

Dataset Acquisitions for USAR Environments

François Pomerleau and Benoit Lescot and Francis Colas and Ming Liu and Roland Siegwart

ETH Zurich, Autonomous Systems Lab
Tannenstrasse 3
8092 Zurich, Switzerland

Abstract

Common representation of the environment is usually depicted as a global or local map. To increase robustness of autonomous creation of such a representation, novel datasets are required. We introduce a system to record precise datasets targeting semi-structured environments which is most likely to be encountered in USAR situations. The global positioning, recorded with a theodolite, is precise in the order of millimeters. Extensive field tests resulted in 4 datasets of challenging outdoor and indoor environments.

Introduction

Team-work implies communication with shared references and symbols. The collaboration between robot and human is therefore highly dependent on a common representation of the environment. Part of this representation is a map, either global or local, that can serve both the robot to do its own task and the human to increase his situation awareness, to collaboratively plan and observe the evolution of a situation. An important issue in mapping is therefore the consistency of the map which relies, on a lower level, on the quality of the registration.

Many registration solutions exist in the literature but few of them were evaluated in 3D semi-structured environments. The lack of dataset is mainly due to the fact that none of the available point cloud datasets target those type of environments while, at the same time, providing reasonable ground truth measurements.

In this article, we present datasets that highlight critical situations for registration algorithms mainly targeting dynamic elements and semi-structured elements. We also explain the motivation behind the different location selection and the methodology used to achieved millimeters precision for the ground truth measurements in outdoor settings.

Related Work

The Iterative Closest Point (ICP) is among the most used algorithm concerning point cloud registration. Although it is a simple algorithm, it may often converge a local minima when used in autonomous systems. Some limitations

come from the hypothesis that the scanned environment is mostly planar from which specialized registration error metrics emerge. For example, the one of the earliest ICP paper present a point-to-plane error (Chen and Medioni 1991) while recent work demonstrated plane-to-plane error (Pathak et al. 2010). Forest with dense vegetation or other environments with many small element can hardly be resumed by plane. There is some needs for semi-structured and unstructured datasets to challenge this hypothesis. An other common hypothesis in registration is that the overlap between scans is constant, which can help to reject outlier matches like in (Chetverikov et al. 2002). Scanned volume can change rapidly within an environment and registration solutions should also adapt to those situations. Earlier work also evaluates the robustness of ICP against low constrained environments (Rusinkiewicz and Levoy 2001). This was mainly done in simulation so real word datasets targeting this limitations could bring the analysis farther. An other problem, recently raised in vision registration (Mortensen, Deng, and Shapiro 2005), is the problem of repetitive elements in the environment which required more robust matching algorithms. All those limitations can be hard to achieved within one dataset so we believe that multiple targeted and challenging datasets would ease the evaluation of registration algorithm against challenging environments.

In the currently available datasets, different laser grade can be found. Surveying equipment for architects now includes 3D scanners. The main advertised systems are the Riegl series and the Leica HDS Laser Scanner series. They provide accurate and long range measurements. This comes with the inconvenient of high weight and low refreshing rates. Cyberware systems target object scanners where it is possible to turn around the subject. Those scanners also provide good precision and even globally consistent scans but are hardly usable in large environments. Security systems provided mainly by SICK also use laser range finders but have typically small opening angle, which required some modifications to create 3D point clouds. Among the most popular laser range finders in Robotics, SICK, Hokuyo and Velodyne are the most used now a day. Recently, the Microsoft Kinect was also used as range measurement systems for autonomous systems (Pomerleau et al. 2011).

To enable registration evaluation, a 'ground truth' evaluation of the scanner pose must be provided. Some datasets

provide positioning of the scanner based on registration process using anchor points coupled with global minimization of the alignment. Then, an expert proceeds to a visual assessment of the data. Such global positioning makes the result dependent of an other registration method and it make it harder to evaluate the achieved precision. Precise global positioning can be reached using a fixed base with multiple degrees of freedom arm holding a scanner, but this solution offer a limited motion range. On the other side, GPS and DGPS system can offer large range of motion but are limited to location having clear sky condition. The precision of such systems can be at best in the range of decimeters in good conditions, which limits the evaluation of registration precision to that range.

