

ロボカップ@ホームにおける ロボット向けベンチマークテスト構築の試み

RoboCup@Home: A Benchmark Test for Domestic Service Robots

杉浦孔明
Komei Sugiura

(独) 情報通信研究機構
National Institute of Information and Communications Technology

1 はじめに

日常環境において、生活支援ロボットが人と対話し行動するには、移動、把持、音声処理、画像処理など広範囲にわたって解くべき技術的課題が多い。一方、日本発の RT ミドルウェアや、Willow Garage が提唱する ROS(Robot Operating System) の登場により、生活支援ロボットの(ソフトウェア)開発コストは急速に減少しており、分野の拡大が期待されている。実際に、ロボカップ@ホーム [1]、つくばチャレンジ [2]、ロボ LDK [3] など、日常環境へのロボットの応用を目指した競技会・イベントは多くの参加者を集めている。

本稿では、日常環境で人間を支援する自律ロボットのベンチマーク競技である、ロボカップ@ホームについて概説する。ロボカップ@ホームは、競技を通じて、生活支援ロボットに必要な技術を競争的に開発することを目的としている。図 1 に、2010 年シンガポールで行われた世界大会の様子を示す。図中の白丸は、探索すべきオブジェクトを示す。2006 年に 11 チームでスタートしたロボカップ@ホーム世界大会は、現在 20 チーム、200 人ほどの参加がある。

ロボカップ@ホームでは、家庭・オフィス・スーパーマーケットなどにおけるロボットの応用を想定したタスクが設定されている。具体的には、日用品の探索、棚からユーザに言われたものを取ってくる、人を追従する等、日常生活に役立つ機能を制限時間内にどれだけ達成できるかを競う。人間がロボットを直接操作することは許可されていないため、高い自律性が要求される。

2 タスク

2.1 タスクの概要

2011 年の世界大会におけるタスクの概要を表 1 に示す。中心課題は、移動マニピュレーションとヒューマン



図 1 2010 年シンガポール大会の競技中の様子

ロボットインタラクションの 2 種類に大別できる。タスクは計 11 種類あり、各タスクには制限時間と配点が設定されている。例えば、Who Is Who タスクでは、音声認識に成功すると 25 ポイント、人検出に成功すると 50 ポイントが与えられる。

一次予選の総合得点により上位半数のチームが二次予選に進み、二次予選の上位 5 チームが決勝に進む。決勝では、各チームが自由デモを行う。デモに対する審査員の評価点と、それまでの得点を合計して最終順位が決定する。

公式ルールの最新版は公式サイト [4] からダウンロードできる。2009 年グラーツ大会のタスクに関しては [5] が詳しい。これまでの世界大会における得点傾向が [6] にまとめられているほか、各チームの獲得スコア情報が各大会のウェブサイト上で公開されている。

2.2 タスク環境

図 2 に 2009 年グラーツ大会のフィールドを示す。競技フィールドは 2LDK 程度のモデルルームで、家具や食器等を配置された、実際の生活環境に近い複雑な環境が用意される。一部のタスクでは、フィールド外で競技を行うこともある。例えば、Shopping Mall タスクでは、ロボットを実際の店舗まで運搬して、競技を行う。

タスクで用いられるオブジェクトや名称は準備日に公開される。オブジェクトは、ペットボトルや菓子などの日用品である。Who Is Who タスクなど、顔と名前を記憶するタスクでは、人名を記憶する必要がある。理想的には、どのような名前でも音声認識できることが望ましいが、あらかじめ音声認識辞書にすべての人名を登録しておくことは現実的でないため、準備日に Michael, Sofia など一般的な英語名のリストが公表される。人名が必要な場合、このリストからランダムに選択される。



図 2 2009 年グラーツ大会のフィールド

表1 ロボカップ@ホーム 2011 イスタンブール大会のタスク

ステージ	タスク名	概要
1	Robot Inspection and Poster Session	競技形式で移動機能・停止スイッチ等の動作を確認する
	Follow Me	未知ユーザに対して1分以内にキャリブレーションを行い、追従する
	Go Get It!	フィールド内にある4個のオブジェクトを探索する
	Who Is Who	フィールド内にいる5人を見分けて、名前と顔を覚える
	General Purpose Service Robot 1	ランダムに生成された複文の指示を実行する
2	Open Challenge	ロボットの性能に関する、研究面で優れている点のデモ
	Enhanced Who Is Who	フィールド内にいる5人を見分けて、飲み物を届ける
	General Purpose Service Robot 2	曖昧な指示を明確化(例:「何の飲み物を持ってきますか」)し、指示を実行する
	Shopping Mall	未知環境(実際の店舗)で棚から3個のオブジェクトを持ってくる
決勝	Demo Challenge	毎年変更されるテーマに沿ったデモ
	Final	自由に選んだテーマに沿ったデモ

