Robust Camera Motion Estimation using Direct Edge Alignment and Sub-gradient Method

Manohar Kuse & Shaojie Shen

Robotics Institute (RI)
Hong Kong University of Science & Technology

mpkuse@connect.ust.hk

October 26, 2015

Overview

Introduction

Proposed Formulation
- Relative Pose Estimation
- Energy Formulation
- Lie Manifolds
- DT Relaxation

Numerical Solution
- Sub-Gradient Method
- Computation of a Sub-gradient
- Fast Convergence

Results

Concluding Remarks

References

Visual Odometry : Problem Definition

- RGB Images : $I_0, I_1, \cdots I_n$.
- Depth Images : $Z_0, Z_1, \cdots Z_n$.
- Determine camera pose (${}^{k}_{k-1}\mathbf{R}$, ${}^{k}_{k-1}\mathbf{t}$) between adjacent frames.
- ${}^{k}_{k-1}\mathbf{R} \in SO(3)$ is a rotation matrix.
- ${}^{k}_{k-1}\mathbf{t} \in \mathbb{R}^3$ is a translation matrix.

Literature Review for Odometry from RGB-D Camera

- Feature-based methods.
- Direct methods (or feature-less methods).

Feature-based methods

- Extract features eg. SIFT, SURF, Harris etc.
- Find correspondences between images.
- Solve Perspective-N-Point problem.
- Huang *et al.* [4], Dryanovski *et al.* [2].
- **Disadvantages** :
  Inaccurate, since it discards most of image information.
  Sensitive to incorrect correspondences.
  Comparitively unstable

Direct Methods

- Do not involve feature extraction.
- Use of entire image information for pose determination.
- Based on minimizing the photometric error to get best pose.
- Relatively new approach Kerl et al. [5].
- **Disadvantages** :
  Sensitive to illumination variation and noise.
  Small basin of attraction.
  Non-differentiable cost function.

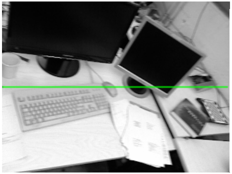
Analysis on Non-differentiability of Cost functions

- Previous work by Kerl et al. [5] suggest to minimize the linearized photometric error cost.

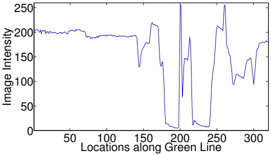
$$r(\xi, {}^{r}\mathbf{P}_i) = I_n(\Pi[exp(\xi)\ {}^{r}\mathbf{P}_i]) - I_r({}^{r}\mathbf{u}_i)$$

$$\xi^* = \operatorname*{argmin}_{\xi} \sum_i r_{lin}(\xi, {}^{r}\mathbf{P}_i)^2.$$

- They proposed to use the Gauss-Newton method to minimize this function.
- Gauss-Newton method is not guaranteed to converge.
- We argue about the non-differentiability of the cost function proposed by Kerl et al. [5].
- Thus, we propose to use the sub-gradient method (numerical method for optimization of non-differentiable functions).



(a)



(b)

Figure : Highlighting the non-differentiability of the function I_n at object transition points

$$r(\xi, {}^r\mathbf{P}_i) = I_n \circ \Pi \circ \tau(\xi, {}^r\mathbf{P}_i)$$

Theorem (Continuity and Composition)

if g is continuous at a and f is continuous at $g(a)$, only then $f \circ g$ is continuous at a.

