

日常空間で対話できるコミュニケーションロボット Robovie-IV

光永法明[†] 宮下敬宏[†] 吉川雄一郎[†] 石黒浩^{††,†}
小暮潔^{†††} 萩田紀博[†]

[†] (株) 国際電気通信基礎技術研究所 知能ロボティクス研究所

^{†††} (株) 国際電気通信基礎技術研究所 メディア情報科学研究所

〒619-0288 「けいはんな学研都市」光台 2-2-2

^{††} 大阪大学大学院工学研究科 知能・機能創成工学専攻 〒565-0871 吹田市山田丘 2-1

E-mail: †{mitunaga,miyasita,y-yoshi,kogure,hagita}@atr.jp, ††ishiguro@ams.eng.osaka-u.ac.jp

あらまし コンピュータやセンサが環境中に埋め込まれたコピキタス環境において、人と情報の接点として人サイズのヒューマノイドロボットは重要な役割を果たすことになると期待される。そのためにはヒューマノイドロボットには日常空間で音声発話・認識ができるだけでなく存在感をもつことが重要である。我々は存在感を持つ日常空間で対話できるコミュニケーションロボットとして Robovie-IV を開発した。本論文では対話ロボットに求められる機能について議論し、開発した Robovie-IV のハードウェアとソフトウェアについて説明する。そして現在我々が取り組んでいる日常のオフィス環境下での対話実験について紹介、現状について報告し、最後に結論と今後の課題を述べる。キーワード コミュニケーションロボット、対話、ヒューマノイド、日常活動型

Robovie-IV: An Every Day Communication Robot

Noriaki MITSUNAGA[†], Takahiro MIYASHITA[†], Yuichiro YOSHIKAWA[†],

Hiroshi ISHIGURO^{††,†}, Kiyoshi KOGURE^{†††}, and Norihiro HAGITA[†]

[†] ATR Intelligent Robotics and Communication Laboratories

^{†††} ATR Media Information Science Laboratories

2-2-2 Hikaridai Keihanna-gakken-toshi, Kyoto, 619-0288

^{††} Dept. of Adaptive Machine Systems, Graduate School of Engineering, Osaka University

2-1 Yamadaoka Suita City, Osaka, 565-0871

E-mail: †{mitunaga,miyasita,y-yoshi,kogure,hagita}@atr.jp, ††ishiguro@ams.eng.osaka-u.ac.jp

Abstract People establish and enhance their social relationships through sharing experiences with others. Among media used to promote such experiences, a human-size humanoid is the most promising medium. In addition to capabilities of other media, such as the ability to collect and provide information from the Internet and via ubiquitous sensors connected through a network, a human-size humanoid has the potential to help people share experiences or information among them owing to its shape. It enables people to accept a sense of humanlike expressions or emotions, wishes, and so forth. We have recently developed a human-size humanoid, called "Robovie-IV," as a ubiquitous medium to enhance co-experience. This paper discusses the design requirements of Robovie-IV and introduces an overview of its hardware and software architectures. There is also a brief description and preliminary results of an experiment that we are currently conducting with Robovie-IV, in addition to conclusions.

Key words communication robot, dialogue, humanoid, everyday life

1. はじめに

コンピュータやセンサが環境中に埋め込まれたコピキタス環

境において、人と情報の接点として人サイズのヒューマノイドロボットは重要な役割を果たすことになると期待される。ヒューマノイドロボットは人に似た外観を持つことから、人のような

表現を示すことに違和感が少ないメディアであると考えられる。そのようなメディアを用い、人と情報、人と人をつないでいくことにより、日常生活がより豊かになると考えられる。一方で人と似た外観を生かすためにはヒューマノイドロボットには日常空間で音声発話・認識ができるだけでなく存在感をもつことが重要であると考えられる。我々は存在感を持つ日常空間で対話できるコミュニケーションロボットを目指して Robovie-IV を開発した。

