RAWSEEDS: Robotics Advancement through Web-publishing of Sensorial and Elaborated Extensive Data Sets

A. Bonarini*, W. Burgard†, G. Fontana*, M. Matteucci*, D. G. Sorrenti‡, J. D. Tardos§

* Department of Electronics and Information, Politecnico di Milano, I-20133 Milan, Italy
Email: bonarini,matteucci,fontana@elet.polimi.it

† Department of Computer Science, University of Freiburg, D-79110 Freiburg, Germany
Email: burgard@informatik.uni-freiburg.de

‡ Dip. Informatica Sistemistica e Comunicazione, Università degli Studi di Milano - Bicocca, I-20126, Milan, Italy
Email: sorrenti@disco.unimib.it

§ Depto. Informática e Ingeniería de Sistemas, Universidad de Zaragoza, S-50018 Zaragoza, Spain
Email: tardos@unizar.es

***Abstract*— The absence of standard benchmarks is an acknowledged problem in the field of robotics, and is doubly harmful to it. First, it prevents recognition of scientific and technical progress, thus discouraging research and development; second, it prevents new actors (and particularly companies) from entering the robotic sector, as heavy investments are needed to compensate for that absence. The need for benchmarking in advanced robotics embraces a wide range of topics, from, e.g., dexterous manipulation to, e.g., emotional interfaces with humans. The RAWSEEDS project will focus on sensor fusion, localization, mapping and SLAM in autonomous mobile robotics. The project will provide a comprehensive Benchmarking Toolkit, including high-quality multisensorial data sets, well defined Benchmarking Problems (BPs) based on the data sets, state-of-the-art Benchmarking Solutions (BSs) in the form of algorithms, software, methodologies and instruments for the assessment of the BSs.**

I. Introduction

Progress in the field of robotics requires that robots and robotic systems gain the ability to operate with less and less direct human control, without detriment to their performance and, most importantly, to the safety of the people interacting with them. We are convinced that when robots will be able to safely navigate through environments designed for human beings, and to effectively execute tasks in those environments, beside and in collaboration with people, we will witness the birth of a new phase in the industrial development of the world. Just as the gradual transformation of computers from laboratory equipment to everyday, ubiquitous appliances created a gigantic market for Information Technology, when robots will lose their status of costly, complex, unsafe and "dumb" mechanisms, commercial applications for robotics will face an explosive growth. This, in turn, will lead to an enormous surge of interest - and financial resources - towards scientific and applied research in the whole field of robotics.

A key factor for a rapid progress towards this "robotic spread" is a substantial advancement in the performances of robots associated to the concept of autonomy, i.e., the set of abilities needed to perform the following activities without human intervention:

1) perceive information about the environment;
2) extract from that information the elements needed to execute an assigned task;
3) decide which actions are needed to proceed towards the goal of the task;
4) correctly execute the selected actions;
5) manage the gap between the expected effect of the actions on the environment and their real effect, also taking into account the modifications of the environment not associated to the actions of the robot.

Among these, we consider that moving through its environment safely and without collision as well as being able to reach a goal location, is the basic ability that a robot must necessarily possess to autonomously operate. This requires, in particular, that the robot is capable to localize itself in the environment: this is usually done by constructing some form of internal representation of the environment, i.e., a map, and locating the position of the robot and its goal on the map. Any mobile autonomous robot must have the abilities needed to perform activities of mapping and self-localization, or even SLAM (Simultaneous Localization And Mapping), a well known problem in literature [7][8][9]. These abilities are not sufficient to ensure that the robot is also able to execute a task, but they can be thought of as necessary conditions for a mobile robot to be capable of effective autonomous behavior.

Solving the problems of mapping and self-localization is, unfortunately, not an easy task. One of the main problems is the fact that data elaborated by the robot come from sensors affected by imperfections, such as:

- limited spatial range and/or field-of-perception;
- noise;
- sensibility to spurious effects;

- low dynamic range;
- systematic errors or drift effects;
- failures.