Several point cloud datasets are available to researchers. The most popular one is the Stanford 3D Scanning Repository (<http://graphics.stanford.edu/data/3Dscanrep/>) from which the *bunny* was reused multiple times. The repository mainly aims at surface reconstruction research. It provides precise ground truth but on small scanned volumes. The Oakland 3D Point Cloud Dataset (http://www.cs.cmu.edu/~vmr/datasets/oakland_3d/cvpr09/doc/) proposes fixed lasers mounted on a car, which moved in an urban environment. The global positioning of the scans was insured using Trimble AgGPS 114 (Munoz et al. 2009). The Kevin Lai’s Urban Scenes 3D Point Cloud Dataset (<http://www.cs.washington.edu/homes/kevinlai/datasets.html>) mainly aims at urban object recognition and proposed point cloud without overlapping and no global positioning information is available (Lai and Fox 2010). The Rawseeds project (<http://www.rawseeds.org>) was created specially to provide common comparison basis for registration algorithms for indoor and outdoor environments. For the outdoor dataset, the global positioning information is provided by using a Trimble 5700 GPS receiver with Zephyr GPS antenna. The 3D scans were achieved with fixed lasers (Hokuyo URG-04LX and Sick LMS200 and LMS291) on a moving platform (Ceriani et al. 2009). In the same order of idea, the Robotic 3D Scan Repository (<http://kos.informatik.uni-osnabrueck.de/3Dscans/>) provides multiple datasets among which precise scans using surveying equipments is present. The global positioning information was realized with reflective marker and registration algorithms evaluated by an expert.

System description

Sensors

The main sensor of the scanner is its laser range finder. We designed the tilting base to handle 4 main laser systems: Hokuyo URG-04LX, Hokuyo UTM-30LX, SICK LMS100 serie and the Microsoft Kinect. Surveying lasers were rejected from the selection mainly because of their weight. They can be used one at the time and metal guides ensure that the laser will always be fixed at the same position. On the left side of the tilting axis, a Point Grey Flea2 color camera is installed to provide texture information and to aug-

ment the reach of recorded datasets to visual mapping. Gravity vector, magnetic north position and GPS information are provided by an Xsens Mtig Inertial Measurement Unit.

The precise control of the motor is ensured by a Maxon Motor EPOS controller. The control system put in place use a dual regulation loop based on 2 encoders. One encoder is located directly to the motor shaft to provide stability while the second is located at the end of the transmission block allowing precise position control. Encoders has respectively 2000 and 48000 tick count. The system is self-powered with two 50Wh lithium-polymer batteries.

Theodolite

To achieve a global positioning method independent of registration algorithm, we used the TS15, a theodolite from Leica Geosystems. Theodolites are used in road and building surveys and the system used ensures 1 mm precision over 1 kilometer distance. A theodolite only measures a position, so 3 measurements are necessary to retrieve the complete pose (translation and orientation). A specialized reflective prism is mounted on a pole, which can be secured at 3 different location on the scan, namely p_0 , p_1 and p_2 (see fig. 1). The pole is higher than the scanner to enhance visibility from the theodolite.

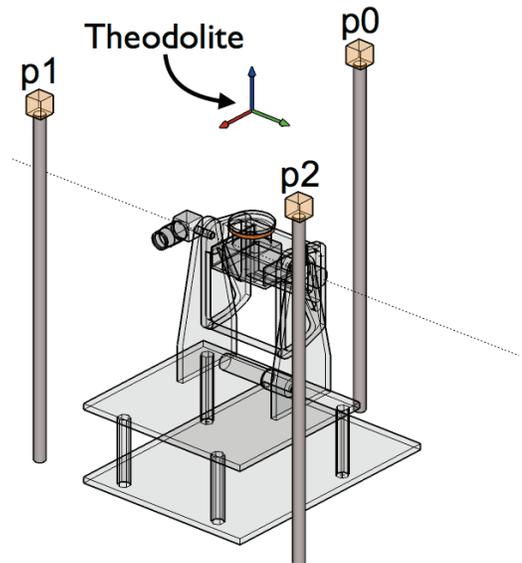


Figure 1: Perspective view of the scanner with positions of the 3 prisms used to reconstruct the global pose of the scanner. The dashed line correspond to the rotation axis.

