VSAC: Efficient and Accurate Estimator for H and F (Supplementary material)

Maksym Ivashechkin[1], Daniel Barath[2], and Jiri Matas[1]
[1] Centre for Machine Perception, Czech Technical University in Prague, Czech Republic
[2] Computer Vision and Geometry Group, Department of Computer Science, ETH Zürich
{ivashmak, matas}@cmp.felk.cvut.cz dbarath@ethz.ch

1. Point Correction

We demonstrate how correcting the ground truth point correspondences, as proposed in Section 6, affects the results of the tested methods. To do so, we corrected the ground truth correspondences provided in datasets EVD and HPatches (homography estimation), and in Kusvod2 and PhotoTour (fundamental matrix estimation). The results of the methods for homography and fundamental matrix estimation are shown in Table 1.

In all cases, using the ground truth corrected by being projected to the model manifold, reduces the median and average errors of the tested method, allowing more accurate comparison. For $\varepsilon_{\max}$, the error is dominated by inaccuracies of the estimated model and the relatively small change between provided and corrected GT points randomly changes the error in either direction, either + or -, by a small amount.

As expected, the corrected correspondences have zero cross-validation (X-val) error – all the corrected points are consistent with an **H** or **F** model, and this model is recovered in this pseudo-noise free setting, regardless of the point left out. For **H** estimation, the errors $\varepsilon_{\text{avg}}, \varepsilon_{\text{med}}$ dropped by about 0.1-0.2 pixels, which is a reasonable value for the positional noise of GT points. For PhotoTour, the GT points were selected from image correspondences perfectly fitting a model estimated from hundreds of points; their correction is minimal. For Kusvod2, the error is reduced by 0.01-0.07 pixels. Note that this is a 1D geometric error w.r.t. **F**, not euclidean in 2D as in homography estimation. These results confirm that the cross-validation error, X-val (provided) is a loose upper bound on the real error.

The ordering of the methods used for homography estimation became clearer than one the provided ground truth points – VSAC with MAGSAC++ ($\text{VSAC}_{\text{MGSC}}$) is always the most accurate and MAGSAC++ is then second most accurate method. For fundamental matrix estimation, ORSA provides the most accurate results on the PhotoTour dataset, but the difference is negligible, only 0.01-0.02 of a pixel w.r.t. $\text{VSAC}_{\text{MGSC}}$ which is the second most accurate algorithm. On Kusvod2, $\text{VSAC}_{\text{MGSC}}$ has the lowest errors.

2. Gauss Elimination for Fundamental Matrix

The estimation of the fundamental matrix from seven point correspondences, consists of two main steps. First, constraint $p_2^{\text{T}}\mathbf{F}p_1 = 0$ that each correspondence imply is used to build a linear system $Af = 0$, where p_i is the point in the ith image, **F** is the fundamental matrix, A is the coefficient matrix of the system and f contains the elements of **F** in vector form [1]. Coefficient matrix A is of size 7×9. Gaussian Elimination is then used to make A an upper triangular matrix as follows:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} & a_{17} & a_{18} & a_{19} \\ 0 & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} & a_{27} & a_{28} & a_{29} \\ 0 & 0 & a_{33} & a_{34} & a_{35} & a_{36} & a_{37} & a_{38} & a_{39} \\ 0 & 0 & 0 & a_{44} & a_{45} & a_{46} & a_{47} & a_{48} & a_{49} \\ 0 & 0 & 0 & 0 & a_{55} & a_{56} & a_{57} & a_{58} & a_{59} \\ 0 & 0 & 0 & 0 & 0 & a_{66} & a_{67} & a_{68} & a_{69} \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{77} & a_{78} & a_{79} \end{pmatrix}.$$

Since the fundamental matrix has 8 degrees-of-freedom the two null-vectors can have the last element fixed to one as $f_9^{(1)} = f_9^{(2)} = 1$.

Let us for the first null-vector fix the eighth element to zero $f_8^{(1)} = 0$, thus, seventh element becomes $f_7^{(1)} = -a_{79}/a_{77}$. Similarly, for the second null-vector the seventh element can be fixed to zero $f_7^{(2)} = 0$ and, thus, the eighth one is $f_8^{(2)} = -a_{79}/a_{78}$.

All other values of null-vectors can be found by substituting the previously found elements:

$$f_i^{(\{1,2\})} = \frac{-1}{a_{ii}} \sum_{j=i+1}^{9} f_j^{(\{1,2\})} a_{ij} \qquad \forall i \in \{1, \dots, 6\} \quad (1)$$

The final fundamental matrix is $f = \alpha f^{(1)} + (1-\alpha) f^{(2)}$.

