

## Supplementary Material

Anonymous ICCV submission

Paper ID

### 1. Derivations of Computational Complexity

	EKF		SRIF	
	Step	Complexity	Step	Complexity
<b>Update</b>				
	$\mathbf{PH}^T$	$288N^2$	$\mathbf{A} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$	$144N^3$
	$\mathbf{S} = \mathbf{HPH}^T + \mathbf{Q}_v$	$100N^2$	$\mathbf{Q}^T \Delta \mathbf{e}_{k+1}$	$120N^2$
	$\mathbf{S} = \mathbf{R}_s^T \mathbf{R}_s$	$8N^3/3$	$\mathbf{R}^{-1} \mathbf{Q}^T \Delta \mathbf{e}_{k+1}$	$36N^2$
	$\mathbf{K} = \mathbf{PH}^T \mathbf{R}_s^{-1} \mathbf{R}_s^{-T}$	$48N^3$		
	$\mathbf{P} = \mathbf{P} - \mathbf{KHP}$	$144N^3$		
<b>Predict</b>				
		Negligible		Negligible
<b>Feature removal</b>				
			$\text{QR}(\mathbf{R}_\Pi)$	$324Nk(N-k)$ $162N^{3\dagger}$
<b>Total</b>				
		$194.67N^3$		$306N^{3\dagger}$

Table 1. Computational complexity of EKF and SRIF, <sup>†</sup>Upper Bound.

Derivations for the estimates of computational complexity in Table 2 of the main paper are presented here. The table is reproduced here for convenience of the reader. Various results from the classic text of Golub and Van Loan are used [1].

- Computing  $\mathbf{AB}$  where  $\mathbf{A} \in \mathbb{R}^{m \times k}$  and  $\mathbf{B} \in \mathbb{R}^{k \times n}$  requires  $2mnk$  ops.
- The Cholesky factorization  $\mathbf{A} = \mathbf{R}^T \mathbf{R}$ , where  $\mathbf{A} \in \mathbb{R}^{n \times n}$  is symmetric positive-definite, requires  $n^3/3$  ops.
- Solving  $\mathbf{R}\mathbf{x} = \mathbf{y}$  for upper/lower triangular  $\mathbf{R} \in \mathbb{R}^{n \times n}$   $\mathbf{R}$  requires  $n^2$  ops.
- Computing the QR factorization  $\mathbf{A} = \mathbf{QR}$ ,  $\mathbf{A} \in \mathbb{R}^{m \times n}$  requires  $2n^2(m-n/3)$  ops. Note that the  $\mathbf{Q}$  matrix is never calculated explicitly but is represented by the Householder vectors obtained during factorization.

#### Complexity of $\mathbf{PH}^T$

Direct multiplication of  $\mathbf{P} \in \mathbb{R}^{6N \times 6N}$  and  $\mathbf{H} \in \mathbb{R}^{2N \times 6N}$  will require  $2 \times 2N \times 6N \times 6N = 144N^3$  ops. Note, however, that  $\mathbf{H}$  is sparse and has only 12 non-zero elements per row. Hence, the number of ops reduces to  $2 \times 2N \times 12 \times 6N = 288N^2$ .

#### Complexity of $\mathbf{S} = \mathbf{HPH}^T + \mathbf{Q}_v$

$\mathbf{PH}^T$ , which has already been calculated above, is of dimension  $6N \times 2N$ . Direct multiplication of  $\mathbf{H}$  with  $\mathbf{PH}^T$  would require  $2 \times 2N \times 6N \times 2N = 48N^3$  ops. Again, leveraging the sparsity of  $\mathbf{H}$  yields,  $2 \times 2N \times 12 \times 2N = 96N^2$  ops. Addition of  $\mathbf{Q}_v$  requires an additional  $2N \times 2N = 4N^2$  ops bringing the total to  $100N^2$  ops.

**Complexity of  $\mathbf{S} = \mathbf{R}_s^T \mathbf{R}_s$** 

Cholesky of  $\mathbf{S} \in \mathbb{R}^{2N \times 2N}$ , requires  $(2N)^3/3 = 8N^3/3$  ops.