Robovie-IV を開発するに当たり、我々のこれまでのコミュニケーション研究のためのプラットフォーム Robovie シリーズ開発の過程で得られた知見をとり入れた。以下の章では日常空間で人と対話するロボットに必要な機能について議論し、開発したハードウェアと基本ソフトウェア構造を紹介する。そして Robovie-IV を用いて現在行なっている日常のオフィス内での実験を紹介し、結論と今後の課題を述べる。

2. 日常空間で人と対話するロボット

日常空間で人と対話するロボットには少なくとも以下の機能が求められると考える。第 1 には人に対して威圧感を与えない大きさや外観をもち、自立して動作可能なことである。制御コンピュータやデータベース、センサ類をコミュニケーション部と分離して外部へ設置が必要な場合もあるが、対話ロボットの外観としては自立して見せることが重要である。対人距離などをロボットが調整するには、電源等を内蔵しワイヤレスで自立動作する必要がある。そこで Robovie-IV は小学校低学年の児童程度の身長とし、軟らかさを感じる外観とした。

第 2 には触覚コミュニケーション能力である。触覚コミュニケーションは、音声やジェスチャのような聴覚、視覚を通してのコミュニケーションと同様に重要である。子供の頭をなでたり抱きしめることや、相手に気づかせるために軽く肩をたたくなど、ごく日常に見られる触覚コミュニケーションである。こういった所作にロボットが的確に反応するには全身触覚センサが必要である。我々はこれまでに触覚コミュニケーションの研究のため、軟らかい皮膚内に触覚センサを持つヒューマノイドロボット Robovie-IIS, Robovie-IIF を開発した [1]。これらのロボットは Robovie-II [2] をベースにしている。Robovie-IIS と Robovie-IIF と人とのコミュニケーションの様子を図 1 に示す。Robovie-IV は、これらのロボットに用いた技術を基にした皮膚センサをもつ。

第 3 は無意識な動きを生成可能な移動機構である。人の動きは意識的なものと無意識なものに分けられる [3] ~ [5]。意識的な動きとはタスクや意図を実現するための動き、たとえば目的地への移動、握手のための動作、対話のための発話などである。一方、無意識な動きとは外乱などにより意図せず表出される動きである。たとえば人は直立時に完全な静止はできず、全身に意図しない揺れを生じる。無意識な動きはタスク実現に直接関わらないが、見た目の生物感に関連すると考えられている。無意識な動きを実現するため Robovie-III では倒立振り機構を移動機構として採用した [6]。倒立振りでは常に姿勢と位置のフィードバック制御を常に行なう。そのため外乱等により無意

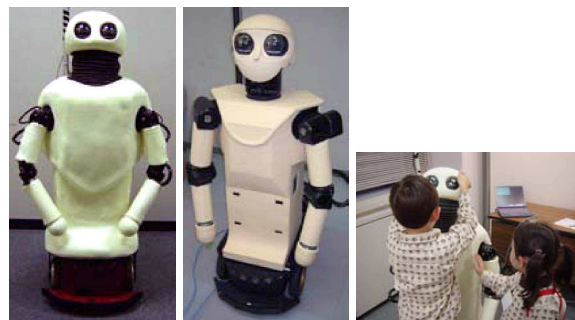


図 1 From left to right, Robovie-IIS, Robovie-IIF, and haptic interaction between children and Robovie-IIS.