Overview

Introduction

Proposed Formulation
- Relative Pose Estimation
- Energy Formulation
- Lie Manifolds
- DT Relaxation

Numerical Solution
- Sub-Gradient Method
- Computation of a Sub-gradient
- Fast Convergence

Results

Concluding Remarks

References

Notations

- RGB Image $I_k : \Omega \subset \mathbb{R}^2 \mapsto \mathbb{R}$.
- Depth Image $Z_k : \Omega \subset \mathbb{R}^2 \mapsto \mathbb{R}$.
- 3D Scene Point ${}^k\mathbf{P} \in \mathbb{R}^3$ at k^{th} timestamp.
- Projection Function ${}^k\mathbf{u} = \Pi({}^k\mathbf{P})$. $\Pi : \mathbb{R}^3 \mapsto \mathbb{R}^2$.
- Inverse Projection ${}^k\mathbf{P} = \tilde{\Pi}({}^k\mathbf{u}, Z_k({}^k\mathbf{u}))$. $\tilde{\Pi} : (\mathbb{R}^2, \mathbb{R}) \mapsto \mathbb{R}^3$.
- Edge image point ${}^r\mathbf{e}_i \in \mathbb{R}^2$ and its inverse projection ${}^r\mathbf{E}_i \in \mathbb{R}^3$.
- Distance transform of edge-map $V^{(k)} : \mathbb{R}^2 \mapsto \mathbb{R}$.

Relative Pose Estimation

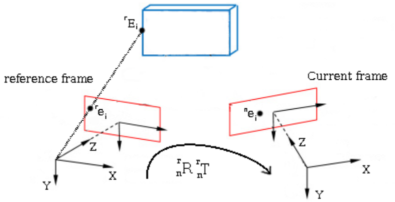


Figure : Notations and Conventions

- Estimation of rigid transformation between a Reference frame (r) and Current Frame (n) as an Optimization Problem.
- Reference frame periodically refreshed.
- Relative poses chained to obtained robot state estimate.

Energy Formulation

- Proposed Geometric Energy function :
  Sum of the squared distances between,
  a) re-projections (of edge points from reference image)
  AND
  b) nearest edge points in current image:

$$f(\mathbf{R},\mathbf{T}) = \sum_i \min_j D^2\left(\Pi[\mathbf{R}^T({}^r\mathbf{P}_i - \mathbf{T})],\ {}^n\mathbf{u}_j \right).$$

$$\underset{\mathbf{R},\mathbf{T}}{\text{minimize}}\ f(\mathbf{R},\mathbf{T})$$

$$\text{subject to } \mathbf{R} \in SO(3)$$

- Non-linear and Non-convex

Optimization with Unitary Constraints

- Theory of Optimization under unitary constraints [6] proposes to use an appropriate manifold.
- Projection of the iterate onto the manifold gives a natural framework to model the non-convex unitary (here orthogonal) constraint.
- Lie manifolds widely used to model a rotation matrix.

$$\xi = (\mathbf{t}; \mathbf{w})^T \in \mathbb{R}^6.$$

$$\tau({}^r\mathbf{P}_i, \xi) = \left[exp(\xi)\right]^{-1} {}^r\mathbf{P}_i = \mathbf{R}^T({}^r\mathbf{P}_i - \mathbf{T}).$$

Unconstrained Reformulation

- Using the Lie manifolds the problem can be reformulated as an unconstrained optimization problem.

$$\underset{\xi}{\text{minimize}} \sum_i \min_j D^2\left(\Pi[\tau({}^r\mathbf{P}_i, \xi)],\ {}^n\mathbf{u}_j \right).$$

Relaxation using Distance Transform

- Observation : If edge points pre-selected (from reference frame), $\min_j D(\mathbf{u}_i, \mathbf{u}_j)$ is exactly the definition of the Distance Transform [3].
- Using the Lie Manifold and the Distance Transform relaxation.

$$\upsilon_{e_i}(\xi) = V^{(n)}\left(\Pi[\tau(\tilde{\Pi}({}^r\mathbf{e}_i, Z_r({}^r\mathbf{e}_i)), \xi)]\right)$$

$$f(\xi) = \sum_{\forall \mathbf{e}_i} (\upsilon_{e_i}(\xi))^2.$$

- Optimal estimates of ξ obtained as

$$\xi^* = \underset{\xi}{\operatorname{argmin}} \sum_{\forall \mathbf{e}_i} (\upsilon_{e_i}(\xi))^2$$

- ξ^* back-projected using logarithm-map from the Lie Manifold to obtain the optimal rigid transform between the frames.

