Projective Bundle Adjustment from Arbitrary Initialization Using the Variable Projection Method

Je Hyeong Hong* University of Cambridge

Christopher Zach Toshiba Research Europe

Andrew Fitzgibbon Microsoft

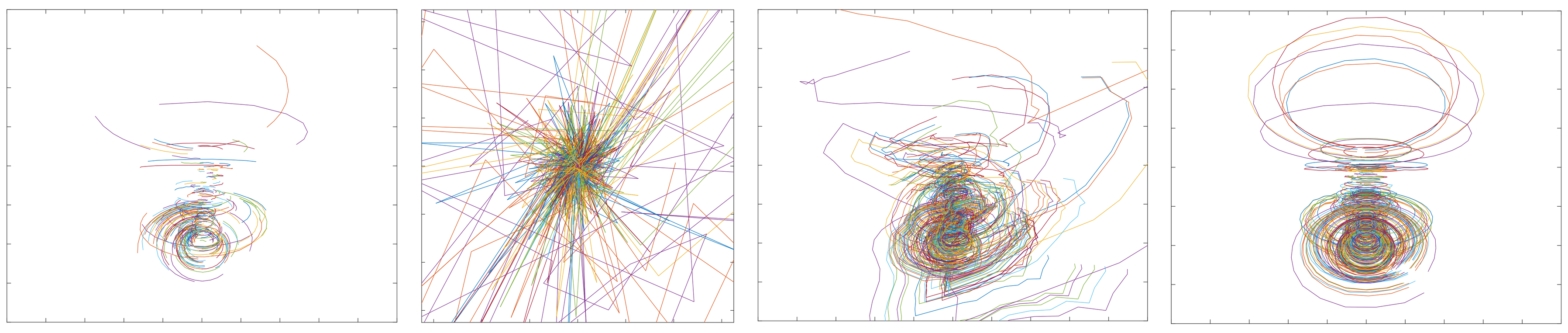
Roberto Cipolla University of Cambridge

1. Paper outline

We apply the **Variable Projection** (**VarPro**, §3) method, recently shown to be effective in matrix factorization [2], to the related problem of **bundle adjustment** (**BA**) **for uncalibrated cameras** (§2). We present a taxonomy of VarPro-based BA algorithms by unifying the affine and the projective camera models (§2).

Assuming cameras and points are initialized arbitrarily, we discover that:

- VarPro performs well on Affine BA as expected but not on Projective BA, and
- a two-stage meta-algorithm (TSMA, §4b) – VarPro on Affine BA, then VarPro on Projective BA – widens the convergence basin (§6).



(a) Di2 tracks (§5) (b) Initial tracks (c) Best affine optimum (d) Best projective opt.

Our two-stage meta-algorithms (TSMAs) are able to reconstruct point tracks accurately from arbitrary initialization even with high missing rate of observations.

2. Bundle adjustment (BA) for uncalibrated camera models (using a unified notation)

Given some 2D point tracks, find a set of camera parameters $\{\mathbf{P}_i\}$ and the 3D structure $\{\tilde{\mathbf{x}}_j\}$ which minimizes

$$\sum_{(i,j)\in\Omega} \left\|\pi(\mathbf{P}_i\tilde{\mathbf{x}}_j) - \mathbf{m}_{ij}\right\|_2^2$$

where $\mathbf{m}_{ij} \in \mathbb{R}^2$ is the observed projection of point j in frame i,

$\mathbf{P}_i := \mathbf{P}(\mathbf{p}_i, \mathbf{q}_i, s_i, \mu_i) := \begin{bmatrix} p_{i1} & p_{i2} & p_{i3} & p_{i4} \\ p_{i5} & p_{i6} & p_{i7} & p_{i8} \\ \mu_i q_{i1} & \mu_i q_{i2} & \mu_i q_{i3} & s_i \end{bmatrix}$ is the camera matrix at frame i,

$\tilde{\mathbf{x}}_j := \tilde{\mathbf{x}}(\mathbf{x}_j, t_j) := [x_{j1} \quad x_{j2} \quad x_{j3} \quad t_j]^\top$ is the homogeneous representation of point j,

$\boldsymbol{\pi}([x \quad y \quad z]^\top) := [x/z \quad y/z]^\top$, and

Ω denotes the set of visible observations.