識な揺れが表出される。また倒立振り機構では外力を姿勢変化として検出することが容易であり、倒立振り制御において位置フィードバックゲインを調整することで、押されたらそのまま下がるような表現も実現が容易である。Robovie-IV にも倒立振り機構を採用した。

第 4 は人認識である。人との対話には人の存在を認識するだけでなく、相手が誰であるかの認識が重要である。人の同定には画像や音声などによる様々な手法が提案されているが、ロボット上のカメラ、マイク、計算資源には制約が多く容易には実現できない。そこで Robovie-IV では、レーザレンジセンサによる人の足の検出と、カメラによる肌色・差分領域検出を合わせた人検出と、光学/RF 式タグによる人同定を用いることとした。

3. Robovie-IV のハードウェア

図 2 に Robovie-IV の外観を示す。Robovie-IV の身長は頭頂部の全方位ミラーを除くと約 1[m] で、Robovie-II/II-S/II-F (約 1.2[m]) よりも低い。図に示すように、上半身には各腕 4 自由あり、頭部はパン・チルト・ロールの 3 自由度ある。移動機構としてはモータで駆動される主輪 2 輪 (直径 200[mm]) と、キャスターによる補助輪 4 輪をもつ。また駆動輪と本体 (補助輪を含む) はジャッキにより接続されており、ジャッキの上下により 6 輪モードと倒立振りモードを稼働中に切替えることが出来る。6 輪移動モードと倒立振りモードの切替えのための構造を図 3 に示す。補助輪は本体に固定されており、駆動輪と本体の間をジャッキで接続している。倒立振りモードではジャッキにより本体を約 50[mm] 持ち上げる。倒立振り時の制御は河と油田 [7] の手法をベースにしている。

頭部にはパン・チルト軸つきのカメラ 2 台と、全方位ミラーつきのカメラ、マイク、スピーカを備える。胸部には RF 式と光学式タグリーダを備えることが出来る。光学式のタグリーダとしては ATR にて体験共有システム [8] として開発したものを、RF 式としては RF Code Inc. [11] のものを採用している。図では光学式タグリーダを外した状態を示している。胴体部には倒立振りモードでの姿勢制御用のジャイロセンサを 2 台搭載している。2 台のジャイロは共に胴体のピッチ軸周りの角速度を測り故障に備えている。また走行面から 150[mm] 程度の高

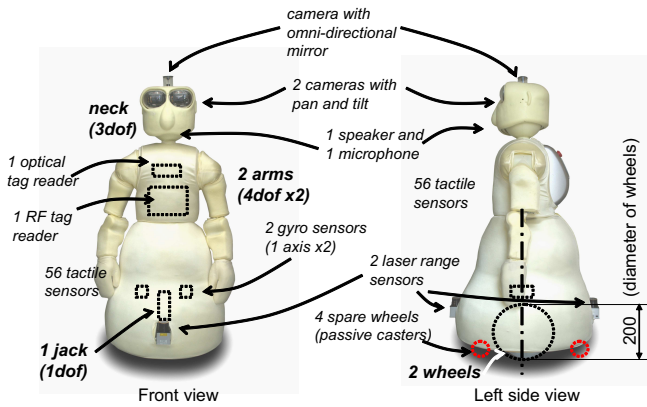


図 2 Front and left-side views of Robovie-IV. Fitted actuators and sensors are shown.

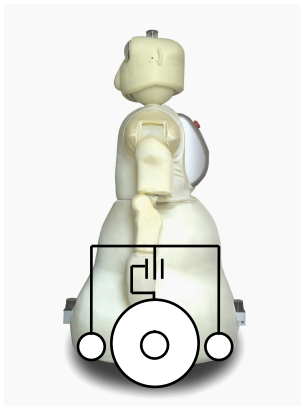


図 3 The powered wheels are connected by a jack to the main body and spare wheels. The figure shows the inverted pendulum mode.

さを水平に測定できるレーザレンジセンサを前後に備える。
全身触覚センサの配置を図 4 に示す。柔らかい皮膚の内部に 56 枚の触覚センサエレメントを埋め込んでいる。指先、鼻、耳のそばなどの人が触りやすい突起部分にもセンサがある。

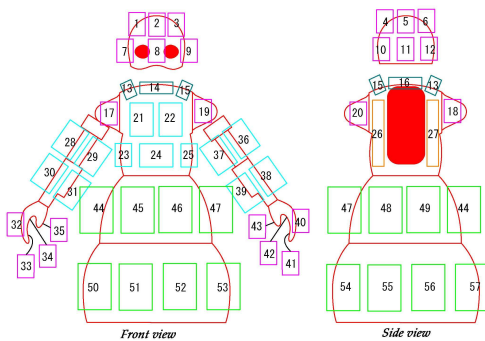


図 4 The arrangement of the skin sensor elements. There are 56 sensors in its soft skin (#16 is not implemented).