**Complexity of  $\mathbf{K} = \mathbf{P}\mathbf{H}^T \mathbf{R}_s^{-1} \mathbf{R}_s^{-T}$** 

Calculating  $\mathbf{K}$  requires solving  $\mathbf{P}\mathbf{H}^T$ , a  $6N \times 2N$  matrix, by an upper-triangular matrix,  $\mathbf{R}_s \in \mathbb{R}^{2N \times 2N}$ , followed by its transposed lower-triangular version  $\mathbf{R}_s^T$ . Consider first, the lower-triangular solving. Solving one row of  $\mathbf{P}\mathbf{H}^T$  by  $\mathbf{R}_s$  requires  $2N \times 2N$  ops. Hence, solving all  $6N$  rows of  $\mathbf{P}\mathbf{H}^T$  requires  $24N^3$  ops. Similarly, the upper-triangular solving will require another  $24N^3$  ops, bringing the total to  $48N^3$  ops.

**Complexity of  $\mathbf{P} = \mathbf{P} - \mathbf{K}\mathbf{H}\mathbf{P}$** 

Here, it is assumed that  $\mathbf{H}\mathbf{P}$  is already available by transposing  $\mathbf{P}\mathbf{H}^T$  above. Multiplication of  $\mathbf{K}$  by  $\mathbf{H}\mathbf{P}$  requires  $2 \times 6N \times 2N \times 6N = 144N^3$ . Subtracting from  $\mathbf{P}$  is an  $O(N^2)$  operation, which is negligible.

**Complexity of  $\mathbf{A} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$** 

Matrix  $\mathbf{A}$  is a  $8N \times 6N$  matrix and has the following form (time subscripts dropped for notational convenience)

$$\mathbf{A} = \begin{bmatrix} \mathbf{R}^x \\ \mathbf{R}^v \mathbf{H} \end{bmatrix} = \begin{bmatrix} \times & \times & \cdots & \times \\ 0 & \times & \cdots & \times \\ 0 & 0 & \ddots & \times \\ 0 & 0 & \cdots & \times \\ \times & \times & \cdots & \times \\ \times & \times & \cdots & \times \end{bmatrix}$$

where,  $\times$  denotes a non-zero element. Computing the QR-factorization of  $\mathbf{A}$  would normally require approximately  $4 \times 8N \times (6N)^2 - (6N)^3/3 = 1080N^3$  ops. However, the upper  $6N \times 6N$  portion of  $\mathbf{A}$ , i.e  $\mathbf{R}^x$ , is already upper-triangular. In order to transform it into  $\begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix}$  by QR factorization, therefore, only the lower  $2N \times 6N$  portion of  $\mathbf{A}$  needs to be zeroed out. A sequence of  $6N$  Householder transformations are applied. The  $l$ th transformation operates on the lower  $(8N - l + 1) \times (6N - l + 1)$  sub-matrix  $\mathbf{A}_l$  of  $\mathbf{A}$ . The Householder vector,  $\mathbf{v}_l$ , therefore, has length  $8N - l + 1$ , but it has only  $2N + 1$  non-zero elements. Hence, in the computation of the Householder transformation

$$(\mathbf{I} - \beta_l \mathbf{v}_l \mathbf{v}_l^T) \mathbf{A}_l = \mathbf{A}_l - \beta_l \mathbf{v}_l (\mathbf{v}_l^T \mathbf{A}_l),$$

$2(1 + 2N)(6N - l)$  ops are required for the computation of  $\mathbf{v}_l^T \mathbf{A}_l$ ; a further  $(1 + 2N)(6N - l)$  ops are required for the computation of the outer-product  $\beta_l \mathbf{v}_l (\mathbf{v}_l^T \mathbf{A}_l)$ ; and finally  $(1 + 2N)(6N - l)$  ops are required for the computation of the difference  $\mathbf{A}_l - \beta_l \mathbf{v}_l (\mathbf{v}_l^T \mathbf{A}_l)$ , which totals  $4(1 + 2N)(6N - l)$ . Summing this, yields

$$\sum_{l=0}^{6N-1} 4(1 + 2N)(6N - l) \approx 2 \times 2N \times (6N)^2 = 144N^3,$$

which is a significant reduction.