These imperfections are very significant for any sensor, even costly state-of-the-art ones, but they become increasingly stringent as the cost of the sensors decrease, thus remarking the need for sophisticated sensor systems in advanced robotics. In the SLAM problem, for example, it is nowadays usual to use Laser Range Finders (LRF) as the main sensing system, also to recover 3D data from the environment (using a 3D LRF [10] system); however, in many mobile robots applications (outdoor navigation, indoor navigation with ramps in the environments, indoor cleaning, indoor navigation with obstacles not perceivable at the LRF height, door opening with door handle, semantic classification of places [11], etc.) an LRF cannot be the main robot sensing system or could integrate other perceiving devices. In many applications a vision-based sensing system (e.g. a trinocular vision system [13] or a correlation-based approach [14]), which could leverage on the huge amount of computer vision algorithms, can be the only complex sensing system of a mobile robot. This is also pushed by economical constraints: *"extensive market analyses show that a complex sensing system for a mobile robot cannot cost more that 10US$, for a consumer-level robot"* [1].

Very sophisticated algorithms are then needed to process sensor output, elaborate it and extract the information needed to solve the mapping and localization problems. These algorithms become much more complex when multiple sensors are used (as is usually done to partially compensate for the intrinsic limitations of each sensor), because they need to include a process of sensor fusion between data coming from different sensors [5] [6]. Sensor fusion is mostly difficult when different kinds of sensors are employed (e.g. cameras and sonars), which is exactly what is generally done to explore different aspects of the environment and to exploit the capabilities of different sensor technologies [12]. Cheap sensors (such as the ones that present and future mass-market robotic applications are forced to employ for cost reasons) have very low performance and so, paradoxically, need the most sophisticated algorithms, as the data they generate must be subject to complex elaboration and interpretation procedures.

The ability to use cheap sensors and nonetheless build high-performance robotic products is absolutely necessary for the diffusion of mass-market robotic applications. However, the use of sophisticated algorithms does not necessarily have a significant impact on the final cost of a robotic product, as the main economic and conceptual effort is required for the development and test phases of the algorithms, while the implementation can be usually rely on inexpensive hardware. As we will show in the following section, presently the tools needed to design and develop such algorithms are not available to the vast majority of the (actually or potentially) interested groups: the objective of RAWSEEDS is to overcome this obstacle by realizing and making freely available such tools.

RAWSEEDS is a project funded by the European Commission as part of the VI EU Framework Program. It will have its official start on November 1st, 2006, and it will end on April 30th, 2009. The authors of this paper are the project's proposers.

II. State of the art

The study, design, engineering and marketing of autonomous robotic systems and solutions relies on the fact that the actors involved (mainly, research groups and companies) possess or can easily acquire the tools to develop and test sophisticated localization, mapping or SLAM algorithms. Such tools can be subdivided into the following categories:

- sensor data sets for the testing of systems on real-world environmental data;
- benchmarks and methodologies for the quantitative evaluation and comparison of algorithms performance;
- proven algorithms, having already demonstrated successful performances, to be used as starting points to develop new solutions and for comparison.

To be fully and readily useful, these elements would need to be integrated into a coherent Benchmarking Toolkit. This in turn requires: common and well-documented interfaces, immediate interoperability, extensive documentation, and accompanying support services.

Presently neither a toolkit of the kind described above nor its constituents are available to the general potential users. Some groups (essentially universities) have made the conspicuous investments needed to create for themselves some of the elements and shared their results with the community [2]. Even in these cases, the produced tools possess very limited performance and/or versatility and the usefulness of the crafted tools is strongly limited by the fact that they do not compare their results with those of others, because these have been obtained using different data sets or methodologies.

In fields not so distant from the topics of interest for roboticists, like computer vision, it is a little bit more common to refer to benchmarking toolkits, usually available on the web, as the reference source of data for the experimental sections of papers describing innovative scientific proposals. Example of such websites are those related to performance evaluation in traffic systems [3] and to 3D reconstruction [4].