Robovie-IV のハードウェア構成を図 5 に示す。車輪、左右の腕、首関節それぞれを PID 制御するサブボードが PC と RS-232C を通して接続されている。車輪制御ボードは移動モードを切替えるジャッキと倒立振り時の姿勢制御も行なう。レー

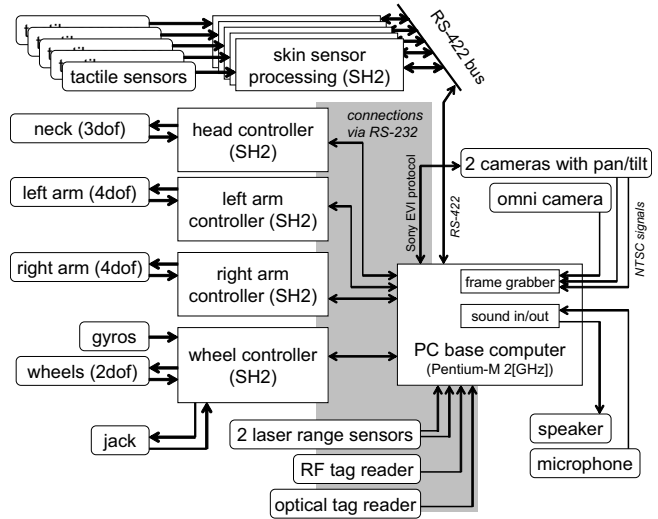


図 5 Robovie-IV's hardware architecture. There are four motor controllers, five skin processors, and one main computer.

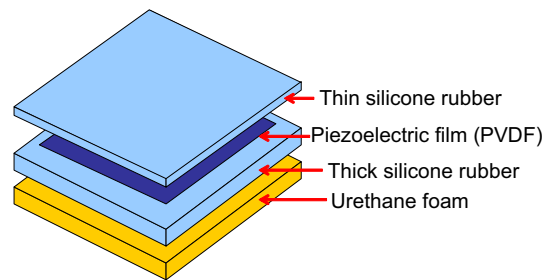


図 6 The skin's structure, which consists of four layers. The piezoelectric sensor is inserted between the outer and middle layers.

ザレンジセンサ、タグリーダー、カメラも RS-232C を通して PC と接続している。触覚センサからの信号は、5 枚の信号処理ボードに部位毎に入力し、処理結果を RS-422 シリアルバスを通して PC へ送っている。各サブボードはルネサスの SH-2 マイコンを用いており、PC の CPU は Pentium-M 2GHz、OS は Linux である。PC は無線 LAN を通して外部へ接続が可能である。全方位ミラーつきカメラを含む 3 台のカメラの画像は、PC 上のビデオキャプチャカードで入力している。マイク及びスピーカーは、PC 上のオンボードサウンド機能を利用している。

触覚センサを埋め込んだ柔らかい皮膚の構造を図 6 に示す。皮膚は表面から薄いシリコンゴム層、PVDF フィルム層 (センサ層)、厚いシリコンゴム層、発泡ウレタン層となっている。シリコンゴムの弾力により触った際に人の皮膚に似た弾力が感じられる。シリコンゴムの密度は約 $1.1[\text{g}/\text{cm}^3]$ 、発泡ウレタンの密度は約 $0.03[\text{g}/\text{cm}^3]$ である。PVDF フィルムは、ピエゾ効果により歪みの変化に対して高い電圧を発生する素子である。PVDF フィルム下に柔らかいシリコンゴムと発泡ウレタン層を用意することで、接触時のフィルムの歪み変化が大きくなり触運動に対する感度が高くなっている。