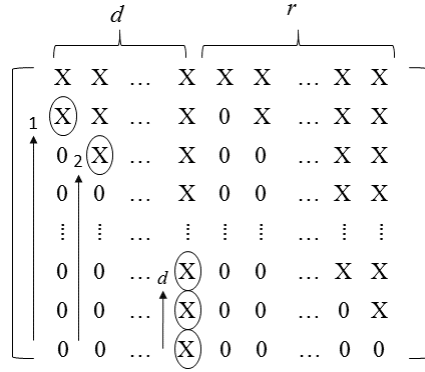
**Complexity of  $\mathbf{Q}^T \Delta \mathbf{e}_{k+1}$** 

From [1, Chapter 5], calculation of  $\mathbf{Q}^T \mathbf{C}$ , where  $\mathbf{Q} \in \mathbb{R}^{n \times n}$  is a product of  $r$  Householder matrices, and  $\mathbf{C} \in \mathbb{R}^{n \times q}$ , requires  $2qr(2n - r)$  ops rather than  $2n^2q$  ops that would be required for direct multiplication. In this case,  $r = 6N$ ,  $n = 8N$ , and  $q = 1$ . Hence, number of ops is  $2 \times 6N \times (16N - 6N) = 120N^2$ .

**Complexity of  $\mathbf{R}^{-1} \mathbf{Q}^T \Delta \mathbf{e}_{k+1}$** 

This is simply solving of a vector,  $\mathbf{Q}^T \Delta \mathbf{e}_{k+1}$ , by a lower-triangular matrix  $\mathbf{R}^{-1}$  which is  $6N \times 6N$ . This requires  $36N^2$  ops.

216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269



270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323

Figure 1. EKF-based Structure-from-Motion.

### Complexity of triangularizing $\mathbf{R}_{\Pi}$

The cost of feature-reduction step in SRIF-SfM comes from triangularizing  $\mathbf{R}_{\Pi}$ . Suppose, for simplicity, that  $\mathbf{x} \in \mathbb{R}^n$ ,  $\mathbf{R} \in \mathbb{R}^{n \times n}$ ,  $d$  elements from  $\mathbf{x}$  are discarded and  $r = n - d$  elements are retained. Recall, from the main manuscript, that  $\mathbf{R}_{\Pi} = \mathbf{R}\mathbf{\Pi}^T$ , where  $\mathbf{\Pi}$  is a permutation matrix; hence  $\mathbf{R}_{\Pi}$  is a column-permutation of an upper-triangular  $\mathbf{R}$ . The permutation matrix  $\mathbf{\Pi}$  is such that the columns corresponding to the  $d$  elements to be discarded are placed first, followed by the  $r$  columns corresponding to the elements to be retained. The general structure of  $\mathbf{R}_{\Pi}$  is shown in Figure 1. Triangularizing  $\mathbf{R}_{\Pi}$  entails applying a sequence of Givens rotations in the order shown in the figure to zero out the elements below the principal diagonal. It is clear that greater the number of non-zero elements below the principal diagonal, more will be the number of Givens rotations required to triangularize  $\mathbf{R}_{\Pi}$ . In general, this number will depend not just on the number of elements to be discarded from  $\mathbf{x}$ , but on the specific order of the elements being removed. Given this dependency, it is not possible to derive a closed-form expression for the number of operations required. However, it is possible to obtain an expression for the upper bound.

The upper-bound will occur when the last  $d$  elements of  $\mathbf{x}$  need to be removed, in which case the permutation matrix  $\mathbf{\Pi}$  would be such that the last  $d$  columns of  $\mathbf{R}$  are permuted to the first  $d$  positions. In this case, it is required to zero out the lowest  $n - d = r$  non-zero elements in the first  $d$  columns. Consider the lowest non-zero element in the first column. Zeroing this out via a Givens rotation will impact the corresponding elements in the last  $d$  columns. Thus the  $2 \times 2$  Givens rotation would have effectively been applied to  $d + 1$  columns. Next, we proceed to the element above this. Zeroing this out would impact the corresponding elements in the last  $d$  columns, resulting in effectively applying the Givens rotation to  $d + 2$  columns. Proceeding thus to the element just below the topmost in the first column, it can be seen that zeroing this would involve Givens rotation to  $n$  columns. Thus for the 1st column, the total number of Givens rotations involved are:

$$\sum_{i=d+1}^n i \quad (1)$$

Proceeding similarly for the  $j$ th column yields the total number of Givens rotation as

$$\sum_{i=d-j+2}^{n-j+1} i \quad (2)$$

Thus, the total number of Givens rotations for the first  $d$  columns is

$$\sum_{j=1}^d \left( \sum_{i=d-j+2}^{n-j+1} i \right) = \frac{d(n-d)(n+2)}{2} \quad (3)$$