Prospective actors (and especially companies), interested in entering the robotics field, find a strong deterrent in the difficulty and cost of acquiring the tools for the algorithm design described above. Presently, the only practical means to access to them (and to the know-how needed to develop new applications) is to set up a heavily funded (and years long) research program, without any possibility to evaluate in advance its eventual economical revenue. If a toolkit for the design, test and evaluation of sophisticated mapping, localization and/or SLAM algorithms would be made available at a low cost, this would in turn facilitate and speed up the activities of present operators in the robotic field, and encourage potential actors to join the drive towards new high-technology robotic research and products. This would then help to create the critical mass of knowledgeable operators and successful applications that will lead, in a short time, to the "robotic spread" that we have prefigured at the start of this Section.

RAWSEEDS will participate to this process by creating such a Benchmarking Toolkit, making it available for free and disseminating it. Moreover RAWSEEDS (and particularly its website, described in the following sections) will propose itself as an instrument for the exchange of scientific results through the whole robotics community.

III. The RAWSEEDS Benchmarking Toolkit

Advancement in any scientific and technical discipline relies on two basic mechanisms: competition between groups and exchange and dissemination of results among the research community. Both require that the results obtained by one group can be quantitatively evaluated by the other, and that the results obtained by the groups can be compared in order to find the best solutions. In the context of RAWSEEDS, the problem is essentially that of evaluating and comparing algorithms, which requires: (i) that the algorithms are applied to the same data and (ii) that an evaluation methodology exists. As we already discussed, even the first of these two conditions is presently very rarely fulfilled, but RAWSEEDS will offer a solution to this problem by providing comprehensive and validated multisensorial data sets. So the second condition, i.e. the availability of tools for the quantitative evaluation of algorithms, must be tackled: and this is what RAWSEEDS will do by creating suitable benchmarks.

In the context of the development of algorithms and software, a benchmark can be usually defined as a standard problem to which any algorithm in the considered class can be applied, together with a set of rules to evaluate the output produced. RAWSEEDS will generate and publish the data sets needed to define problems, and also two categories of structured benchmarks: Benchmark Problems (BPs) and Benchmark Solutions (BSs).

- A Benchmark Problem (BP) is defined as the union of: (i) a detailed and unambiguous description of a task; (ii) an extensive, detailed and validated collection of multisensorial data, gathered through experimental activity, to be used as the input for the execution of the task; (iii) a rating methodology for the evaluation of the results of the task execution. The application of the given methodology to the output of an algorithm or piece of software designed to solve a Benchmark Problem produces a set of scores that can be used to assess the performance of the algorithm or compare it with other algorithms.
- A Benchmark Solution (BS) is defined as the union of: (i) a BP; (ii) the detailed description of an algorithm for the solution of the BP (possibly including the source code of its implementation and/or executable code); (iii) the complete output of the algorithm applied to the BP; (iv) the set of scores of this output, obtained with the methodology specified in the BP.

The complete set of BPs and BSs published by RAWSEEDS is what we made reference to, in this document, as "RAWSEEDS Benchmarking Toolkit". For instance: a Benchmark Problem may be a precise description of the task of extracting a map of an environment composed of line segments from the point-based representation of the environment produced by a laser range scanner, plus the complete scanner data recorded on location, plus the rating methodology to be applied to the results. The union of this BP with an algorithm solving the problem (and possibly a software implementation of it), its results and their rating (obtained with by the given methodology) may then be a BS.

The main use of a BP is to clearly test existing (or in the course of development) algorithms. On the other hand, a BS can be very useful in many ways, as it will be possible to:

- compare the rating of the results obtained by the algorithm included in the BS with the rating obtained by another algorithm applied to the same BP (please remember that the rating methodology is defined by the BP itself, and so can be applied to different BSs);
- use the output of the algorithm included in the BS to get pre-processed input data for higher level algorithms to be tested, such as planners;
- use the algorithm included in the BS as a "building block" to design a complete multi-layer system for the processing of sensor data;
- use the algorithm included in the BS (and, if available, the source code of its implementation) as a source for the design of new, more sophisticated algorithms.

