Model Acquisition by Registration of Multiple Acoustic Range Views

Andrea Fusiello, Umberto Castellani, Luca Ronchetti, Vittorio Murino
University of Verona
{fusiello,castellani,murino}@sci.univr.it

Objective

Obtain a 3D model of an an object from multiple range views.

Assume that there are M overlapping views (or point sets) $V^1 \dots V^M$, each taken from a different viewpoint.

The objective is to find the best rigid transformations $\mathbf{G}^1 \dots \mathbf{G}^M$ to apply to each set, bringing them into a common reference frame where they are seamless aligned.

Compostition of rigid transformations yields sub-optimal results: a *global* approach is needed.

Context

The final goal is to display a 3D scene model to the human operator(s) of an *underwater* Remotely Operated Vehicle (ROV), in order to facilitate the navigation and the understanding of the surrounding environment.

Range data comes from a high frequency acoustic 3D camera, called Echoscope, which outputs a 64×64 range image. Resolution is about 3 cm at 500 KHz. Speckle noise affect images.

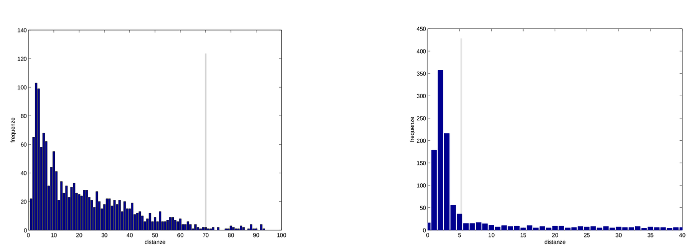
Robust pairwise registration

First we calculate the registration matrix $\mathbf{G}^{i,j}$ between every pair of overlapping views $V^i\ V^j$:

$$V^i = \mathbf{G}^{i,j} V^j \qquad (1)$$

The Iterative Closest Point (ICP) algorithm can give very accurate results when a set is a subset of the other, but results deteriorate with *outliers*, created by non-overlapping areas between the two sets.

Good correspondences can be discriminated by using an outlier rejection rule on the distribution of closest point distances ϵ_i.



Distances histograms. Zhang's (left) and X84's.

The *X84* rejection rule uses robust estimates for location and scale of a corrupted Gaussian distribution to set a rejection threshold.

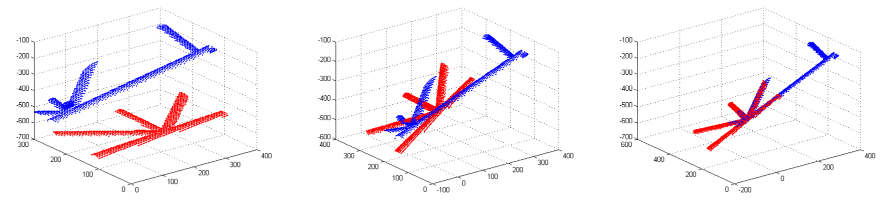
The median is a robust location estimator, and the Median Absolute Deviation

$$\mathrm{MAD} = \operatorname*{med}_i\{|\epsilon_i - \operatorname*{med}_j \epsilon_j|\} \qquad (2)$$

is a robust estimator of the scale.

Values that are more than 5.2 MADs away from the median are rejected.

X84 has a *breakdown point* of 50%: any majority of the data can overrule any minority.



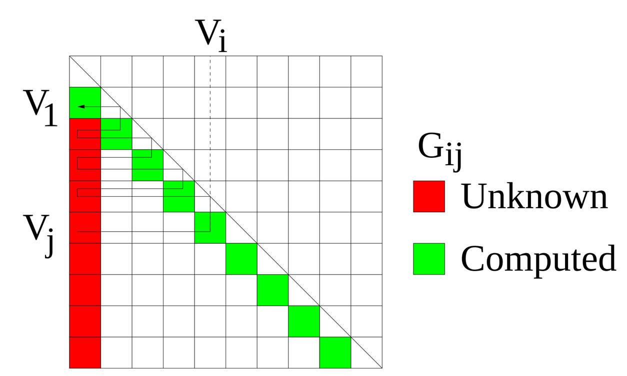
Registration. An example of two point sets that ICP failed to align, but ICP+X84 succeded.

Chaining pairwise transformations

Then we compute a starting guess for the global registration $\mathbf{G}^i$, that map V^i into the space the reference view, V^1 by chaining pairwise transformations between consecutive frames $\mathbf{G}^{i-1,i} \dots \mathbf{G}^{1,2}$:

$$\mathbf{G}^i = \prod_{j=2}^{i} \mathbf{G}^{j-1,j}. \qquad (3)$$

The combination of pairwise registration does not yield the optimal result.



For example, if $\mathbf{G}^{k,i}$ and $\mathbf{G}^{i,j}$ are optimal on the sense that they minimize the mean square error distance between the respective sets, then $\mathbf{G}^{k,j} = \mathbf{G}^{k,i}\mathbf{G}^{i,j}$ does not necessarily minimizes the mean square error between views V^j and V^k.