The above expression attains a maximum at  $d = n/2$  and its value is  $\approx n^3/8$ . Now, each Givens rotation requires 6 ops [1, chapter 5]. Hence, total number of ops is  $3n^3/4$ . Since  $n = 6N$  in our case, the number of ops is

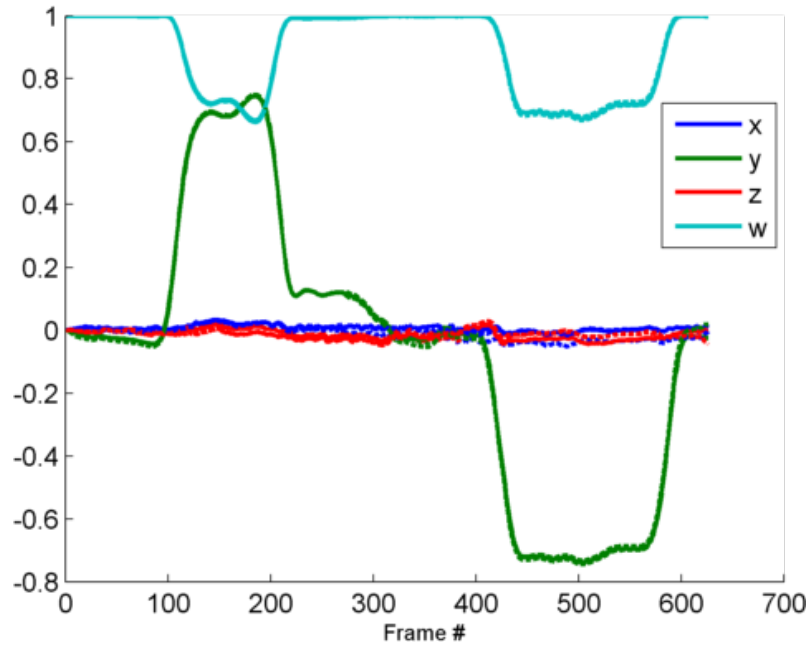
$$3 \frac{(6N)^3}{4} = 162N^3$$

324 **2. Dataset and Results** 378

325  
326 In this section we present the *pose-only* plots for ten Kitti odometry clips for which pose and position errors are given in 379  
327 Table 1 of the manuscript. The pose is represented by a 4-element quaternion  $\mathbf{q} \triangleq (x, y, z, w)$ . Thus in the following plots, 380  
328 the four elements of the pose-quaternion are plotted against the frame index. For the selected ten Kitti clips, the ground-truth 381  
329 information is available. We generated outputs for four different cases: EKF single precision, EKF 64bit fixed-point, SRIF 382  
330 single precision, and SRIF 32bit fixed-point implementations. The errors in the pose quaternion are a good indication of the 383  
331 quality of the quality of ego-motion tracking and also the estimated 3D scene structure. Since our algorithm is a monocular 384  
332 SfM, the scene structure and self-position values suffer from scale ambiguity. Hence the position plots do not lend themselves 385  
333 to an intuitive visual comparison. We have compared the ground-truth pose (solid line) values with the four implementations 386  
334 387

- 335 1. EKF single precision (long-dash) 388  
336 2. EKF 64bit fixed-point (dots and dash) 389  
337 3. SRIF single precision and (dots) 390  
338 4. SRIF 32bit fixed-point implementations (short dash) 391  
339 392  
340 393  
341 394

342 It can be observed from the plots that the estimated outputs are match so closely with the ground-truth as to be largely 395  
343 indistinguishable from it. A slightly higher deviation in the estimated values for sequence 02 (Figure 3) results in higher 396  
344 position errors (Table 1 in the manuscript). 397  
345 398



346 400  
347 401  
348 402  
349 403  
350 404  
351 405  
352 406  
353 407  
354 408  
355 409  
356 410  
357 411  
358 412  
359 413  
360 414  
361 415  
362 416  
363 417  
364 418  
365 419  
366 420  
367 421  
368 422  
369 423  
370 424  
371 425  
372 426  
373 427  
374 428  
375 429  
376 430  
377 431

Figure 2. Pose results for kitti odometry clip 00 (Seq#1 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