The translation components are selected to be the center of mass of the 3 points. Defining the vectors $v_{21} = p_2 - p_1$ and $v_{01} = p_0 - p_1$, we express the rotation matrix R , defining the scanner orientation, as followed:

$$\begin{aligned} v_y &= v_{21} \\ v_z &= v_y \times v_{01} \\ v_x &= v_y \times v_z \end{aligned}$$

$$R = [v_x^T \quad v_y^T \quad v_z^T]$$

We evaluated the precision of this positioning method over 181 scanner positions. We moved the scanner over different types of ground. For every pose, we measured the position of p_0 , p_1 and p_2 to evaluate the stability of distances between the 3 prisms (d_{01} , d_{02} and d_{12}). The mean and standard deviation (std) for every segment are presented in Table 1.

Table 1: Precision of the distances between the prisms.

	Mean	Std
d_{01}	534.4 mm	1.4 mm
d_{02}	503.5 mm	1.4 mm
d_{12}	412.5 mm	1.2 mm

Fig.2 shows the estimate Gaussian distributions against the sampled distributions. The std can mainly be due to manipulation errors and can exceed the manufacturer specification. The scanner can move while the prism is installed to another position, the pole supporting the prism can vibrate, the person moving the prism can confuse the sequence of position, etc. Several systems were set in place to minimize the impact of those phenomenas. The weight of the scanner has been increased to augment its inertia. The feet of the scanner can either be 5 cm spikes for soft ground or rubber disks for hard ground. A software monitors the theodolite readings and rejects every sequence of 3 points if the distance variations is larger than 3.5 mm. Moreover, since the mean distance between the prism is significantly different, the expected sequence of measurements can be validated by removing the chances of confusing the order of 2 positions while changing the prism position. All those elements insure a high precision measurement of the scanner pose in a multitude of environments.

Sometimes, a sequence of scanner poses cannot be measured from a single theodolite position. For example, this happens when the scanner needs to turn a corner. To overcome this limitation, we use the scanner as a marker to relocalize the theodolite in its new pose. This increases slightly the localization error due to cumulation but the precision is still in the range of millimeters.

Calibration

As presented in fig. 4, several transformation frames need to be computed in order to achieve reasonable consistency of the measurements. The transformation between the Axis and the Base $T_{B \leftarrow A}$ is insured by the rotational encoder positioned at the end of the transmission system to avoid any dead zone in the reading that could be caused by gear play or flexible transmission chain. This encoder gives us a precision of 0.00013 radian. Moreover, the homing procedure, which transforms the relative measurement of ticks in an absolute reading, is done using one special marked tick of that encoder, which also insures a precision of 0.00013 radian during the calibration.

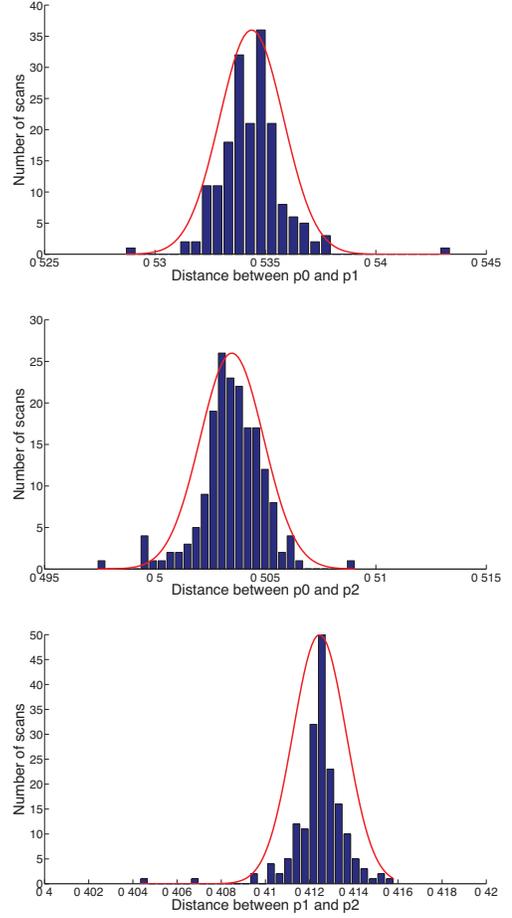


Figure 2: Gaussian estimations (red line) compared to real distributions (blue bars) of measurements obtained with the theodolite. Top: distances between p_0 and p_1 . Middle: distances between p_0 and p_2 . Bottom: distances between p_1 and p_2 . Distances are in meter.