Global transformations adjustment

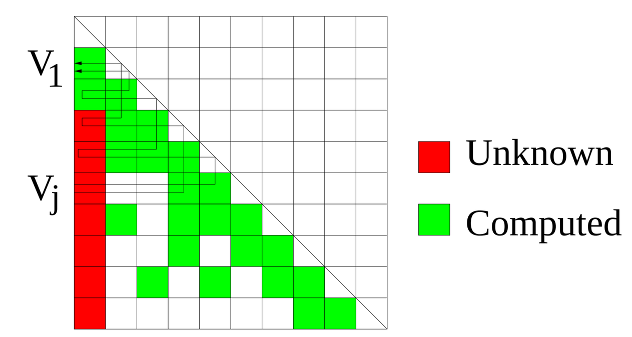
The aim of the global registration is to improve the quality of the mosaic, by using the information coming from *every* overlapping view pairs, not just consecutive.

Key idea: obtain the global registration by introducing algebraic constraint *on the transformations*, instead of data points.

By considering all the $\mathbf{G}^{i,j}$ matrices we can write a number of equations as

$$\mathbf{G}^j = \mathbf{G}^i \mathbf{G}^{i,j}. \qquad (4)$$

where the $\mathbf{G}^{i,j}$ are known, and $\mathbf{G}^i = \mathbf{G}^{1,i}\ \ (2 \le i \le N)$ are the sought unknowns.



This non-linear least squares problem can be cast as the minimization of the following objective function:

$$\min \sum_{i,j} \left[\left(\frac{\angle(\mathbf{R}^i \mathbf{R}^{i,j} (\mathbf{R}^j)^\top)}{\sigma_\alpha} \right)^2 + \left(\frac{||\mathbf{R}^i \mathbf{t}^{i,j} + \mathbf{t}^i - \mathbf{t}^j||}{\sigma_t} \right)^2 \right] \qquad (5)$$

where $\angle(\cdot)$ is an operator that takes a rotation matrix and returns the angle of rotation around a suitable axis, σ_α and σ_t are normalization factors.

Starting from the global registration obtained by chaining pairwise transformation, a Least-Squares solution is iteratively sought, using Gauss-Newton algorithm.

The complexity of the proposed algorithm is *independent from the number of points involved.*

As the objective function includes only the matrix components, the complexity depends only on the number of (overlapping) views.

Dealing with rotations

Rotations are represented with unit quaternions, which are particularly well suited to the requirements of an iterative search algorithm, namely that a small increment of the parameters corresponds to a small increment of the rotation.

Instead of requiring the quaternion $\mathbf{q}$ to be a unit vector, we enforce the constraint that the corresponding rotation matrix is orthonormal:

$$\mathbf{R}(\mathbf{q}) = \frac{1}{\mathbf{q} \cdot \mathbf{q}} \mathbf{R}_u(\mathbf{q}) \qquad (6)$$

where $\mathbf{R}_u(\mathbf{q})$ is the rotation matrix associated to the quaternion.

Algorithm: global registration

Step 1. *compute pairwise registration between every view pair using ICP algorithm, with X84 rejection rule;*

Step 2. *if the pairwise registration is good (i.e. views are overlapping), accept the registration matrix* $\mathbf{G}^{i,j}$*;*

Step 3. *compute a starting guess for the global registration by chaining pairwise transformation (Eq. (3));*

Step 4. *minimize the objective function defined in Eq. (5) with a Gauss-Newton method (MATLAB* `lsqnonlin` *function); At each step enforce orthogonality of rotation matrix with Eq. (6)*

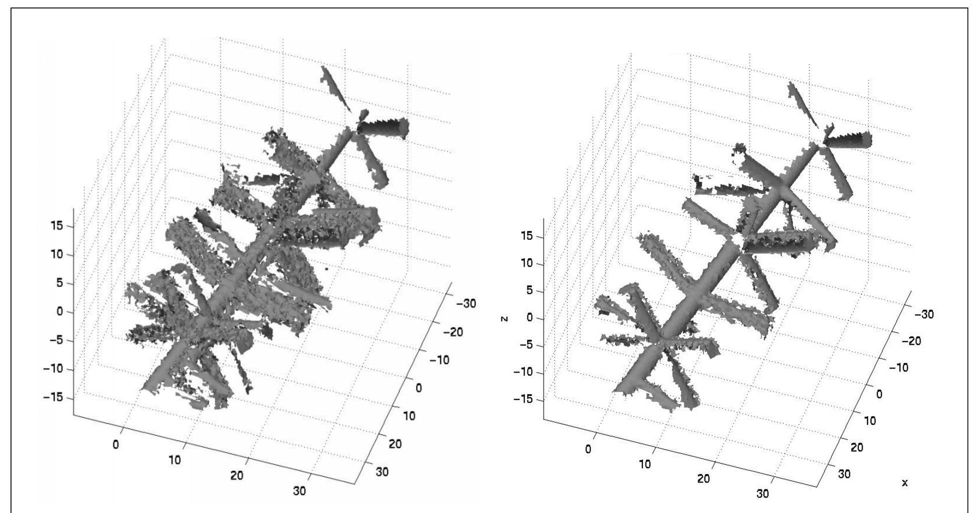
Step 5. *apply the transform defined by* $\mathbf{G}^i$ *to the view* V^i, $(2 \le i \le N)$.