370 **References** 424

371 [1] G. Golub and C. V. Loan. *Matrix Computations*. Johns Hopkins University Press, third edition, 1996. 1, 2, 3 425  
372 426  
373 427  
374 428  
375 429  
376 430  
377 431

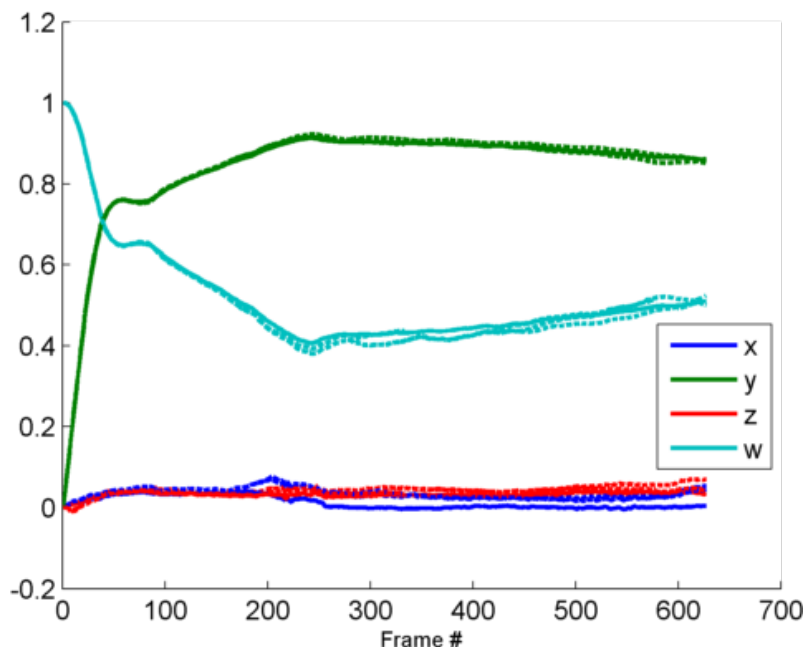


Figure 3. Pose results for kitti odometry clip 01 (Seq#2 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

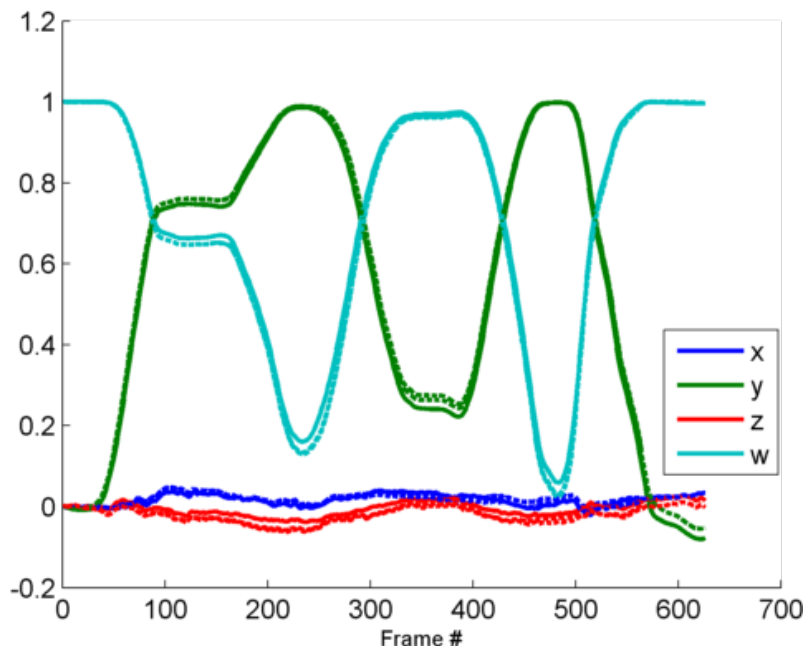


Figure 4. Pose results for kitti odometry clip 02 (Seq#3 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593