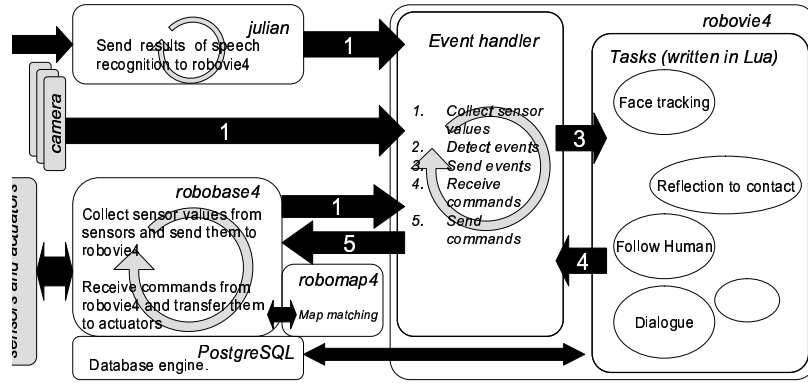


図 7 Robovie-IV's software architecture. The PC's OS is Linux. There are six processes robovie4, robobase4, robomap4, PostgreSQL, and julian.

4. Robovie-IV の基本ソフトウェア

Robovie-IV のソフトウェアアーキテクチャを図 7 に示す。OS として Linux を採用している。図のグレイの四角はセンサやアクチュエータなどのハードウェアを、白い四角は Linux 上のプロセスを表している。主として 5 つのプロセス *robovie4*, *robobase4*, *robomap4*, *PostgreSQL*, *julian* が Robovie-IV を制御している。プロセス間は FIFO, ソケットで通信している。*robovie4* が他のプロセスからのセンサ情報を統合して行動決定する。関節・車輪制御ボード, 皮膚センサボード, レーザレンジセンサ, タグリーダとの通信は *robobase4* が行ない, 統一したプロトコルで *robovie4* へ情報を渡す。*robomap4* は, 車輪の回転量とレーザレンジセンサの情報を用いてマップマッチングにより自己位置を計算する。観測や行動などの体験共有に関わるデータベースとしては PostgreSQL [13] を使用している。音声認識エンジンとしては Julius [9] をベースとする Julian を採用している。

4.1 皮膚センサ信号処理

代表的な 4 つの触行動に対する皮膚センサからの出力を図 8 に示す。「no touch」「still contact」「hit」「stroke」は、それぞれ接触がない場合、そっと手のひらを皮膚上においた場合、たいた場合、なでた場合のセンサ出力を表す。図から分かるように 4 つの触行動は異なるセンサ出力を導く。「no touch」「still contact」は平均センサ出力が小さく、「hit」「stroke」は大きい。「still contact」では「no touch」で見られなかった商用電源に起因するノイズ(この場合 60[Hz]) が表れている。一方「hit」は「stroke」に対して鋭いパルスとなっている。これは「たたく」触行動がインパルス的な歪み変化をもたらしているためである。そのためセンサ出力は「stroke」と比較して高周波成分を多く含む。そこで皮膚センサ処理ボード上に触行動判別プログラムを用意し、処理結果を PC へ伝えることにした。皮膚センサ処理ボードでは、200[Hz] でセンサ出力を読み、図 9 に示す判別法で触行動を判別する。まずセンサ出力の大きさにより「no touch」「still contact」であるか、「hit」「stroke」であるかを判別する。センサ出力が小さい場合には、センサ出力と 60[Hz] の正弦波との相関の大きさにより「no touch」と「still

contact」を判別する。センサ出力が大きい場合には 5 次のバンドパスフィルタ (カットオフ周波数は 70[Hz] と 100[Hz]) の出力により、「hit」と「stroke」を判別する。処理結果は RS-422 バスを通して PC から定期的にポーリングし読みとる。

