A PROBABILISTIC APPROACH TO TOF AND STEREO DATA FUSION

C. Dal Mutto, P. Zanuttigh, G.M. Cortelazzo - University of Padova
{dalmutto, pietro.zanuttigh,corte}@dei.unipd.it

ICVSS 2010
Registration and Video Analysis

DEPARTMENT OF INFORMATION ENGINEERING
UNIVERSITY OF PADOVA

LTTM

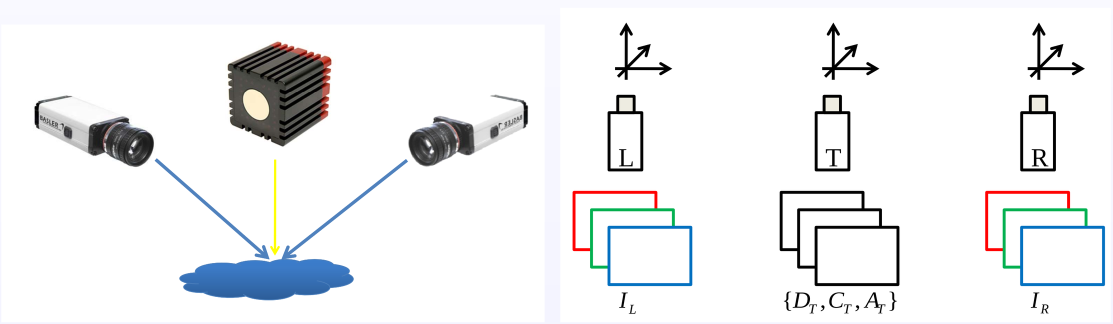
Abstract

Depth information can be acquired real-time by stereo vision systems and ToF cameras. Both solutions present critical issues that can be overcome by their combined use. In this work, a heterogeneous acquisition system is considered, made of two standard cameras and one ToF camera. This paper introduces a novel multi-camera calibration technique based on the combined use of the color information and of the ToF depth data and a probabilistic fusion algorithm to combine the two devices.

Acquisition Setup

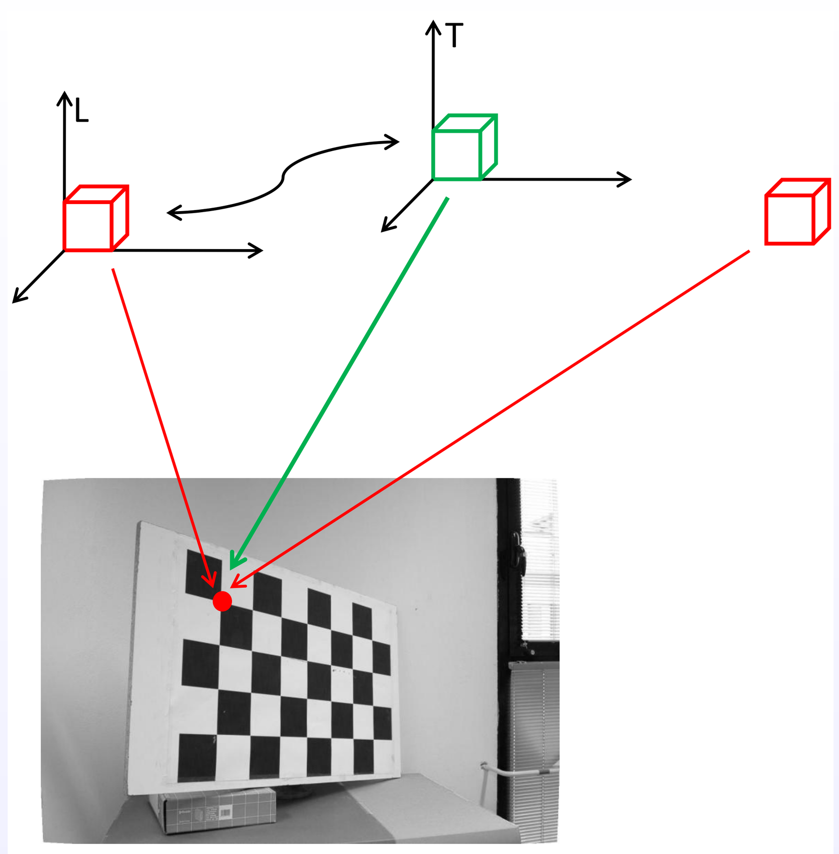
Trinocular heterogeneous acquisition system

- A ToF camera T
- Two standard videocameras $\{L, R\}$ (stereo pair)