Form	Variable	Affine model ($\mu_i = 0$)	Projective model ($\mu_i = 1$)
Homogeneous	Camera ($\mathtt{P}_i$)	$\mathtt{P}(\mathbf{p}_i, 0, s_i, 0)$	$\mathtt{P}(\mathbf{p}_i, \mathbf{q}_i, s_i, 1)$
	Point ($\tilde{\mathbf{x}}_j$)	$\tilde{\mathbf{x}}(\mathbf{x}_j, t_j)$	$\tilde{\mathbf{x}}(\mathbf{x}_j, t_j)$
	Inverse depth	$s_i t_j$	$\mathbf{q}_i^\top \mathbf{x}_j + s_i t_j$
	Model property	Nonlinear in $\mathtt{P}_i$ and $\tilde{\mathbf{x}}_j$	
	Gauge freedom	$\mathtt{P}_i\tilde{\mathbf{x}}_j = (\mathtt{P}_i\mathtt{H})(\mathtt{H}^{-1}\tilde{\mathbf{x}}_j)$	
	Scale freedom	$\boldsymbol{\pi}(\mathtt{P}_i, \tilde{\mathbf{x}}_j) = \boldsymbol{\pi}(\alpha_i\mathtt{P}_i, \beta_j\mathbf{x}_j)$	
Inhomogeneous	Camera ($\mathtt{P}_i$)	$\mathtt{P}(\mathbf{p}_i, 0, 1, 0)$	$\mathtt{P}(\mathbf{p}_i, \mathbf{q}_i, 1, 1)$
	Point ($\tilde{\mathbf{x}}_j$)	$\tilde{\mathbf{x}}(\mathbf{x}_j, 1)$	$\tilde{\mathbf{x}}(\mathbf{x}_j, 1))$
	Inverse depth	1	$\mathbf{q}_i^\top \mathbf{x}_j + 1$
	Model property	linear in $\mathtt{P}_i$ and $\tilde{\mathbf{x}}_j$	nonlinear in $\mathtt{P}_i$ and $\tilde{\mathbf{x}}_j$
	Gauge freedom	$\mathtt{P}_i\tilde{\mathbf{x}}_j = (\mathtt{P}_i\mathtt{A})(\mathtt{A}^{-1}\tilde{\mathbf{x}}_j)$	$\mathtt{P}_i\tilde{\mathbf{x}}_j = (\mathtt{P}_i\mathtt{H})(\mathtt{H}^{-1}\tilde{\mathbf{x}}_j)$
	Scale freedom	None	

3. The Variable Projection Method (VarPro)

a. Linear VarPro [1]

VarPro is a method designed for solving separable nonlinear least squares, e.g.

$$\min_{\mathbf{u},\mathbf{v}} \|\boldsymbol{\varepsilon}(\mathbf{u},\mathbf{v})\|_2^2 = \min_{\mathbf{u},\mathbf{v}} \|\mathbf{A}(\mathbf{u})\mathbf{v} - \mathbf{m}\|_2^2$$

where $\mathbf{m}$ is the measurement vector,
$\mathbf{u}$ and $\mathbf{v}$ are the model parameters and
$\mathbf{A}(\mathbf{u})$ is a matrix which depends on $\mathbf{u}$ only.

Since $\boldsymbol{\varepsilon}(\mathbf{u},\mathbf{v})$ is linear in $\mathbf{v}$, we can obtain a closed-form minimizer for $\mathbf{v}$,

$$\mathbf{v}^*(\mathbf{u}) := \min_{\mathbf{v}} \|\boldsymbol{\varepsilon}(\mathbf{u},\mathbf{v})\|_2^2 = \min_{\mathbf{v}} \|\mathbf{A}(\mathbf{u})\mathbf{v} - \mathbf{m}\|_2^2 = \mathbf{A}^\dagger(\mathbf{u})\mathbf{m}.$$

VarPro uses a Newton-like optimizer to solve the reduced problem

$$\min_{\mathbf{u}} \left\|\boldsymbol{\varepsilon}(\mathbf{u}, \mathbf{v}*(\mathbf{u}))\right\|_2^2 = \min_{\mathbf{u}} \left\|\mathbf{A}(\mathbf{u})\mathbf{A}^\dagger(\mathbf{u})\mathbf{m} - \mathbf{m}\right\|_2^2.$$

Depending on the approximation used for the Hessian, we have the following algorithms [3]:

- **RW1**, which uses the Gauss-Newton matrix,
- **RW2**, which uses an approximated Gauss-Newton matrix [3], and
- **RW3**, which is alternation [2].

b. Nonlinear VarPro [4]

Strelow extends VarPro to solving non-separable nonlinear least squares. e.g. $\min_{\mathbf{u},\mathbf{v}} \|\boldsymbol{\varepsilon}(\mathbf{u},\mathbf{v})\|_2^2$.