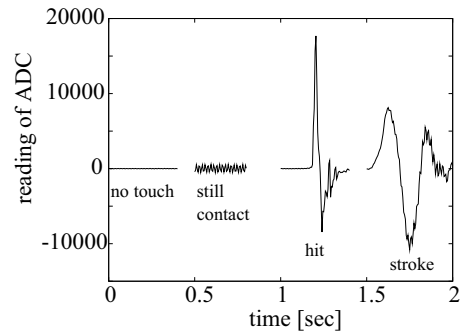


図 8 The raw data from a tactile sensor for four kinds of touch behaviors by a human.

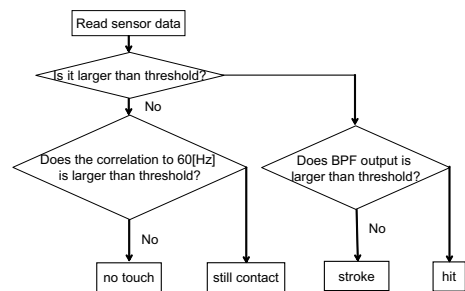


図 9 The algorithm to classify the tactile sensor signal.

4.2 行動決定部のプログラミング

robovie4 プロセスは, C++ で書かれた *robobase4* などとの通信, イベントハンドラ, 画像処理部と Lua 言語 [12] で書かれたタスク部からなる。Lua は C, C++ のプログラムに埋め込みやすさを重視したスクリプト言語である。行動決定部全体を C++ で記述することも可能であるが, インタラクティブな開発を可能にすることで開発効率を高めるためスクリプト言語による記述が可能な構造とした。また Lua の協調的マルチスレッド機能を利用することで複数タスクは並列に動作する。

```

function task1(co)
  local self = TaskTable("task1", co, task1)
  Register(self)

  while (1) do
    CatchEvent("R4Touch", self)
    SleepTask()

    print("Ouch");
  end
end

```

図 10 An example task written in Lua. It repeatedly and infinitely 1) waits for touch events; 2) prints out “touch.”

イベントハンドラでセンサからの主なイベントを検出し、各タスクでイベントに応じた行動を記述する。図 7 に示すように *robovie4* のイベントハンドラは、以下のループを繰り返す。1) *julian*, *robobase4* からセンサ値を受取り、2) イベントを検出し、3) イベント待ちタスクを再開する。そして 4) タスクからアクチュエータへのコマンドを受け取り、5) *robobase4* へ送る。1) から 5) のイベント検出ループは 200[ms] に一度起動している（さらに速い周期での実行も可能である）。また画像処理に関しては、カメラ毎に別スレッドで行ない処理終了後にイベント待ちタスクを再開する (30[Hz] 周期)。

タスク側では 1) イベントハンドラからの起動待ち、2) センサ値を読む (*robovie4* のパツファ内の値)、3) データベースの書き込みや動作の決定を繰り返す。Lua で書いたタスクの例を図 10 に示す。この例では 1) 触覚センサのイベントを待ち、2) 「touch」を表示することを無限に繰り返す。OPEN-R [10] などの他のオブジェクト指向アプローチのようにイベントハンドラを登録してコールバックを待つのではなく、ロボットの振舞いの記述が容易になると考え、この例のように一連のプログラムの中にイベント待ちを記述できるようにしている。

5. オフィス内での日常対話実験

現在 Robovie-IV をオフィス内で毎日動作させ日常空間での対話実験を行なっている。実験環境は著者らの勤務先である ATR 知能ロボティクス研究所内である。研究所内には、研究者、事務職員、インターンの学生など常時 100 名ほどが勤務している。Robovie-IV が動作する対象となる廊下の長さは約 150[m] である。ロボットの身長、音声認識能力などから、対話の内容やしぐさは、親についてきてオフィスにいる子供を想定して作成している。対話実験においてはプログラムを固定せず、改良点が見つければどんどん改良を加えている。図 11 に実験を行なっている研究所内の様子を、図 12 に実験中の対話の様子を示す。

基本的には、人発見をし人から打ち切るまで対話を続けることを繰り返す。人発見は次のように行なっている。まずレーザーレンジセンサによる足首の高さの距離画像から背景差分により



図 11 Robovie-IV is moving in the office in ATR IRC.



