Segmentation of dynamic objects from low acquisition rate range data

Agustin Ortega and Juan Andrade-Cetto
Institut de Robòtica I Informàtica Industrial

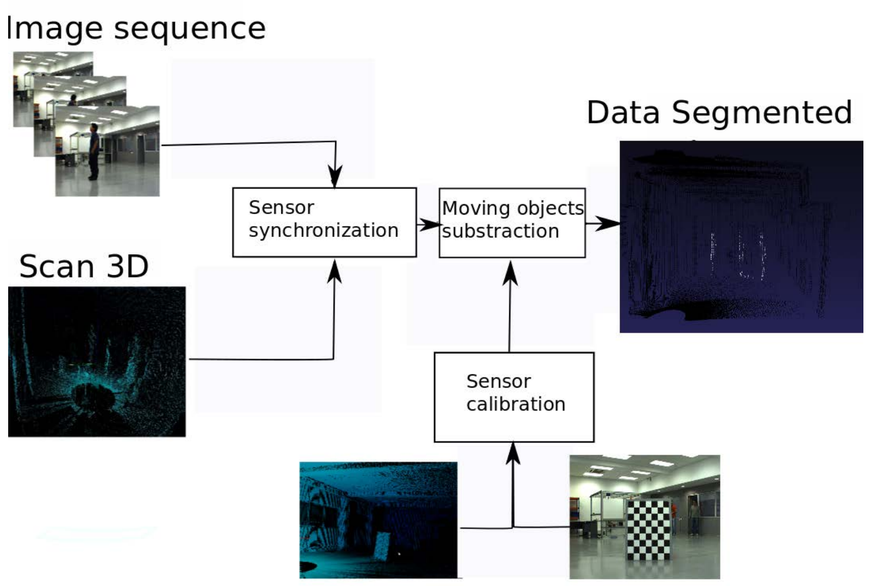
Problem definition

We address the problem of eliminating moving bodies from low acquisition rate 3d range scans by using high rate image data.

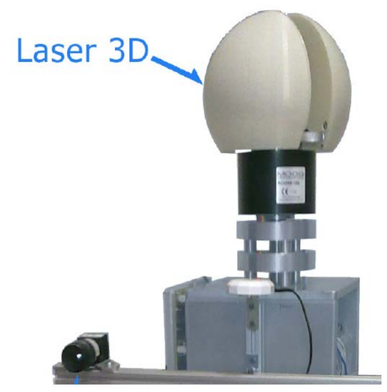
Range points are classified according to temporally consistent corresponding images per- pixel class distributions are learned on-line as Gaussian Mixture.

Approach

To appear ECMR'11 [1].



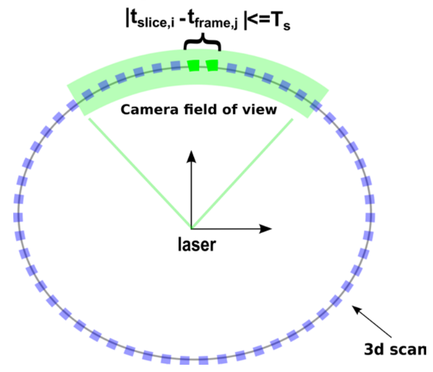
Sensor specifications



A Hokuyo UTM-30LX laser is mounted on a slip-ring, with angular resolution of 0.5 degrees in azimuth and elevation, and 360 degrees omnidirectional field of view and elevation range of 270 degrees.

A Pointgray Flea color camera with M1214-MP optics and 40.4 degrees field of view is rigidly attached to the sensor structure.

Sensor synchronization



The 3D scanner takes 18 seconds to complete a full scan. The camera frame rate is set to 17 fps.

The timestamps between consecutive laser slices, and grabbed images are compared and set to lie within a fixed threshold Ts in milliseconds.

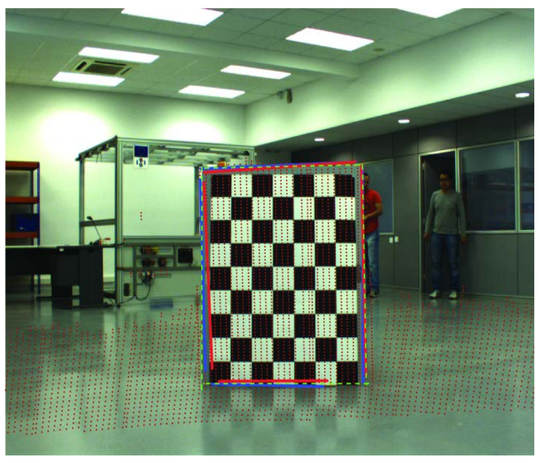
Camera-laser calibration

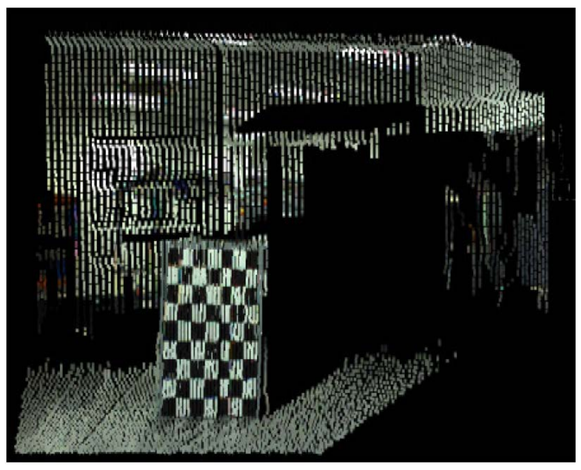
Intrinsic camera parameters are computed using Zhang's method [3].