- For initial $\mathbf{u}$, we first compute $\hat{\mathbf{v}}$, which is an estimate of $\mathbf{v}^*(\mathbf{u}) := \min_{\mathbf{v}} \|\boldsymbol{\varepsilon}(\mathbf{u},\mathbf{v})\|_2^2$, using an iterative optimizer.
- Linearizing the residual in $\mathbf{v}$ at $(\mathbf{u},\hat{\mathbf{v}})$ gives $\boldsymbol{\varepsilon}(\mathbf{u},\hat{\mathbf{v}}+\Delta\mathbf{v}) \approx \boldsymbol{\varepsilon}(\mathbf{u},\hat{\mathbf{v}}) + \mathbf{J}_\mathbf{v}(\mathbf{u},\hat{\mathbf{v}})\Delta\mathbf{v}$, where $\mathbf{J}_\mathbf{v}(\mathbf{u},\hat{\mathbf{v}}) := \partial\boldsymbol{\varepsilon}(\mathbf{u},\hat{\mathbf{v}})/\partial\mathbf{v}$.
- At each iteration, we perform:
  1. 1 iteration of linear VarPro to solve $\min_{\mathbf{u},\Delta\mathbf{v}} \|\boldsymbol{\varepsilon}(\mathbf{u},\hat{\mathbf{v}}) + \mathbf{J}_\mathbf{v}(\mathbf{u},\hat{\mathbf{v}})\Delta\mathbf{v}\|_2^2$, then
  2. iterative computation of new $\hat{\mathbf{v}}$, which estimates $\min_{\mathbf{v}} \|\boldsymbol{\varepsilon}(\mathbf{u}+\Delta\mathbf{u},\mathbf{v})\|_2^2$.

We propose the RW counterparts for the nonlinear extension of VarPro.

4. Proposed algorithms

- None of the Projective BA algorithms succeeds from arbitrary cameras and points.
- VarPro-based algorithms for Affine BA have a large convergence basin.
- Hence, we propose to first perform Affine BA using VarPro then warm-start Projective BA.

a. BA algorithms for uncalibrated camera models

ID	Camera	Form	Strategy	Algorithm
AHRW2P	Affine	Homogeneous	Nonlinear VarPro	RW2
AIRW2P	Affine	Inhomogeneous	Linear VarPro	RW2
PHRW1P	Projective	Homogeneous	Nonlinear VarPro	RW1
PHJP	Projective	Homogeneous	Joint optimization	LM

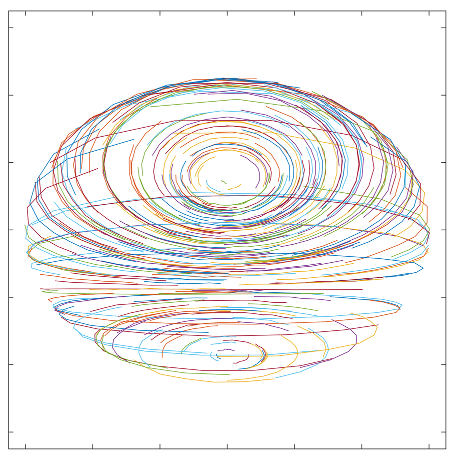
b. Two-stage meta-algorithms (TSMAs)

ID	First-stage (affine) algorithm	Second-stage (projective) algorithm
TSMA1	AHRW2P	PHRW1P
TSMA2	AIRW2P	PHRW1P
TSMA3	AHRW2P	PHJP
TSMA4	AIRW2P	PHJP

5. Datasets

a. Synthetic datasets

Dataset	d	Loop closed	Missing (%)	Best affine cost	Best projective cost
S30L	30.0	Yes	77.56	2.993821	0.861392
S30	30.0	No	76.92	2.947244	0.842539
S20L	20.0	Yes	77.61	7.261506	0.862520
S20	20.0	No	76.92	7.157783	0.842511
S13L	13.0	Yes	77.61	25.100919	0.863871
S13	13.0	No	76.92	24.831476	0.844125
S12L	12.0	Yes	77.61	34.871149	0.863867
S12	12.0	No	76.92	34.730023	0.844817
S11L	11.0	Yes	77.61	55.782547	0.863274
S11	11.0	No	76.92	56.946272	0.845271
S10.5L	10.5	Yes	77.61	80.700734	0.862545
S10.5	10.5	No	76.92	85.970046	0.844771