図 12 A scene from the experiment. A person is communicating with Robovie-IV. The robot is asking if the person watches a baseball game.

足候補を抽出する。図 13 にレーザーレンジセンサによる距離画像を地図データ上にプロットしたものを示す。右側の円と直線が Robovie-IV のマップマッチング後の位置と姿勢を表す。距離画像上に壁や、段ボールなどの障害物が検出されていることが分かる。センサの最大測定距離は約 4[m] であり、遠くの壁等は検出されていない。この図ではロボットの正面に人の足が 2 つ検出されている (図中ではロボット正面の 2 つの小円)。ロボット左側の円は足として誤検出した障害物である。足候補の抽出後、そのうちの 1 つの方向にカメラを向け肌色領域または差分領域が十分大きければ、足候補を人と判断する。またタグリーダにタグの反応がある場合も人とする。対話が始めれば、最短の足候補の方向の肌色領域をカメラでトラッキングすると共に、車輪で移動しながら足候補をトラッキングする。一定時間、足候補が発見できなければ、対話相手を探しに移動を開始する。無線タグリーダは、ロボットから約 1.5[m] 以内程度で反応するようにアッテネータを調整している。

人発見後、Robovie-IV から対話を試みる。声をかけ一定時間内にならぬ音声入力があれば対話相手とする。音声入力がない場合には対話をあきらめる。各対話は主として、Robovie-IV からの発話内容およびジェスチャ、人から期待される返答を予測して記述した Julian の音声認識用辞書、Julian の出力に応じた次の対話への分岐記述からなる。現在は試行錯誤で対話を用意しているため、プログラムの形で対話を用意し、音声認識用辞書もあらかじめ用意している。対話ルーチンは Lua で記述し長さは 5000 行を越え、対話の種類は 50 種類を越えてい

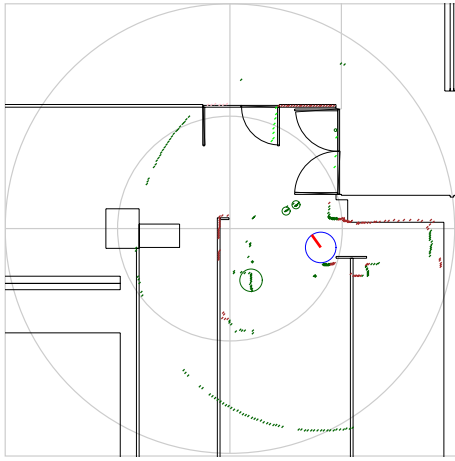


図 13 A depth image plotted on the floor map by a laser range sensor. A circle and a line on the right indicate the Robovie-IV's position and posture. Two circles in front of the robot are detected legs. A larger circle on the left of the robot is a false detection.

る．返答に対する対話の分岐は，内容によるが平均して 3 以上ある．現在のところ対話相手は同時に一人を想定している．

対話に関して，どの対話をしたか，対話の開始と終了時刻，相手の返答，人（足）の移動距離，一日に何度あったか等をデータベースに記録している．また一定時刻毎に，見えた足候補について位置と移動距離，画像上での肌色領域の位置，観測したタグの ID など記録している．現在はデータベースは，同じ対話を繰り返さないようにする，一日に見えた足の数から人の人数が多いか少ないかといった発話を生成するといった目的に利用している．またインターネットを通して研究所付近の天気予報にアクセスし，天気的话题に利用している．