Registered sets of points must be fused in order to get a single 3D model. We used the public domain implementation of Hoppe and De Rose algorithm.

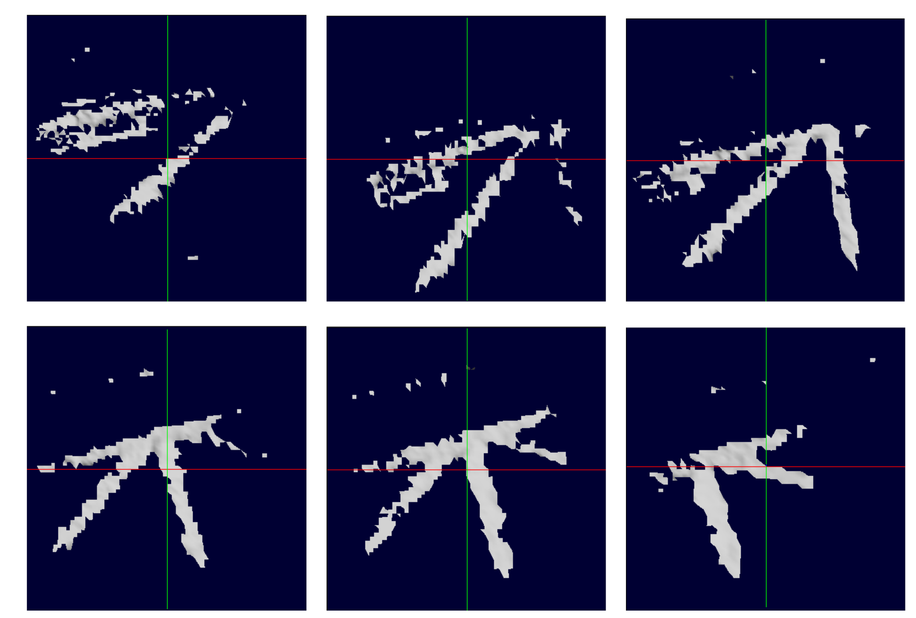
Results

Synthetic exp. to compare the average rotation error and its variance over the views.

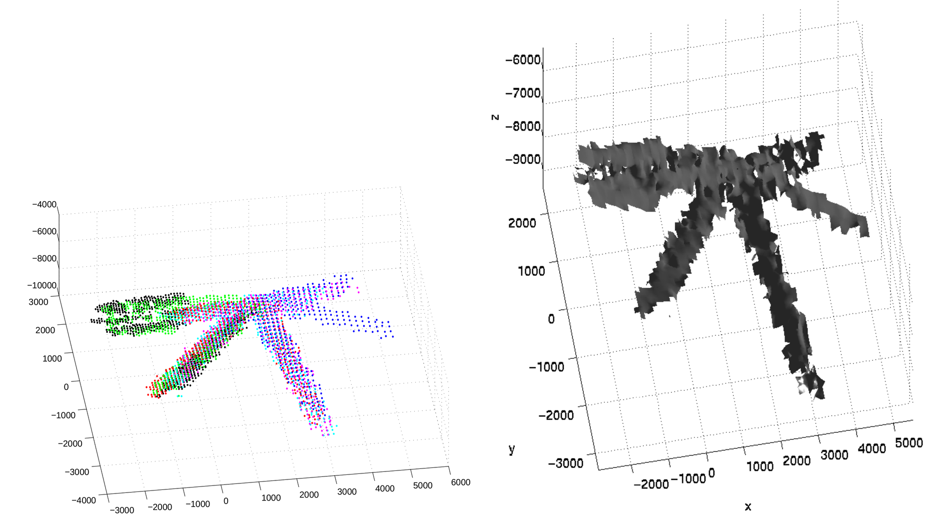
	Pairwise	Global	% diff.
avg error	0.0463	0.0381	17.7%
variance	0.00243	0.00108	55.6%



Synthetic views. Surface reconstructed before (left) and after (right) the global registration.



Sample Echoscope frames.



Mosaic. Clouds of points (left) and reconstructed surface (right).

Conlusions

Complexity does not depend on the number of points involved, but only on the number of views.

Error is only spread among the views, but does not get reduced significantly.

This technique is well suited for all the application where speed can be traded for accuracy.

Acknowledgements: This work was supported by the European Commission under the project no. GRD1-2000-25409 ARROV (Augmented Reality for Remotely Operated Vehicles based on 3D acoustical and optical sensors for underwater inspection and survey). Echoscope images are courtesy of Dr. R.K. Hansen of Omnitech A/S (Norway).