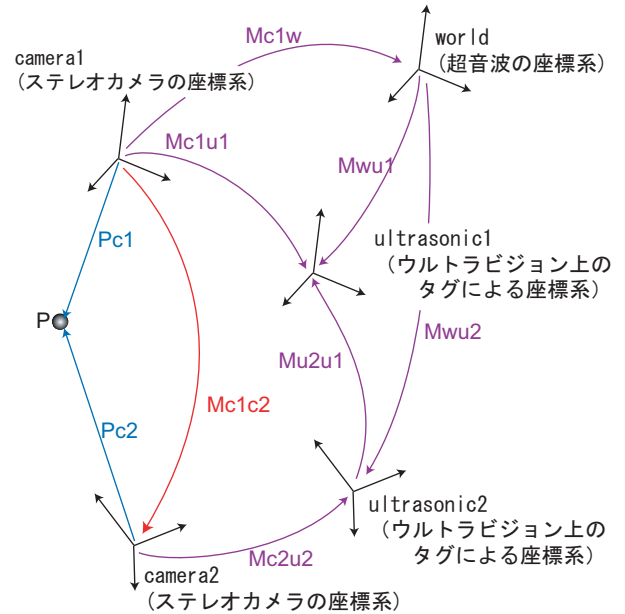


図4 ウルトラビジョンにおける座標変換
Fig. 4 Coordinate Conversion in UltraVision System

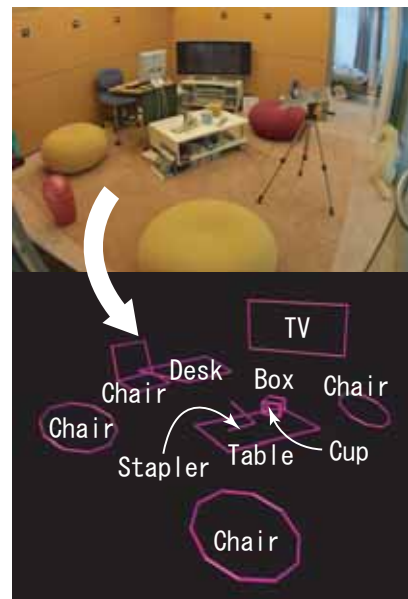


図5 ウルトラビジョンにより作成したモデル
Fig. 5 Example of Virtualized Objects

$$\begin{aligned} P_{c1} &= M_{c1c2} \cdot P_{c2} \\ &= M_{c2u2}^{-1} \cdot M_{u2u1} \cdot M_{c1u1} \cdot P_{c2} \\ &= (M_{cu}^{-1} \cdot M_{wu1} \cdot M_{wu2}^{-1} \cdot M_{cu}) \cdot P_{c2} \end{aligned}$$

この変換手順をもとに、ウルトラビジョンを用いて、実際にセンサールームのモデル化を行った様子を図5に示す。

3.2 簡便な環境モデリングと行動イベントの検出システムのソフトウェア

(a) 物理構造のモデル化手法の実現

ウルトラビジョンシステムを用いて、図6に示すように、左右のステレオ画像中の対象物の特徴的な対応点をマウスで対話的に指定することにより、簡単に線形・多角形・円形といった簡易な3次元形状にモデル化することが可能である。また、対