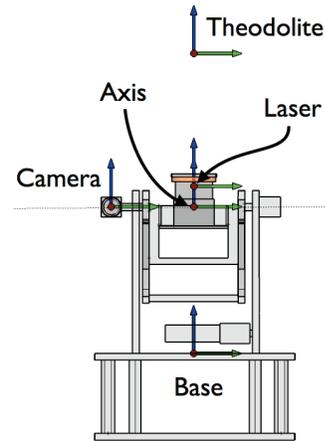


Figure 3: Front view of the scanner with its different reference frames. The dashed line correspond to the rotation axis.

One of the weakest link of the calibration procedure is the transformation between the Laser to the Axis $T_{A \leftarrow L}$. The laser was aligned by construction but the center of the Hokuyo laser is not defined in their documentation as opposed to SICK LSM151. In order to find the height of the beam, we used a camera without the infrared filter, which enables us to detect the laser and measures its position in the range of millimeters. The origin of the 3D point clouds is Base, which is constructed with the transformation chain $T_{B \leftarrow L} = T_{B \leftarrow A} \cdot T_{A \leftarrow L}$ for every 2D scan received.

Our global pose is given with respect to the frame Theodolite while the point cloud are constructed in the reference frame Base. To compute the fixed correction between those two frames, we scanned a corridor room several times and computed the global pose of Base as if the transformation $T_{B \leftarrow T}$ would equal the identity matrix. We then computed the alignment error e_{align} using an Iterative Closest Point Algorithm (Pomerleau et al. 2011). A global minimization algorithm reduces that error, which gives us the transformation between those frames:

$$T_{B \leftarrow T} = \operatorname{argmin}(e_{align})$$

Fig 4 presents the impact of the calibration procedure. The color represents 3D scans of the same corridor with different poses. Given that the calibration environment is controlled, the ICP results ease the evaluation of the calibration but further experiments are needed to assess the precision of this calibration procedure.

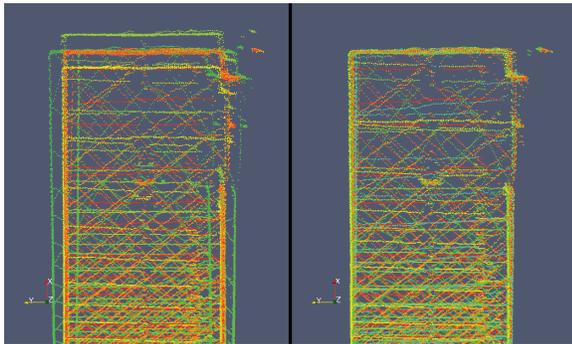


Figure 4: Positioning of the 3D scans. Right: before calibration. Left: after calibration.

Results

We present here the first results recorded with the scanner. Those datasets will be available publicly. The aim is to provide unregistered data for researcher willing to evaluate their registration solutions on a common base. The environments were selected in order to highlight difficult situations where the registration could fail. At this point, we focus mostly on static environments where the configuration is expected to be challenging. The point clouds are provided in local coordinates, which can be compared to the measured global path on processed. We also provide globally consistent point clouds for researchers doing environment modeling.

Our preliminary datasets described below target explicitly: rapid variation of scanning volume, semi-structured environments, unstructured environments and repetitive elements.

Rapid variation of scanning volume

The dataset named *Stairs* aimed at evaluating the robustness of registration algorithms facing rapid variation of scanning volume. This can typically happen when entering in a room from a long corridor. The scanned volume will rapidly drop from the size of the corridor to the size of the doorway, to finally augment to the size of the room. Fig. 5 highlights 3 main zones with different volumes: a large corridor, a small staircase and an open zone with the building facade. Two small doorways separate those zones.