594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647

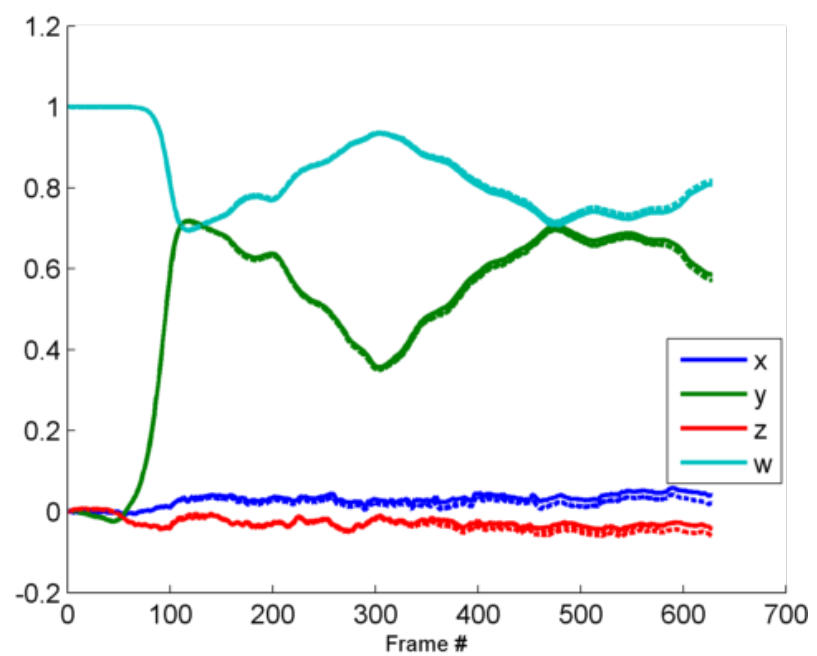


Figure 5. Pose results for kitti odometry clip 03 (Seq#4 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

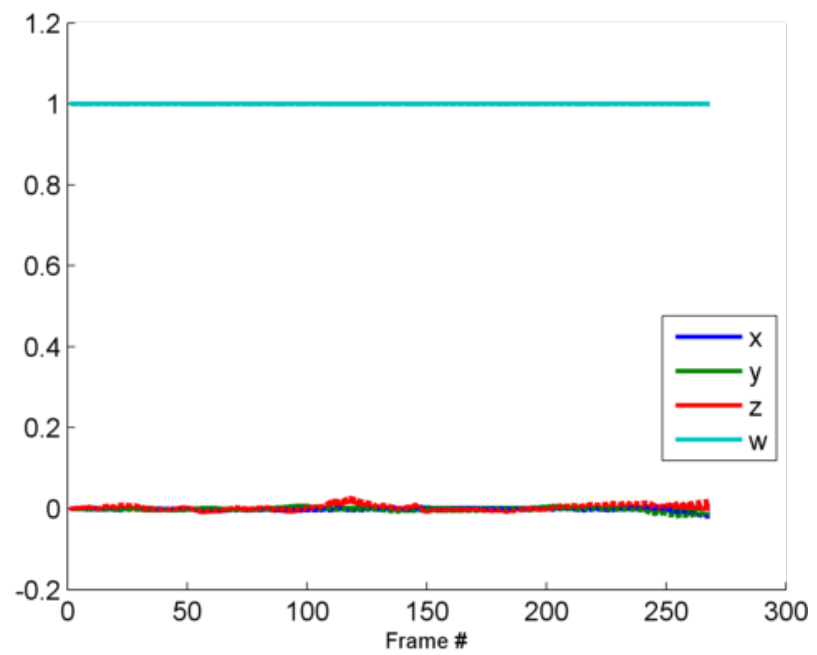


Figure 6. Pose results for kitti odometry clip 04 (Seq#5 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701