象物に取り付けられた超音波式 3 次元タグの個数や ID 情報モデルと結び付けて登録する。図 7 に作成された物理構造モデルの例を示す。

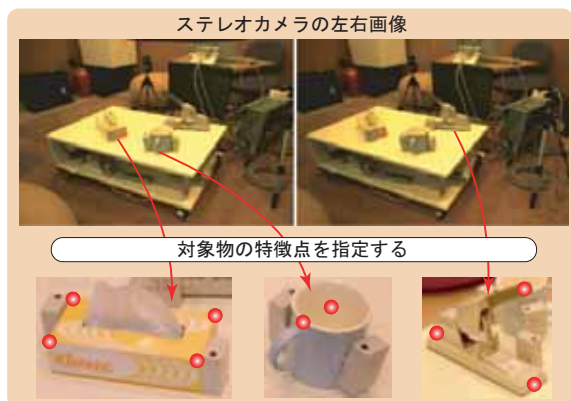


図 6 ステレオカメラによるフォトモデリング
Fig. 6 Photo-modeling by Stereoscopic Camera System

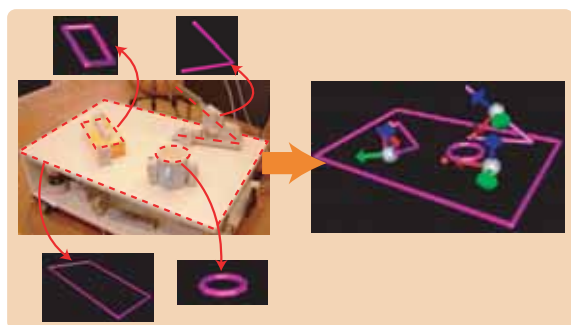


図 7 対象物の物理構造のモデル化
Fig. 7 Virtualizing Real Objects

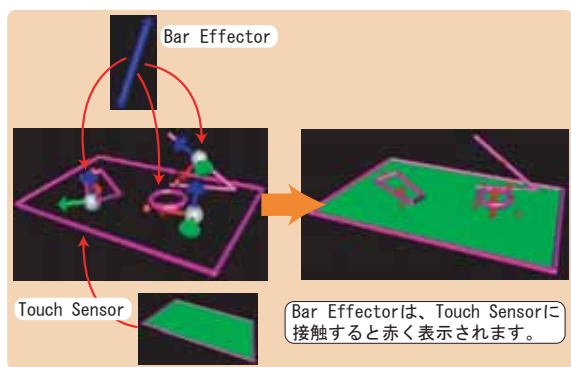


図 8 仮想センサ・効果器による機能モデルの作成
Fig. 8 Creating Functional Model using Virtual Sensors/Effectors

(b) 物理現象のモデル化手法の実現

(a) で作成された物理構造モデルに、本研究で開発された仮想センサ・効果器を取り付けることで、対象物の物理的機能をモデル化する。仮想センサ・効果器とは、計算機上でセンサとして振舞える属性を与えるためのものである。図 8 に示すように、仮想センサを取り付ける物理構造モデルをマウスで選択し、取り付けるべき仮想センサを選択することで、計算機上で

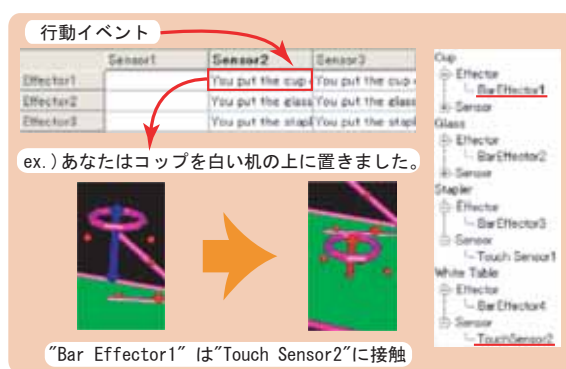


図 9 仮想センサの出力と行動イベントの対応付け
Fig. 9 Associate Output of Virtual Sensors with Target Behavior Event

対話的に取り付け作業を行うことが可能である。現在、開発したシステムで用意されている仮想センサ・効果器の種類は、回転を検出する「Angle Sensor」、接触という事象を起こさせる「Bar Effector」、接触という事象を検出する「Touch Sensor」である。

(c) 物理現象モデルの出力と行動イベントの対応付け手法の実現

図 9 に示すように (b) で作成された物理現象モデルの組み合わせの表に、その現象と対応する行動意味情報を書き込むことで、物理現象モデルの出力と行動イベントの対応付けを行う。例えば、机とコップが接触すれば、「机にコップを置いた」などと書き込み、ノートとペンが接触していれば、「何が書いている」などと書き込む。

(d) 3次元位置計測システムとセンサ化された対象物を用いた行動認識手法の実現

(a) (b) で作成された物理構造モデル・物理現象モデルに、3次元位置計測システムから得られる位置情報を入力し、その結果得られる物理現象モデルの出力をもとに (c) で対応付けた行動イベントを参照することによって人間行動の検出を意味情報として取り出すことが可能である (図 10)。