Overview

Introduction

Proposed Formulation
- Relative Pose Estimation
- Energy Formulation
- Lie Manifolds
- DT Relaxation

Numerical Solution
- Sub-Gradient Method
- Computation of a Sub-gradient
- Fast Convergence

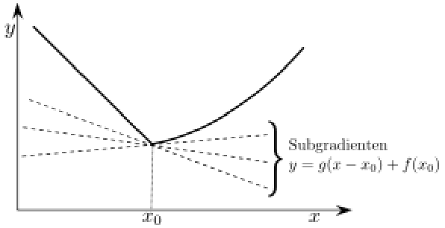
Results

Concluding Remarks

References

Sub-Gradient Method

- Numerical method for minimizing non-differentiable functions (Boyd *et al.* [1])
- First order method
- Concept of sub-gradients (a set). Generalizing the concept of gradients.



Sub-Gradient Method

- Start with an initial estimate $(\xi^{(0)})$.
- Update along negative sub-gradient direction

$$\xi^{(k+1)} = \xi^{(k)} + (-\alpha_k)\ \hbar^{(k)}.$$

$$\xi_a + \xi_b := log(exp(\xi_a)\ exp(\xi_b)).$$

- Step length fixed ahead of time.

$$\alpha_p := \frac{\eta}{(p+1)};\ p = 0, \cdots, k$$

- Keep track of best estimates.

$$f_{best}^{(k)} = min(f(\xi^{(0)}), f(\xi^{(1)}), \ldots, f(\xi^{(k)})).$$

Computation of a Sub-gradient

- All **c** which satisfy following inequality are the sub-gradients of $\upsilon_{e_i}(\xi)$ wrt ξ at $\xi = \xi^{(k)}$.

$$\lim_{\xi \to \xi^{(k)}} \upsilon_{e_i}(\xi) - \upsilon_{e_i}(\xi^{(k)}) \geq \lim_{\xi \to \xi^{(k)}} \mathbf{c} \cdot (\xi - \xi^{(k)}).$$

- Since, $\upsilon_{e_i}(\xi) = V^{(n)} \circ \Pi \circ \tau(\xi, {}^r\mathbf{P}_i)$

$$\mathbf{J}_{e_i} := \frac{\partial V^{(n)}}{\partial e_i} \cdot \frac{\partial e_i}{\partial E_i} \cdot \frac{\partial E_i}{\partial \xi} \bigg|_{\xi = \xi^{(k)}}.$$

- 3^{rd} term computed using approximation of rotation matrix

$$E_i(\xi^{(k)} + \delta\xi) \approxeq [\hat{\mathbf{R}}(\mathbf{I}_3 + \lfloor \delta\mathbf{w} \rfloor_x)]^T \left[{}^r\mathbf{E}_i - \hat{\mathbf{T}} - \delta\mathbf{t} \right].$$

$$\frac{\partial E_i}{\partial \xi} \bigg|_{\xi = \xi^{(k)}} = \hat{\mathbf{R}}^T \left[\ -\mathbf{I}_3 \ \middle| \ \lfloor {}^r\mathbf{E}_i - \hat{\mathbf{T}} \rfloor_x \ \right]$$

Computation of a Sub-gradient

$$f(\xi) = \sum_{\forall \mathbf{e}_i} (\upsilon_{e_i}(\xi))^2.$$

$$\hbar^{(k)} = \sum_{\forall \mathbf{e}_i} 2\ \upsilon_{e_i}(\xi^{(k)})\ \mathbf{J}_{e_i}.$$

Heavy Ball Method

- Technique to speedup convergence of a gradient method
- Use the update direction $\mathbf{s}^{(k)}$ instead of $\hbar^{(k)}$.