702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755

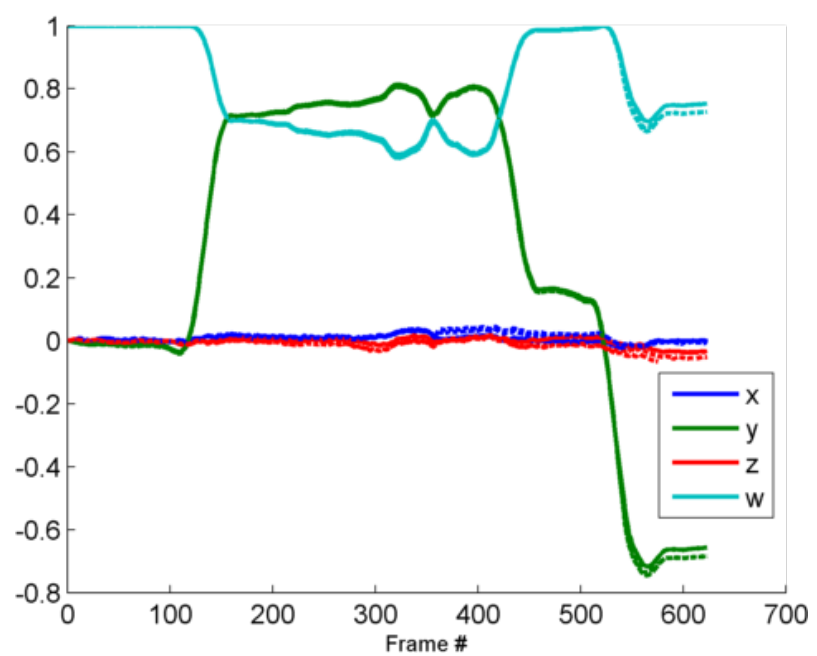


Figure 7. Pose results for kitti odometry clip 05 (Seq#6 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

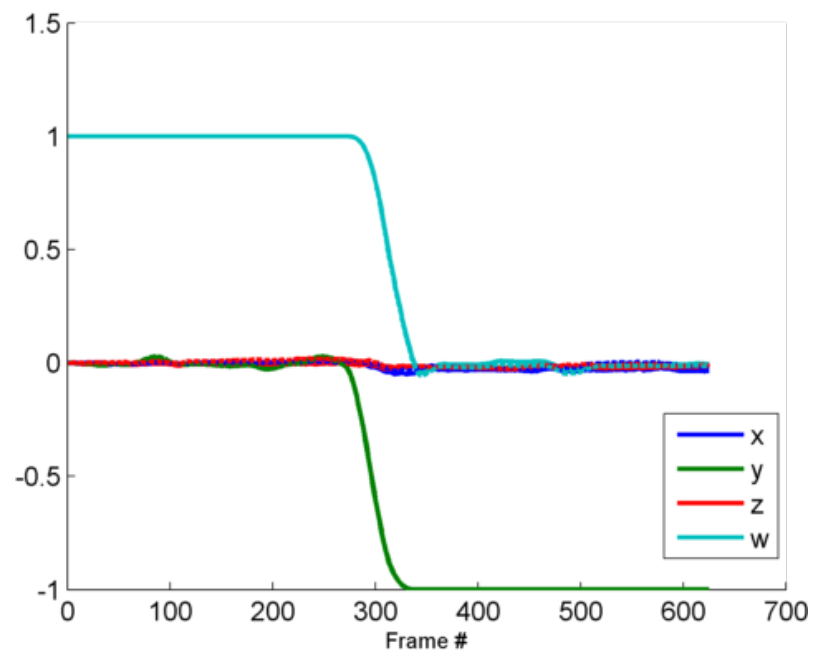


Figure 8. Pose results for kitti odometry clip 06 (Seq#7 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809

810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863

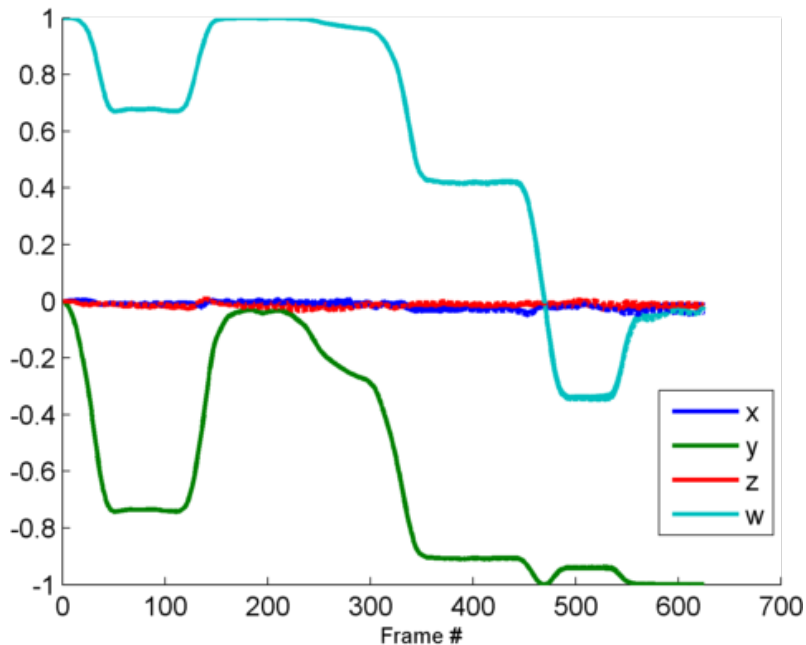


Figure 9. Pose results for kitti odometry clip 07 (Seq#8 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).