実験環境がオフィス内であり仕事の手を休めることが出来ない人も多く Robovie-IV は声をかけても必ずしも返答をもらえない．また話かけても無視されることもある．一方で Robovie-IV の動作については注目はされており，少しずつ改良していることに気づいている人も多い．まだまだこれから改良の余地は大きいが何か気になる存在として扱われているようである．

6. まとめと今後の課題

本論文では，まず日常環境で対話するロボットに求められる機能について議論し，開発した Robovie-IV のハードウェアとソフトウェアについて紹介した．Robovie-IV は小学校低学年の児童程度の背の高さの完全自立型のロボットであり，全身に軟らかい皮膚を持つ．皮膚内には 56 の触覚センサを，移動機構として倒立振り機構を備える．倒立振り機構により無意識の動きを実現することが可能である．

現在，オフィス内での日常対話実験を継続的に行なっている．前述したように日々改良を加えることでより，違和感なく日常にとけこみ存在感のあるロボットを目指している．現在でも一定の存在感を人に与えられているようである．今後の課題としては対話の充実，さらなるデータベースの有効利用，対話記録の解析などに取り組んでいく予定である．

謝 辞

本研究は独立行政法人情報通信研究機構の委託研究「超高速知能ネットワーク社会に向けた新しいインタラクション・メディアの研究開発」により実施した．

文 献

- [1] Takahiro Miyashita, Taichi Tajika, Hiroshi Ishiguro, Kiyoshi Kogure, and Norihiro Hagita. Haptic communication between humans and robots. In *Proceedings of the 12th International Symposium of Robotics Research*, 2005.
- [2] H. Ishiguro, T. Ono, M. Imai, T. Maeda, T. Kanda, and R. Nakatsu. Robovie: A robot generates episode chains in our daily life. In *Proc. of Int. Symposium on Robotics*, pp. 1356–1361, 2001.
- [3] H. Bruno. Repp.: phase correction in sensorimotor synchronization: nonlinearities in voluntary and involuntary responses to perturbations. *Hum. Mov. Science*, Vol. 211, No. 1, pp. 1–37, 2002.
- [4] Y. Zheng. A neural gait synthesizer for autonomous biped robots. In *Proceedings of the IEEE International Workshop on Intelligent Robots and Systems (IROS90)*, pp. 601–608, 1990.
- [5] C.Wu. Analysis of voluntary movements for robotic control. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA91)*, pp. 326–331, 1991.
- [6] Takahiro Miyashita and Hiroshi Ishiguro. Human-like natural behavior generation based on involuntary motions for humanoid robots. *International Journal of Robotics and Autonomous Systems*, Vol. 48, pp. 203–212, 2004.
- [7] 河潤秀, 油田信一. 屋内環境を自律的にナビゲーションする平行二輪倒立振り型自立移動ロボットの機能分散アーキテクチャによる実現. *日本ロボット学会誌*, Vol. 14, No. 2, pp. 238–248, 1996.
- [8] 角康之, 伊藤禎宣, 松口哲也, Sidney Fels, 間瀬健二. 協調的なインタラクションの記録と解釈. *情報処理学会論文誌*, Vol. 44, No. 11, pp. 2628–2637, 2003.
- [9] A. Lee, T. Kawahara, and K. Shikano. Julius — an open source real-time large vocabulary recognition engine. In *Proceedings of the European Conference on Speech Communication and Technology (EUROSPEECH)*, pp. 1691–1694, 2001.
- [10] Masahiro Fujita. Development of an autonomous quadruped robot for robot entertainment. *Autonomous Robots*, Vol. 5, No. 1, pp. 7–18, 1998.
- [11] RF Code, Inc. <http://www.rfcode.com/>
- [12] The programming language Lua, <http://www.lua.org/>
- [13] PostgreSQL Global Development Group. <http://www.postgresql.org/>