It must be noted that different BSs can be constructed for a single BP, so the number of BPs is not a limiting factor for the number of BSs that can be defined. Additionally, it is important to stress that the ratings of all the BSs based on the same BP, including the ones uploaded by users of the RAWSEEDS website, will be directly comparable. The BSs defined as part of the RAWSEEDS Benchmarking Toolkit will use state-of-the-art, well-proven algorithms that will constitute a corpus of "standard solutions" for the BPs and for similar problems. As we will see in Section 5, we foresee a conspicuous contribution of new BSs (and possibly new BPs) from the users of the RAWSEEDS website.

The sensory data (including vision data) contained in the BPs will be raw, i.e. they will not be subject to the application of any preliminary elaboration and/or compression procedure, to avoid *a priori* choices about what has to be considered "superfluous" or "redundant". Data will be the result of exploration of a few representative environments by mobile test robots, capable of indoor and outdoor activity and fitted with an extensive set of sensor equipment, both of high and low quality, in order to accomodate the needs of both consumer and high-grade developments. This equipment will be chosen both to cover most of the common sensor schemata and quality levels used in autonomous robotics and also to provide the community with the most useful sets of high and low resolution raw data, that may be successively elaborated to extract higher-level environmental features and descriptions. RAWSEEDS plans to mount on the test robots the following sensor systems: 2 x 180 laser range scanners, trinocular and stereoscopic B/W camera systems, omnidirectional catadioptric camera, color cameras, sonars and GPS, in addition to suitable proprioceptive sensors (e.g. odometric systems, gyroscopes, accelerometers). Data acquisition will be performed by on-board PCs fitted with suitable data acquisition cards. If the necessity emerges, other sensors will be added to this list during the project.

For the construction of the BPs, typical instances of different indoor and outdoor environments will be used, in both static (i.e., excluding moving elements such as people) and dynamic conditions. Each sensor data set will be collected moving the test robot through the environment on a complex exploratory path. The path will be covered with a succession of elementary "steps": at the end of each step the raw output of each sensor or set of sensors (including cameras) will be recorded, and all the motions performed by the robot through the step (including accidental ones) will be logged. Each environment will be covered by multiple data sets, generated by performing exploration sessions on different paths with the same test robot; in this way it will be possible to use multiple data sets associated to the same environment to simulate a multi-robot data set. In our view the inclusion of outdoor locations is particularly significant, since many research groups do not own robot platforms capable to navigate through unstructured terrain and thus research results in this field are very scarce, even if many possible scientific and commercial applications can be envisaged.

Simply collecting high-resolution sensorial data is not sufficient to guarantee their precision and consistency, i.e. the fact that the data obtained from different sensor devices are coherent with each other, with the (logged) actions performed by the robot and with the physical environment explored. Moreover, advanced robotics applications require time coherence between different sensor data streams, which usually is neither guaranteed nor verifiable. To overcome these problems RAWSEEDS will include an extensive data validation phase, with the aim of verifying and certifying the consistency of the data produced by each sensor and their coherence with the ground truth. Statistical analysis of the data against the ground truth will be performed during the validation process, and its results (e.g. noise levels and distribution) added to the BPs.