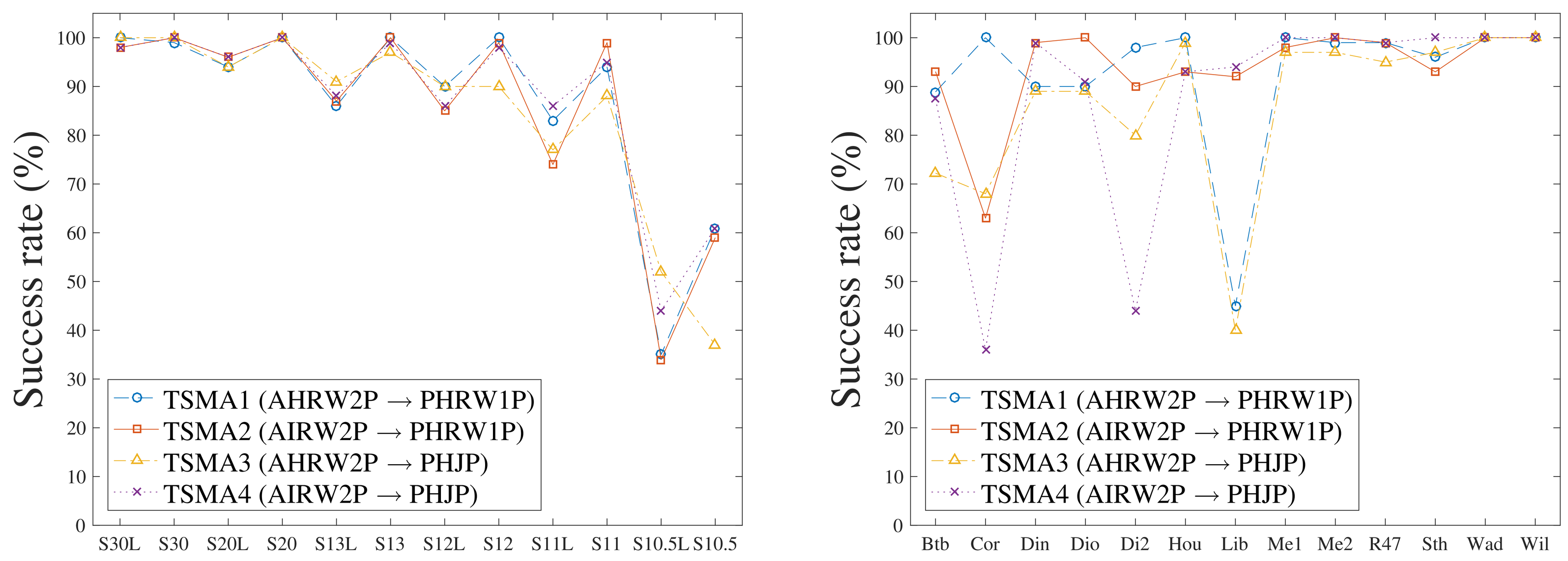
(e) S12 tracks

b. Real datasets

ID	Dataset	f	n	Missing (%)	Best affine	Best projective
Btb	Blue teddy bear (trimmed) (Ponce)	196	827	80.71	0.530633	0.489283
Cor	Corridor (VGG)	11	737	50.23	2.237213	0.272462
Din	Dinosaur (trimmed) [5]	36	319	76.92	1.270153	1.114493
Dio	Dinosaur (VGG)	36	4983	90.84	1.217574	1.166165
Di2*	Dinosaur (trimmed) closer cameras	36	319	76.92	9.380473	0.838414
Hou	House (VGG)	10	672	57.65	2.750877	0.441660
Lib	Oxford University Library (VGG)	3	667	29.24	4.180297	0.172830
Me1	Merton College 1 (VGG)	3	717	22.13	3.176176	0.118450
Me2	Merton College 2 (VGG)	3	475	21.61	3.995869	0.158851
R47	Road scene point tracks #47	11	150	47.09	4.402777	3.344768
Sth	Stockholm Guildhall (trimmed) [18]	43	1000	18.01	8.833195	5.619975
Wad	Wadham College (VGG)	5	1331	54.64	3.424812	0.135711
Wil	Wilshire (Ponce)	190	411	60.73	2.703663	0.423892

6. Results and conclusion

- Each run is initialized from arbitrary cameras and points.
- We evaluate 100 runs per dataset per two-stage meta-algorithm (TSMA, §4).
- TSMA1 and TSMA2 return global optimum in a large fraction of runs on most datasets.
- The convergence basin can be greatly enhanced using the right combination of methods.



References

[1] G. H. Golub and V. Pereyra. The Differentiation of Pseudo-Inverses and Nonlinear Least Squares Problems Whose Variables Separate. SINUM, 1973.
[2] J. H. Hong and A. W. Fitzgibbon. Secrets of Matrix Factorization: Approximations, Numerics, Manifold Optimization and Random Restarts. ICCV, 2015.
[3] A. Ruhe and P.A. Wedin. Algorithms for Separable Nonlinear Least Squares Problems. SIREV, 1980.
[4] D. Strelow. General and Nested Wiberg Minimization: L2 and Maximum Likelihood. ECCV, 2012.

Acknowledgement The work was supported by Microsoft, Toshiba Research Europe and JNE Systech.

The conference travel was supported by the British Machine Vision Association (BMVA), Cambridge University Engineering Department (Ford of British Trust) and Christ's College (University of Cambridge).

Project repository https://github.com/jhh37/projective-ba/

** Much of the work was done while the first author was an intern at Toshiaba Research Europe.*