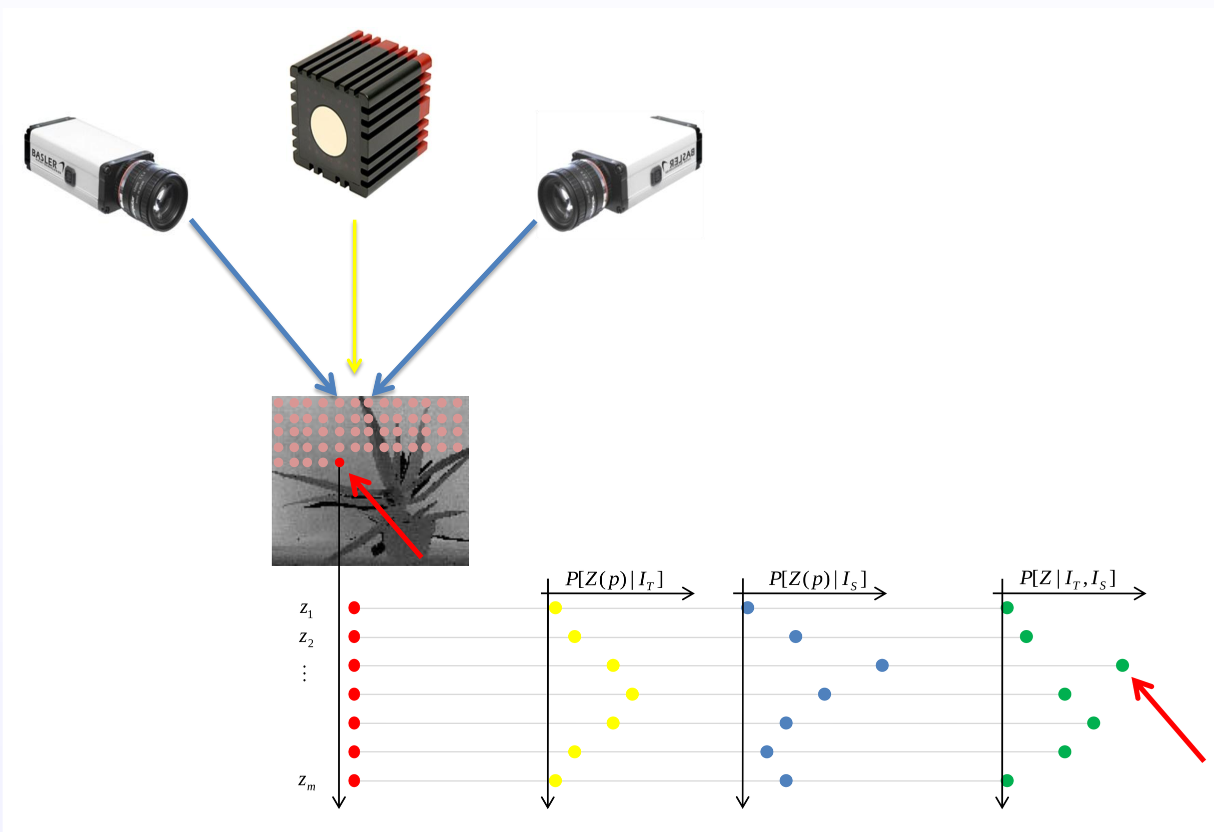
Calibration

- Intrinsic Parameters (distortions, systematic depth error for ToF)
- Extrinsic Parameters (rototranslations between cameras)
  - Stereo already calibrated (*OpenCV*, Camera Calibration Toolbox for Matlab)
  - Rototranslation between the stereo pair reference system (camera L) and the ToF camera T reference system: M_L
    - n points acquired with a calibration checkerboard
    - 3D points from the stereo pair: P_S^i (triangulation)
    - 3D points from the ToF camera: P_T^i (backprojection)
    - $\arg\min_{M_L} \sum_{i=1}^{n} ||P_T^i - M_L P_S^i||_2$
    - 3D Absolute Orientation Problem
    - Horn Algorithm + RANSAC
    - Average Calibration Error: $0.7[cm]$