$$\mathbf{s}^{(k)} = (1-\beta)\hbar^{(k)} + \beta\mathbf{s}^{(k-1)}.$$

- We derive the convergence

$$f_{best} - f(\xi^*) \leq \frac{R^2 + G^2 \sum_{p=0}^{k} \alpha_p^2 (1-\beta^p)^2}{2\sum_{p=0}^{k} \alpha_p \mu^{(p)}}$$

Summary of Proposed Algorithm

Algorithm 1 RelativePoseEstimation($I_n, I_r, Z_r, \xi^{(0)}$)

E_n = EdgeMap(I_n)
V_n = DistanceTransform(E_n)
E_r = EdgeMap(I_r)
${}^r\mathbf{P_i}$ = InverseProjectAllEdgePixels(E_r, Z_r)
$k = 1$
$\hbar^{(0)} = 0$
for $k : 1 \rightarrow M$ **do**
  $\hbar^{(k)}$ = GetSubGradient($V_n, {}^r\mathbf{P_i} \forall \ \mathbf{i}, \xi^{(\mathbf{k-1})}$) (Eq. 8, 10, 16)
  $\mathbf{s^{(k)}} = (1-\beta)\hbar^{(k)} + \beta \mathbf{s^{(k-1)}}$
  $\Delta\xi = \Gamma(-\alpha_k \mathbf{s^{(k)}})$
  if $||\Delta\xi||_2 < \Delta$ **then**
    break
  else
    $\xi^{(k)} = \xi^{(k)} + \Delta\xi$
    $f(\xi^{(k)}) = GetFunctionValue(V_n, {}^r\mathbf{P_i}\forall\mathbf{i}, \xi^{\mathbf{k}})$ (Eq. 3)
  end if
end for
return $f_{best}^{(M)} = min(f(\xi^{(0)}), f(\xi^{(1)}), \ldots, f(\xi^{(M)}))$

Overview

Introduction

Proposed Formulation
- Relative Pose Estimation
- Energy Formulation
- Lie Manifolds
- DT Relaxation

Numerical Solution
- Sub-Gradient Method
- Computation of a Sub-gradient
- Fast Convergence

Results

Concluding Remarks

References

Relative Pose Error

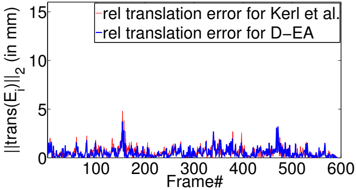
- GT Poses : $\mathbf{Q_1}, \cdots, \mathbf{Q_n} \in \mathbf{SE(3)}$
- Estimated Poses : $\mathbf{B_1}, \cdots, \mathbf{B_n} \in \mathbf{SE(3)}$
- For an integer δ

$$\mathbf{E}_i = \left(\mathbf{Q_i}^{-1}\mathbf{Q_{i+\delta}}\right)^{-1}\left(\mathbf{B_i}^{-1}\mathbf{B_{i+\delta}}\right)$$

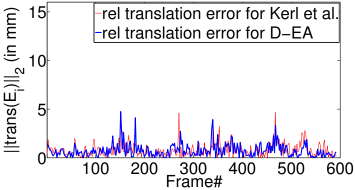
- RMSE of the translation component of the sequence $\mathbf{E_1}, \cdots, \mathbf{E_{n-\delta}}$

Sequence	D-EA		Kerl *et al.* [5]	
	$\delta = 1$	$\delta = 20$	$\delta = 1$	$\delta = 20$
fr2/desk	0.0324	0.1529	0.0333	0.2217
fr1/desk	0.0289	0.0948	0.0346	0.4286
fr1/desk2	0.0335	0.1818	0.0343	0.3658
fr1/floor	0.0355	0.1988	0.0330	0.3380
fr1/room	0.0353	0.2514	0.0307	0.3399
fr2/desk_with_person	0.0125	0.0594	0.0137	0.1516
fr3/sitting_halfsphere	0.0208	0.1462	0.0181	0.2599
fr2/pioneer_slam2	0.0593	0.4447	0.0847	0.4707

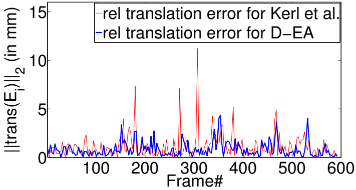
Table : RMSE values of the Relative Pose Errors for various sequences.