4. 検証実験

本研究で提案した手法の有効性を検証するために、日常生活空間を模擬したセンサルームにおいて、構築したシステムを用いた検証実験を行った。検証実験の際にモデル化を行った対象物は、机・コップ・ティッシュ箱・ホッチキスの 4 点である。人間が使用する際に動かす可能性の高い、机以外の 3 点には超音波 3 次元タグを取り付けた。4 点の対象物を、ウルトラビジョンシステムにより、机とティッシュ箱を四角形、コップを円形、ホッチキスを線形といった簡易な物理構造にモデル化を行った。次に、机には「Touch Sensor」、コップ・ティッシュ箱・ホッチキスには、「Angle Sensor」と「Bar Effector」といった、仮想センサ・効果器を取り付けることにより、物理現象モデルを作成し、対象物のセンサ化を行った。その後、仮想センサの出力と行動イベントの対応付けを行った。超音波 3 次元タグの 3 次元座標値の入力を開始し、机以外の 3 点を持ち上げたり、回転させることによって、行動イベントが確実に検出できることを

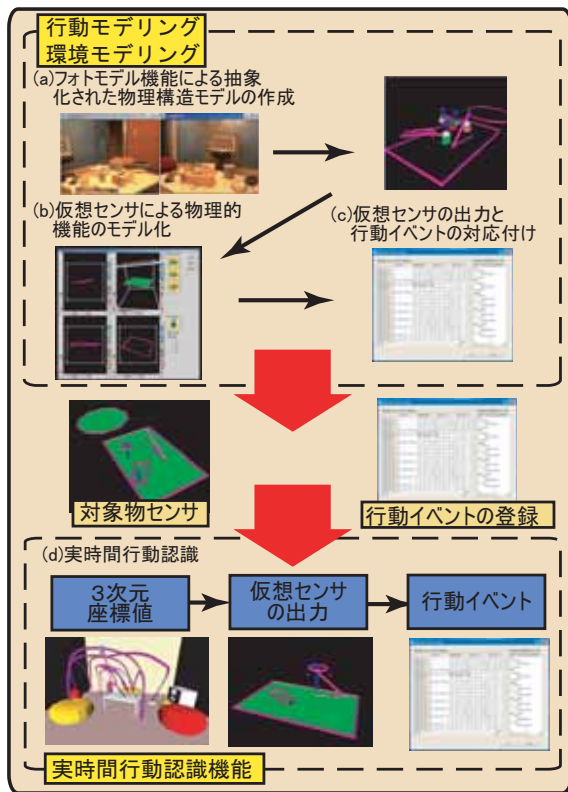


図 10 提案手法のフローチャート

Fig. 10 Flowchart of proposed method

確認した (図 11)。この検証実験により、本研究で提案したシステムは、簡便な環境モデリング機能を実現し、頑健性、実時間性、システム構築の容易性といった問題点の解決も可能であることを確認した。

5. 結 論

本研究では、人間の行動が生じる現場で、行動観察システムを手早く実現する手法を提案した。提案手法は、1) 対象物の物理構造のモデル化機能、2) 対象物の物理現象のモデル化機能、3) 物理現象モデルの出力と行動イベントの対応付け機能、4) 対象物の認識・位置計測機能から構成される。1) ~ 3) の機能を統合することで、対象物の性質や機能を維持しつつ、対象物を人間の行動を認識するためのセンサとして振舞えるように、対象物のセンサ化を容易に行えた。また、対象物センサに4)の機能を統合することで、仮想センサの出力を得られ、人間の行動を観察・認識することができる。