Probabilistic Fusion Model

- ToF Images:
  $I_T = \{C_T, D_T\}$
- Stereo Images:
  $I_S = \{I_L, I_R\}$
- For each point on the lattice of T images, the estimated depth is:
  $\hat{Z} = \arg\max_Z P[Z|I_T, I]$
- Bayes Rule:
  $P[Z|I_T, I_S] = \frac{P[I_T, I_S|Z]P[Z]}{P[I_T, I_S]}$
- $P[Z|I_T, I_S] \propto P[I_T, I_S|Z]P[Z]$
- $P[Z]$ is uniformly distributed
- $P[Z|I_T, I_S] \propto \frac{P[I_T, I_S|Z]P[Z]P[Z]}{P[I_T]P[I_S]}$
- Hp: $\{I_S|Z\}$ is independent from $\{I_T|Z\}$
- $P[Z|I_T, I_S] \propto \frac{P[I_T|Z]P[Z]}{P[I_T]} \frac{P[I_S|Z]P[Z]}{P[I_S]}$
- $\hat{Z} \approx \arg\max_Z P[Z|I_T]P[Z|I_S]$



ToF and Stereo Models

ToF Camera Model:

The ToF camera model takes into account:

- Thermal noise component, distributed as $\mathcal{N}(0, \sigma_t^2)$
- Scattering generated error, approximated by the variance of the depth (σ_s^2) in the second order neighborhood of $\mathbf{p}$

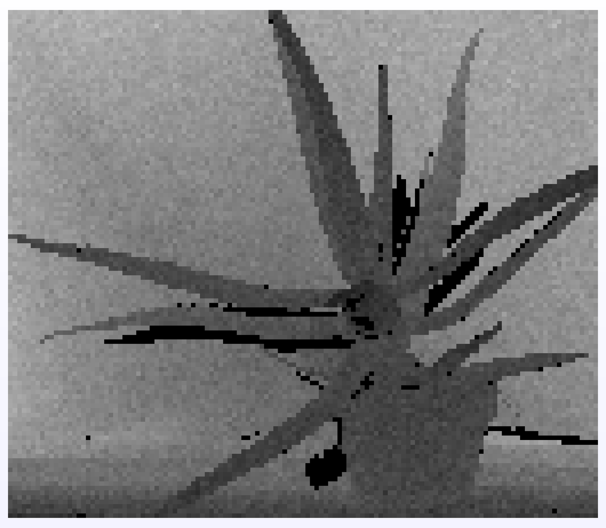
For each pixel $\mathbf{p}$ in the Z lattice ($\mathcal{Z}$):

- $P[Z(\mathbf{p})|I_T] \sim \mathcal{N}(d, \sigma_w^2)$
- $d = D_T(\mathbf{p})$, and $\sigma_w = \max(\sigma_t, \sigma_s)$.

Stereo Pair Model:

- Each sample $\mathbf{p}_i$, $i = 1, \ldots, m \in [d - 3\sigma_w, d + 3\sigma_w]$ is reprojected into I_L and I_R
- A TAD cost function of the coupling $\mathcal{C}_i$ is calculated
- $P[Z(\mathbf{p}) = z_i(\mathbf{p}_i)|I_T, I_S] \propto \exp -\frac{\mathcal{C}_i(\mathbf{p})}{\sigma_I}$, where σ_I is the noise standard deviation in $\{I_L, I_R\}$

Experimental Results

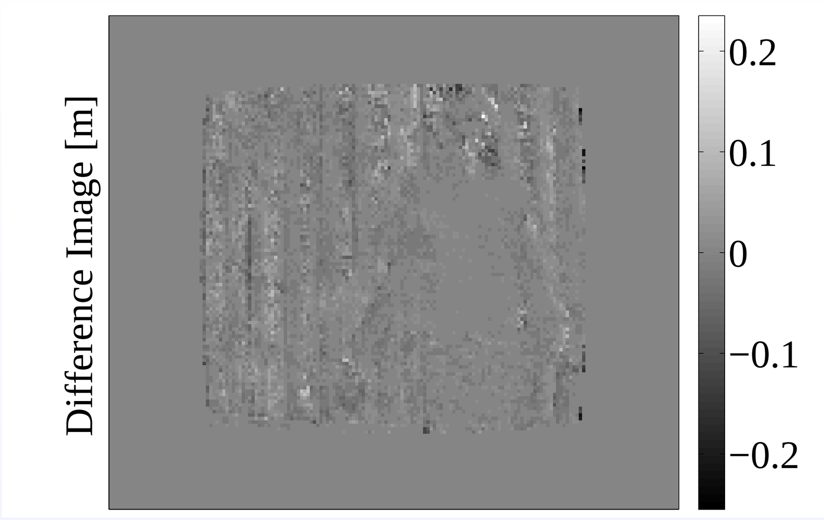


Synthetic ToF image I_T (noisy)

Synthetic depth image Z after the fusion algorithm

Real scene image I_L

Real scene depth image Z after the fusion algorithm



Difference between Z and I_T on the real scene

http://lttm.dei.unipd.it