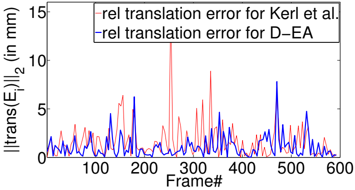
(a) Processing all frames



(b) Processing only frames $0, 2, 4, 6 \cdots$



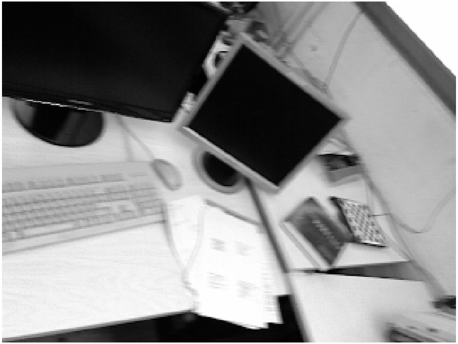
(c) Processing only frames $0, 3, 6, 9 \cdots$



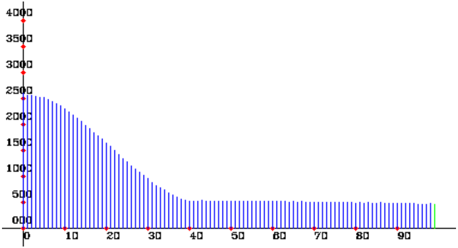
(d) Processing only frames $0, 4, 8, 12 \cdots$