提案した手法を、超音波3次元タグとステレオカメラを用いて構築した。対象物の抽象化された物理構造モデルの作成機能は、超音波3次元タグとステレオカメラを統合したウルトラビジョンで実現した。次に、物理現象モデルの作成機能は、本研究で開発した仮想センサ・効果器を物理構造モデルに取り付けることで実現した。行動イベントの検出機能は、仮想センサの出力と行動を表す文章と対応付けることで実現した。最後に、対象物の認識・位置計測機能は、超音波式3次元タグがIDを持つことができるので、そのIDと対象物の名称を結びつけることによって、対象物の認識機能を実現し、そのIDを持った超音波3次元タグの位置を計測することで、対象物の位置計測

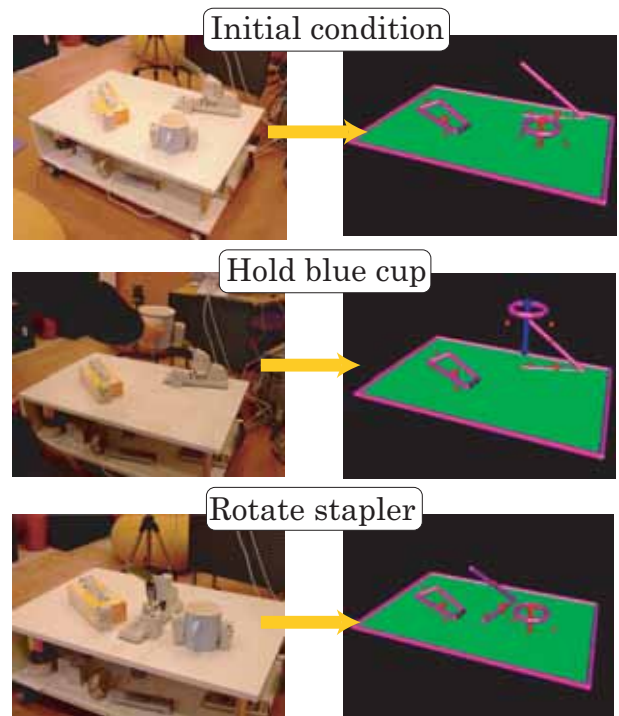


図 11 行動検出の様子

Fig. 11 Behavior Detection

機能を実現した。このようにして、簡便な環境モデリング機能と行動イベントの検出機能を実現した。

また、日常生活空間を模擬したセンサルームを構築し、提案手法を用いて、実際に環境のモデリングを行い、センサルーム内でモデリングを行った対象物を使って、人間が行動をしてみることににより、環境モデリングが容易であることと行動イベントを確実に検出できることを検証した。

文 献

- [1] 独立行政法人産業技術総合研究所編, "デジタル・サイバー・リアル 人間中心の情報技術,"丸善, 2002
- [2] H. Noguchi, T. Mori and T. Sato, "Network Middleware For Utilization of Sensors in Room," IEEE International Conference on Intelligent Robots and Systems (IROS2003), pp. 1832-1838, 2003
- [3] 野口 勝則, ボンサック ソンウォン, 松原 隆, 中内 靖, 柏木 英一, "センサ・ルームにおける人の行動認識に関する研究,"日本機械学会 ロボティクス・メカトロニクス講演会'03 講演論文集, 1P1-3F-E7, 2003
- [4] Y. Nishida, H. Aizawa, T. Hori, N.H. Hoffman, T. Kanade, M. Kakikura, "3D Ultrasonic Tagging System for Observing Human Activity," IEEE International Conference on Intelligent Robots and Systems (IROS2003), pp. 785-791, 2003
- [5] T. Hori, Y. Nishida, T. Kanade, K. Akiyama, "Improving Sampling Rate with Multiplexed Ultrasonic Emitters," IEEE International Conference on Systems, Man & Cybernetics (SMC2003), pp. 4522-4527, 2003
- [6] 西田 佳史, 西谷 哲史, 相澤 洋志, 堀 俊夫, 溝口 博, "ポータブルな超音波3次元タグ-簡便なキャリブレーション手法-", "日本ロボット学会第21回学術講演会, 1A15, 2003
- [7] A. Nakamura, T. Ogasawara, K. Kitagaki, T. Suehiro, "Using Robust and Simplified Geometric Models in Skill-Based Manipulation," IEEE International Conference on Intelligent Robots and Systems (IROS2001), pp. 138-145, 2001