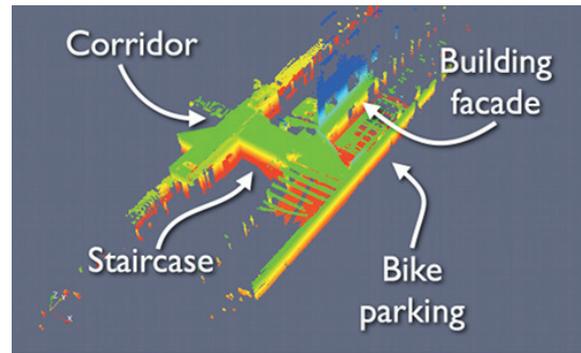


Figure 5: Top view of the dataset with the up left section being an indoor corridor, the middle part a staircase and the bottom left the outdoor facade of the building.

This environment is highly structured and the selected path for the scanner crosses stairs. Fig. 6 presents a cut view of the staircase, which is the zone where the scanner changes progressively its height of about half a floor. The total path length is about 9 m long on which the stairs are climbed. Once outdoor, the path turns at 90 degrees and continues for an other 3 m.

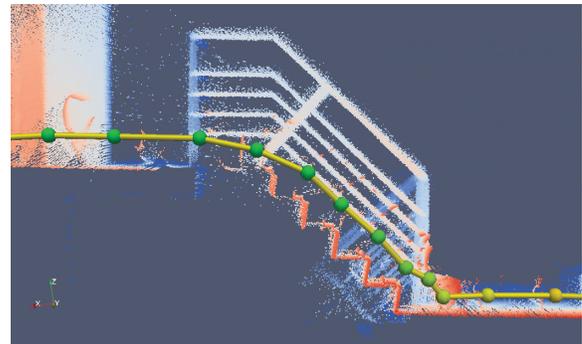


Figure 6: A cut view of the staircase with the scanned path (green balls and yellow tube).

Repetitive elements

The dataset named *ETH* was recorded inside the main building of ETH Zurich. It consists mainly of a straight path following a balcony surrounding the central exhibition hall. The ceiling is curved and, from Fig. 7, we can observe 2 different types of repeating elements: 1) large piers supporting arches, and 2) smaller piers supporting the fence. The realized path is about 25 m long.

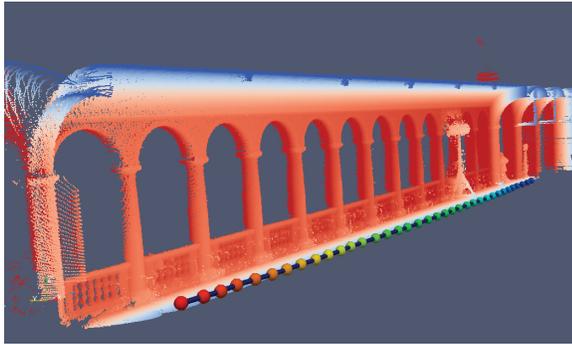


Figure 7: A cut view of the dataset with the realized path (balls with tube) and highlighting the arches and piers.

Semi-structured environment

The dataset named *Gazebo* makes the transition from structured to semi-structured environments. The selected path is a loop around one of the piers supporting a wooden gazebo. Grapevines surround the gazebo under which there is a bench and a paved road. Fig. 8 highlights a part of the dataset where one can see the gazebo, but the complete dataset also includes grass and trees as showed in Fig. 9. The path covers a zone of 4 by 5 m.

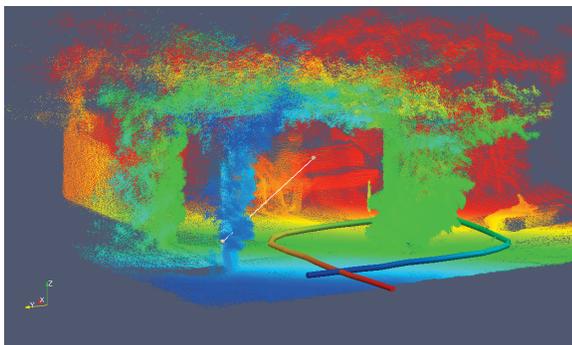


Figure 8: A cut view representing the realized path (tube) under a gazebo.