Reference Frame

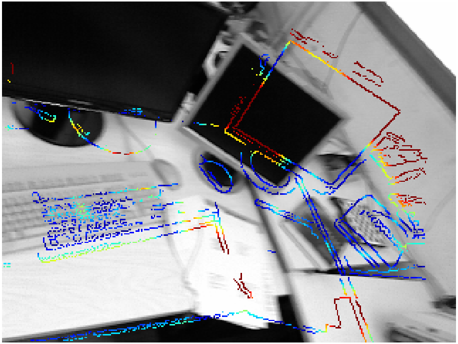
Current Frame



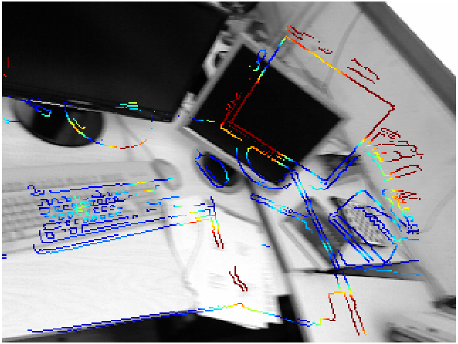
Energy at each Iteration



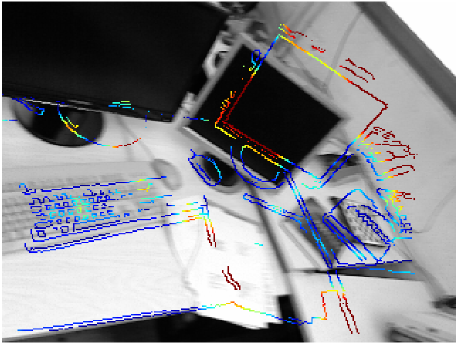
Iteration 0



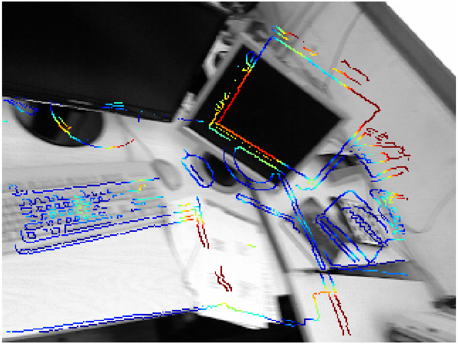
Iteration 5



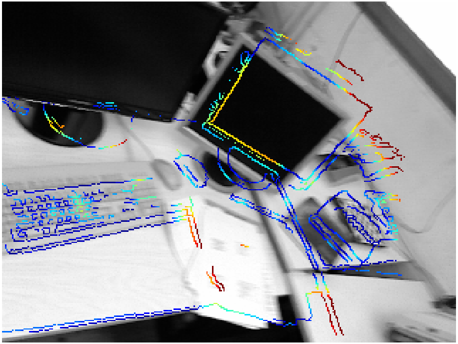
Iteration 10



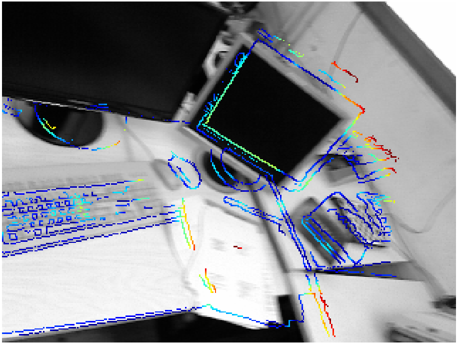
Iteration 15



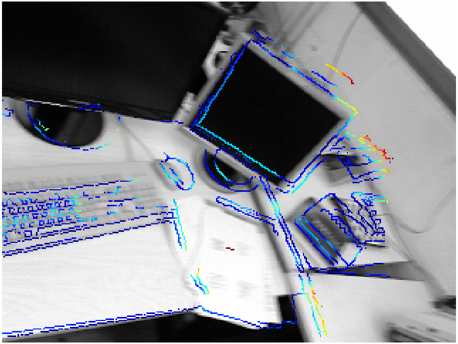
Iteration 20



Iteration 25



Iteration 30



Iteration 35

Iteration 40

Iteration 45

Iteration 50

Overview

Introduction

Proposed Formulation
- Relative Pose Estimation
- Energy Formulation
- Lie Manifolds
- DT Relaxation

Numerical Solution
- Sub-Gradient Method
- Computation of a Sub-gradient
- Fast Convergence

Results

Concluding Remarks

References

Concluding Remarks

- Feature-less approach for 6-DOF pose estimation.
- Novel formulation based on Distance transform.
- Robust to illumination changes, noise and has larger convergence basin.
- Address the problem of non-differentiability of cost function.
- Outperforms previous method for fast motion.

Overview

Introduction

Proposed Formulation
- Relative Pose Estimation
- Energy Formulation
- Lie Manifolds
- DT Relaxation

Numerical Solution
- Sub-Gradient Method
- Computation of a Sub-gradient
- Fast Convergence

Results

Concluding Remarks

References

S. Boyd, L. Xiao, and A. Mutapcic.
Subgradient methods.
lecture notes of EE392o, Stanford University, Autumn Quarter, 2004:2004–2005, 2003.

I. Dryanovski, R. G. Valenti, and J. Xiao.
Fast visual odometry and mapping from RGB-D data.
In *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pages 2305–2310. IEEE, 2013.

P. F. Felzenszwalb and D. P. Huttenlocher.
Distance Transforms of Sampled Functions.
Theory of computing, 8(1):415–428, 2012.

A. S. Huang, A. Bachrach, P. Henry, M. Krainin, D. Maturana, D. Fox, and N. Roy.

Visual odometry and mapping for autonomous flight using an RGB-D camera.
In *International Symposium on Robotics Research (ISRR)*, pages 1–16, 2011.

C. Kerl, J. Sturm, and D. Cremers.
Robust odometry estimation for RGB-D cameras.
In *ICRA*, pages 3748–3754. IEEE, 2013.

J. H. Manton.
Optimization algorithms exploiting unitary constraints.
IEEE Transactions on Signal Processing, 50(3):635–650, 2002.