2.3 タスクの標準化

ルールは、テクニカルコミッティ(TC)による議論に基づき、毎年更新される。タスクの設計に際しては、生活支援ロボットのベンチマークテストとして標準化することを重視している。生活支援ロボットの研究開発においては、独自の環境や評価尺度が用いられることが多く、手法同士を比較することが難しい。一方、音声認識などの分野では、標準化されたベンチマークテストにより手法の比較が容易になっている。ロボティクスにおいても、タスクを標準化して比較評価のコストを低減すれば、コミュニティ全体の研究開発に貢献できると考えられる。

日常生活環境における自律ロボットのベンチマークテストは、移動マニピュレーションを主眼とした Mobile Manipulation Challenge や、物体探索を主眼とした Semantic Robot Vision Challenge など有名であり、ロボカップ@ホームもこれらに刺激を受けている。ロボカップ@ホームの特徴としては、スコアが厳密に定義されている、各チームのスコアが公開されている、世界最大規模の大会であり比較対象の数が十分である、などが挙げられる。

3 ロボカップ@ホームにおける音声対話機能：未登録語の学習

本節では、タスクを解決するにあたり我々が開発した音声対話機能のうち、未登録語を学習する手法 [7] について概説する。本手法を用いることで、ユーザはオブジェクトを見せて名称を発声するのみで名称(未登録語)と画像情報を教示できるようになり、一般ユーザに対する利便性が向上する。

図3に提案手法の概要を示す。図において、四角形、平行四辺形は、それぞれ処理、データを表す。本手法では、未登録語の登録は「この名前はX」など決められた定型文で行なう。音声認識の結果、入力音声の音素アラインメント情報が得られるので、音声から未登録語部分(図3の「OOV」)を切り出す。ここで切り出された部分は、ユーザの声による音声である。そのため、ロボットに「これはXです」のような音声を出力させる場合には、そのままではXの部分のみがユーザの声になり不自然である。そこで、Eigenvoice Gaussian Mixture Model (EGMM) に基づく声質変換 [8] を適用することで、ユーザの声をロボットの声に変換し音声出力する。Comprison Mean Opinion Score(CMOS) [9] により評価を

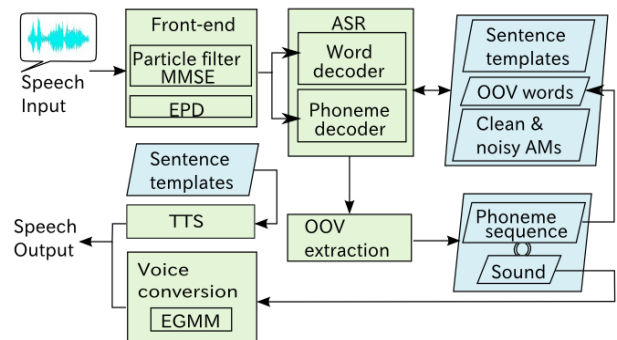


図3 声質変換による未登録語発話手法の概略図
行い、良好な結果を得ている [7]。

4 今後の展望

本稿では、ロボカップ@ホームについて概説した。ルール上の制約により、現状ではネットワーク上のリソースを用いることができないため、出場するロボットはスタンドアロン型のみである。一方、今後のサービス展開を考えると、一部の機能をクラウド化することのメリットは大きい。例えば、音声認識では、高精度な処理を低コストで実現できるうえ、収集されたデータを用いて音響・言語モデルを改善できる。今後、ネットワークの利用に関するルールの改正を議論していきたいと考えている。

参考文献

- [1] 杉浦孔明, “ロボカップ@ホームリーグ; 情報処理, 2012, to appear.
- [2] 油田信一, 水川真, 橋本秀紀, “つくばチャレンジのねらいと成果” 計測と制御, vol.49, no.9, pp.572-578, 2010.
- [3] ロボ LDK 実行委員会(編), ロボットのいる暮らし, 日刊工業新聞社, 2007.
- [4] <http://www.ai.rug.nl/robocupathome/>.
- [5] 岡田浩之, 大森隆司, “ロボカップ@ホーム: 人とロボットの共存を目指して” 人工知能学会誌, vol.25, no.2, pp.229-236, 2010.
- [6] L. Iocchi, and T. van der Zant, “Robocup@home: Adaptive benchmarking of robot bodies and minds,” Proceedings of the International Conference on Simulation, Modeling and Programming for Autonomous Robots, pp.171-182, 2010.
- [7] T. Nakamura, K. Sugiura, T. Nagai, N. Iwahashi, T. Toda, H. Okada, and T. Omori, “Learning novel objects for extended mobile manipulation,” Journal of Intelligent & Robotic Systems, pp.1-18, 2011.
- [8] T. Toda, Y. Ohtani, and K. Shikano, “One-to-many and many-to-one voice conversion based on eigenvoices,” Proceedings of 2007 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007), pp.1249-1252, 2007.
- [9] I.T. Union, “ITU-T P.800,” <http://www.itu.int/rec/T-REC-P.800/en>.