Unstructured environment

The dataset named *Wood* targets unstructured environments and is mainly constituted of vegetation, from small bushes to big trees. The selected path, presented in Fig. 10, starts under trees that are located on the top left side, and quickly

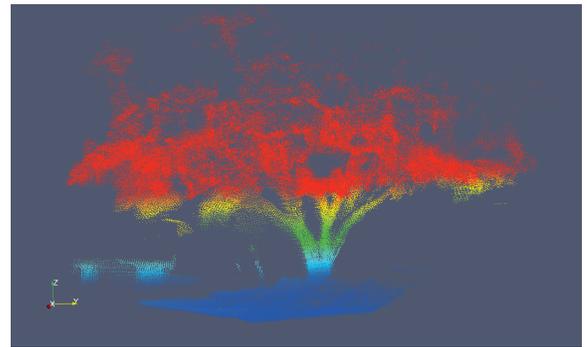


Figure 9: A cut view of the dataset highlighting a tree.

reaches a small road surrounded by dense vegetation. The total path length is around 20 m long.

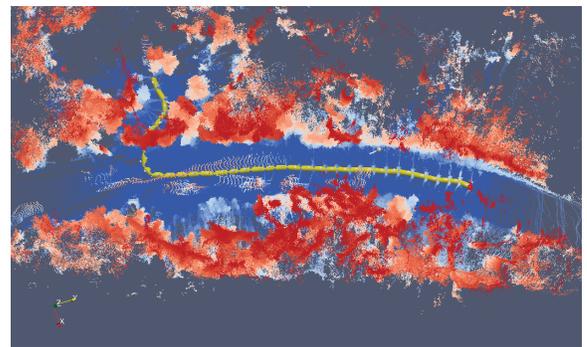


Figure 10: A top view of the realized path (yellow tube) with the ground represented in blue and the vegetation in red.

Conclusion

In this paper, we introduced new datasets aiming at evaluating registration solutions in challenging environments. We achieved global localization of the scanner using a theodolite, which gives us the ability to record datasets in GPS denial environments, indoors or outdoors. The precision achieved is also higher than when using datasets that are already available to the community, which leads to the evaluation of registration algorithms.

While the recorded datasets now mainly cover unstructured and semi-structured environments, we will soon aim at more dynamic environments. The overviewed calibration procedure using ICP in a controlled environment will also be quantified to evaluate its precision in future works.

Acknowledgements

The research presented here was supported by the EU FP7 IP projects Natural Human-Robot Cooperation in Dynamic Environments (ICT-247870; <http://www.nifti.eu>). François Pomerleau was supported by a fellowship from the Fonds québécois de recherche sur la nature et les technologies (FQRNT). Special thanks to Leica Geosystems for their theodolite.

References

- Ceriani, S.; Fontana, G.; Giusti, A.; Marzorati, D.; Matteucci, M.; Migliore, D.; Rizzi, D.; Sorrenti, D. G.; and Taddei, P. 2009. Rawseeds ground truth collection systems for indoor self-localization and mapping. *Autonomous Robots* 27(4):353–371.
- Chen, Y., and Medioni, G. 1991. Object modeling by registration of multiple range images. *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on* 3:2724 – 2729.
- Chetverikov, D.; Svirko, D.; Stepanov, D.; and Krsek, P. 2002. The trimmed iterative closest point algorithm. *Pattern Recognition, 2002. Proceedings. 16th International Conference on* 3:545 – 548.
- Lai, K., and Fox, D. 2010. Object Recognition in 3D Point Clouds Using Web Data and Domain Adaptation. *The International Journal of Robotics Research* 1019 – 1037.
- Mortensen, E.; Deng, H.; and Shapiro, L. 2005. A SIFT descriptor with global context. *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* 1:184 – 190 vol. 1.
- Munoz, D.; Bagnell, J.; Vandapel, N.; and Hebert, M. 2009. Contextual classification with functional Max-Margin Markov Networks. *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* 975 – 982.
- Pathak, K.; Birk, A.; Vaskevicius, N.; and Poppinga, J. 2010. Fast Registration Based on Noisy Planes With Unknown Correspondences for 3-D Mapping. *Robotics, IEEE Transactions on* PP(99):1 – 18.
- Pomerleau, F.; Magnenat, S.; Colas, F.; Liu, M.; and Siegwart, R. 2011. Tracking a depth camera: Parameter exploration for fast icp. *Intelligent Robots and Systems (IROS)* 1–6.
- Rusinkiewicz, S., and Levoy, M. 2001. Efficient variants of the ICP algorithm. *3-D Digital Imaging and Modeling, 2001. Proceedings. Third International Conference on* 145 – 152.