IV. WWW.RAWSEEDS.ORG

A key part of the RAWSEEDS project is the construction of a website (URL: http://www.rawseeds.org). Users of the website (i.e. typically researchers and enterprises) will be able to download the BPs and BSs, as well as all their accessory information, such as: extensive description of all the hardware used to collect the data sets and of its configuration, ground truth, specifications of the data formats used, pictures of the test locations, notes. The site will include an extensive FAQ section, tutorials and instructions, and all the official documents of the RAWSEEDS project. It will also host a moderated forum, where researchers could exchange opinions and results as well as set up collaborations, and a "user section" where it will be possible to upload new BSs associated to the original BPs, or simply the set of results obtained by a novel algorithm applied to one of the BPs (evaluated with the methodology defined by the BP). The first option (upload a complete BS, including the description of the algorithm used to solve the chosen BP, and possibly the source code of the implementation used and/or executable files) will be probably preferred by research groups, willing to share their results and see them acknowledged by the robotic community; the second option (publication of the results obtained, evaluated with the criteria defined by the BPs) is mainly dedicated to enterprises, which would prefer not to share how they can obtain such results. The possibility to upload new BPs will be left open, with the restriction that the environmental data included in the new BPs will have to certify their compliance to the same validation standards used (and documented) for the original RAWSEEDS BPs.

All that will be needed to become a user of the RAWSEEDS website is a free registration. During the registration process the user will be warned that any material submitted for publication will, if approved, become publicly available. The main goal of the RAWSEEDS website is to provide all the spectrum of data, from raw sensor output to the high-level representations obtained from that output by application of suitable (and immediately available as BSs) algorithms. These algorithms will be the ones provided by RAWSEEDS (a selection of the current state of the art) together with the new ones published by the users of the website. In this way, a current or prospective actor in the robotics field will have access not only to the data needed to immediately test its own applications (whichever be their abstraction level) without any need for a costly data-gathering programme, but also to alternative algorithms and results obtained by other groups. For example, a new application implementing symbolic models of environmental features could be tested simply by using as its input one of the maps produced by the published BSs.

As we will later show in more detail, RAWSEEDS will make efforts to promote among its users a general attitude towards the *sharing* of results; we believe, in fact, that it is the most useful for the progress of robotics and, in general, of science and technology.

A. Management of Intellectual Property Rights (IPR)

The upload section of the RAWSEEDS website will be designed to expand the range of possible users as widely as possible: for example, companies are usually very conservative towards the sharing of results with others, and thus need special attentions to be persuaded to do it. On the other side, we will restrict publication to those contributions that are really useful for the community, i.e. it will not be possible to use the site to "advertise" a result without sharing it, at least partially, with the other users. To reach this objective, a carefully tailored IPR policy will be adopted.

The material that will be published by the site can be divided into two broad categories:

- material produced by the RAWSEEDS project itself;
- material voluntarily submitted for publication by the users of the website.

Both will be subject to the same IPR regime, RAWSEEDS will require that any material published on the website complies with the following three requisites:

- R1: its creator chose to make that material publicly available;

- R2: the rights granted by the creator to the users of the material are clearly and explicitly defined;
- R3: the material must qualify as useful and appropriate for publication.

Requisite R1 is automatically satisfied by the fact that it is the free choice of each user of RAWSEEDS if he/she wants to publish any of his/her creations, and which ones.

To meet requisite R2, any contributor to www.rawseeds.org will have to accompany the proposed material with a *license* stating which rights he/she reserves to himself/herself and which are instead granted to the public. So intellectual property of any material published by RAWSEEDS will remain to its creator, who (with the act of submitting the material for publication) will choose to relinquish part of the rights on it to the public. Which rights are actually given to the public is defined by the chosen license. RAWSEEDS will leave the choice of the license to the user, with the only constraint that a copy of the chosen license is sent along with the material submitted for publication.

To meet requisite R3 any contribution will have to be previously examined and approved by the administrators of www.rawseeds.org, prior to publication. They will decide about the publication of each contribution, and their judgement will be based upon the following definition:

Any publishable material is considered *useful and appropriate* for publication on the RAWSEEDS website if all of the following are true:

1) it is related and can help pogress in the field of robotics, expecially regarding Localization, Mapping and SLAM;
2) it is sufficiently detailed to be usable (e.g., the description of an algorithm must be complete enough to allow a reader to implement - that algorithm into a piece of software);
3) it is usable (e.g., executable code is usable only if it is actually working and accompanied by all the information needed to install and configure it);
4) it does not have commercial purposes only (e.g., a company marketing robotic products could publish the description of a product, but that description will need to disclose enough data about the product to be considered a worthwhile contribution to the field in itself rather than a marketing operation).