Extrinsic calibration between the 3D laser and the camera is initialized using Hager's method [4] for pose estimation over a human-selected set of point correspondences.

Camera to laser extrinsic calibration is then refined by minimizing the projection error of line primitives, computed as the weighted sum of both angular and midpoint location errors [2].

$$\epsilon = \sum (\theta_i - \theta_p)^2 + w(m_i - m_p)^T (m_i - m_p)$$





Segmentation of moving objects

The probability of each image pixel to be of the static object class is modeled as a weighted mixture of *K* Gaussian distributions.

$$P(x|\mathcal{S}) = \sum_{k=0}^{K} \omega_k \eta\left(\mu_k, \Sigma_k\right)$$

This mixture is iteratively refined from image data

$$\begin{aligned} \omega_k(t+1) &= (1-\alpha)\omega_k(t) + \alpha \qquad \rho = \alpha\eta\left(\mu_k, \Sigma_k\right) \\ \mu_k(t+1) &= (1-\rho)\mu_k(t) + \rho x \\ \Sigma_k(t+1) &= (1-\rho)\Sigma_k(t) + \rho(x-\mu(t))^T(x-\mu(t)) \end{aligned}$$

Static class

We choose as the static object class, the first *S<K* ordered distributions which add up to a factored weight ωs, where

$$S = \mathtt{argmin}_S\left(\sum_{i=1}^{s} \omega_i \geq \omega_s\right)$$

Point classification

For each point in a slice, its corresponding pixel values for the subset of matching synchronized images is measured and compared to the full set of pixel values for the whole sequence by means of the learned mixture distribution.

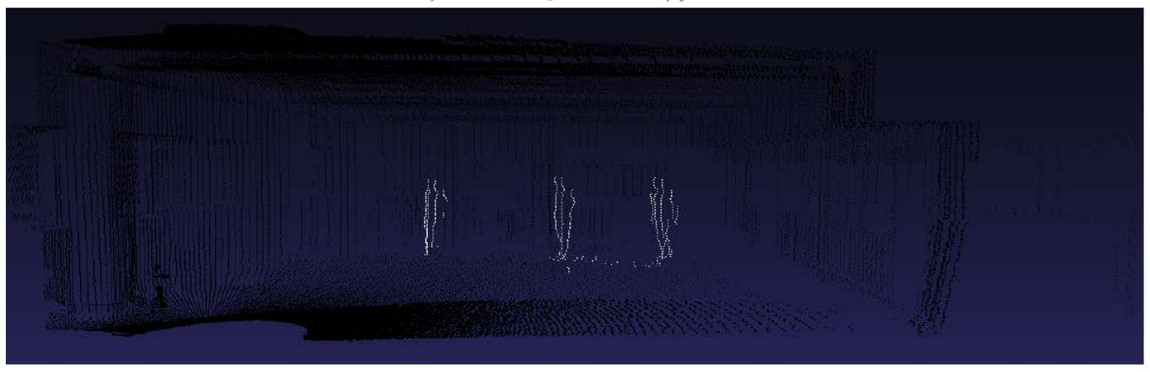
A range point is classified as static if all matching pixels in the synchronized image subset are within the learned static class.

Experiments

(a) $T_s = 1/\text{fps}$. (b) $T_s = 0.5\text{sec}$ (c) Foreground segmentation

three sequence instances segmented at $T_s = 1/\text{fps}$.

Segmented dynamic object



Segmentation results for a sequence with three slowly moving people with random walking trajectories.

Conclusions

- We present a method to segment low-rate 3D range data as static or dynamic using multimodal classification.
- Special attention is paid to the synchronization and metric calibration of the two sensing devices.
- The proposed method was designed to remove spurious data.
- The intended application of the technique is robotic 3d mapping

References

[1] A. Ortega, J. Andrade-Cetto,Segmentation of dynamic objects from laser data, to appear ECMR'11. .
[2]A. Ortega, B. Dias, E. Teniente, A. Bernardino, J. Gaspar, and Juan Andrade-Cetto. Calibrating an outdoor distributed camera network using laser range finder data. IROS'09
[3]Z. Zhang. A flexible new technique for camera calibration. PAMI'00.
[4]C.P. Lu, G.D. Hager, and E. Mjolsness. Fast and globally convergent pose estimation from video images.PAMI'00