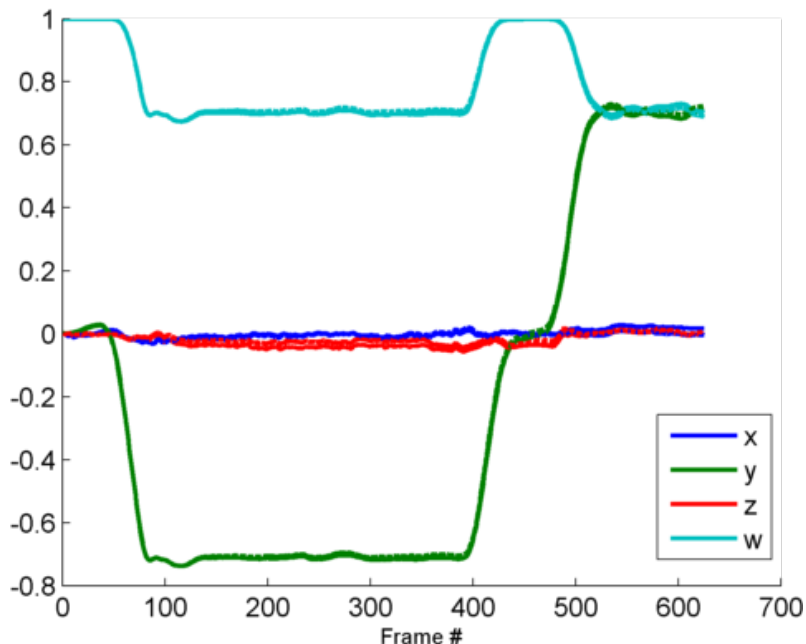


Figure 10. Pose results for kitti odometry clip 08 (Seq#9 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).



864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917

918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971

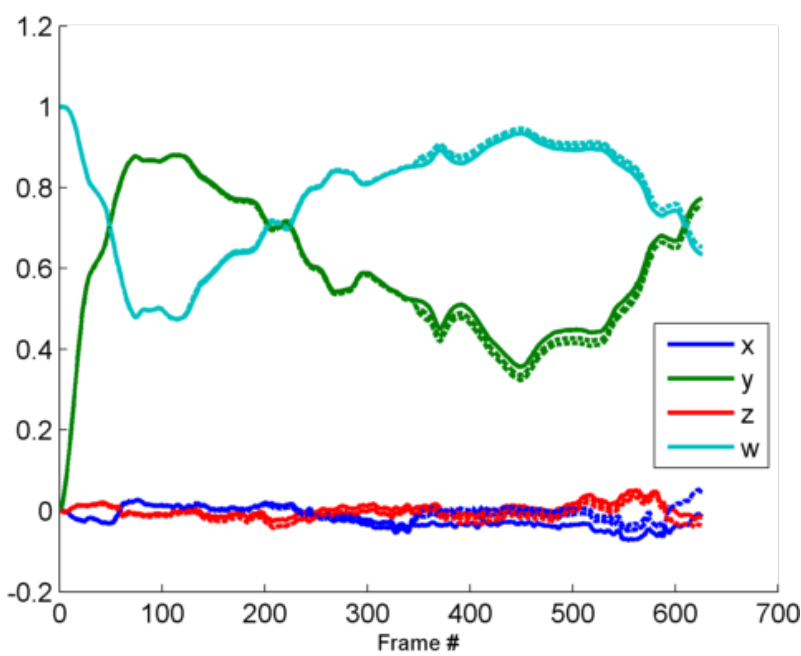


Figure 11. Pose results for kitti odometry clip 10 (Seq#10 in Table 2). The solid-lines are for the ground-truth pose values. The dotted-lines are for the estimated pose values from EKF single precision, EKF 64bit fixed-point, SRIF single precision and SRIF 32bit fixed-point implementations (in most cases overlapping with the ground-truth plots).