B. Dissemination actions and feedback

As the results of RAWSEEDS have the objective of being useful tools for all the actors involved into the development of robotics, rather than scientific achievements *per se*, feedback from the robotics community is explicitly sought. In particular, the RAWSEEDS workplan includes an extensive preliminary phase aimed to the precise definition of the Benchmarking Toolkit that the project will later develop, in order to maximise the usefulness of the toolkit: therefore any motivated suggestion about this topic will be welcome, and (within the limits of the project as defined by the contract between the EU and the proposers) we will take due account of all such suggestions during the benchmark definition phase. Any suggestion can be already sent to suggestions@rawseeds.org.

To maximise the impact of RAWSEEDS, suitable dissemination activities have been expressly included in the workplan; they will be publicized as soon as they are organized. For the same reason RAWSEEDS is proposed by a consortium of partners with an acknowledged expertise in the research fields covered by the project, and coming from different countries.

V. Conclusion

RAWSEEDS is a project aimed at overcoming one of the main limits to research and development in robotics, i.e. the lack of comprehensive, validated and publicly available benchmarks. It will reach its objective through the development of a complete Benchmarking Toolkit and the setup of a website (http://www.rawseeds.org). The website will publish not only the RAWSEEDS Benchmarking Toolkit, but also any useful extension of it coming from the scientific and technical community, and all of this material will be freely available. To help the sharing of results and the general progress of the robotic field, the RAWSEEDS website will also act as a point of aggregation and exchange of information, that will be flanked by a suitable program of dissemination activities.

References

[1] Colin Angle, *Invited talk at EURON meeting*, Amsterdam, 2004.

[2] Andrew Howard and Nicholas Roy, *The Robotics Data Set Repository (Radish)*, http://radish.sourceforge.net/, 2003.

[3] Computational Vision Group, *PETS: Performance Evaluation of Tracking and Surveillance*, The University of Reading, http://www.cvg.cs.rdg.ac.uk.

[4] D. Scharstein and R. Szeliski,*Stereo Vision Research Page* http://www.middlebury.edu/stereo, 2002.

[5] Hugh F. Durrant-Whyte,*Integration, coordination and control of multi-sensor robot systems*,Kluwer Academic,1987.

[6] Juan D. Tardós and Jose A. Castellanos,*Mobile robot localization and map building : a multisensor fusion approach*,Kluwer Academic,1999.

[7] S. Thrun and M. Montemerlo and D. Koller and B. Wegbreit and J. Nieto and E. Nebot,*FastSLAM: An Efficient Solution to the Simultaneous Localization And Mapping Problem with Unknown Data Association*,Journal of Machine Learning Research,2004.

[8] J. Folkesson and Henrik I. Christensen,*Robust SLAM*, 5th IFAC Symp. on Intelligent Autonomous Vehicles, 2004.

[9] F. Lu and E. Milios,*Globally consistent range scan alignment for environment mapping*,Autonomous Robots,1997.

[10] Oliver Wulf and Bernardo Wagner and Mohamed Khalaf-Allah,*Using 3D data for Monte Carlo localization in complex indoor environments*,Proc. of ECMR Conf.,2005.

[11] Rottmann, A. and Martínez Mozos, O. and Stachniss, C. and Burgard, W.,*Place Classification of Indoor Environments with Mobile Robots using Boosting*,aaai,Pittsburgh, PA, USA,2005.

[12] J. Castellanos and J. Montiel and J. Neira and J. Tardos,*The SPmap: A probabilistic framework for simultaneous localization and map building*,IEEE Transactions on Robotics and Automation, 15(5):948–953, 1999.

[13] N. Ayache and F. Lustman,*Trinocular Stereo vision for Robotics*, IEEE Trans. on PAMI,Vol.12-1,1991.

[14] Kurt Konolige *Small vision system. hardware and implementation*,isrr,Hayama, Japan,1997.