References

[1] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004. 1

		Method	GT	ε_{med}	ε_{avg}	ε_{max}
Homography	HPatches (142 pairs)	VSAC	Provided	0.66	0.99	5.83
			Corrected	0.45	0.82	6.04
		$VSAC_{MGSC}$	Provided	0.65	0.82	3.78
			Corrected	0.41	0.62	3.56
		USACv20	Provided	0.66	0.92	4.05
			Corrected	0.47	0.73	4.16
		USAC	Provided	0.67	5.11	370.28
			Corrected	0.56	5.00	384.48
		OpenCV	Provided	0.76	1.25	10.10
			Corrected	0.62	1.09	9.94
		GC	Provided	0.74	1.12	11.42
			Corrected	0.52	0.89	11.28
		MGSC++	Provided	0.66	0.86	4.91
			Corrected	0.42	0.64	4.81
		ORSA	Provided	0.75	55.74	1105.82
			Corrected	0.76	54.42	1104.78
		X-val	Provided	0.58	0.71	6.94
			Corrected	0.00	0.00	0.00
	EVD (10 pairs)	VSAC	Provided	3.23	3.62	8.99
			Corrected	3.07	3.51	9.92
		$VSAC_{MGSC}$	Provided	2.80	3.37	7.05
			Corrected	2.51	3.27	9.25
		USACv20	Provided	3.26	3.78	10.88
			Corrected	3.00	3.53	11.76
		USAC	Provided	6.56	117.73	474.08
			Corrected	6.31	130.14	485.75
		OpenCV	Provided	3.68	4.53	8.80
			Corrected	3.55	4.22	9.16
		GC	Provided	3.72	4.17	13.28
			Corrected	3.49	4.18	16.84
		MGSC++	Provided	2.85	3.51	7.99
			Corrected	2.56	3.41	10.66
		ORSA	Provided	143.69	170.65	438.44
			Corrected	190.48	181.46	482.97
		X-val	Provided	1.79	1.80	2.29
			Corrected	0.00	0.00	0.00

		Method	GT	ε_{med}	ε_{avg}	ε_{max}
Fundamental matrix	PhotoTour (500 pairs)	VSAC	Provided	0.16	0.18	0.80
			Corrected	0.16	0.18	0.82
		$VSAC_{MGSC}$	Provided	0.15	0.17	0.75
			Corrected	0.15	0.17	0.73
		USACv20	Provided	0.17	0.22	3.44
			Corrected	0.17	0.21	3.43
		USAC	Provided	0.42	0.63	8.01
			Corrected	0.42	0.63	8.03
		OpenCV	Provided	0.39	0.73	25.25
			Corrected	0.39	0.73	25.24
		GC	Provided	0.16	0.25	13.31
			Corrected	0.16	0.25	13.31
		MGSC++	Provided	0.20	0.23	1.49
			Corrected	0.20	0.23	1.48
		ORSA	Provided	0.14	0.15	0.64
			Corrected	0.14	0.15	0.63
		NG-RSC	Provided	0.17	0.18	1.60
			Corrected	0.17	0.18	1.60
		X-val	Provided	0.06	0.06	0.16
			Corrected	0.00	0.00	0.00
	Kusvod2 (15 pairs)	VSAC	Provided	0.55	0.77	3.47
			Corrected	0.51	0.74	3.47
		$VSAC_{MGSC}$	Provided	0.52	0.76	3.47
			Corrected	0.45	0.73	3.47
		USACv20	Provided	0.60	1.01	5.42
			Corrected	0.56	0.98	5.41
		USAC	Provided	2.09	2.85	15.07
			Corrected	2.08	2.84	15.09
		OpenCV	Provided	1.51	6.26	63.05
			Corrected	1.55	6.26	63.06
		GC	Provided	0.55	3.94	48.48
			Corrected	0.54	3.92	48.48
		MGSC++	Provided	0.58	1.18	5.69
			Corrected	0.58	1.16	5.69
		ORSA	Provided	0.51	14.29	307.42
			Corrected	0.49	14.26	307.42
		NG-RSC	Provided	0.48	2.31	50.04
			Corrected	0.46	2.28	50.04
		X-val	Provided	0.91	1.12	2.34
			Corrected	0.00	0.00	0.00

Table 1: The median (ε_{med}), average (ε_{avg}) and maximum (ε_{max}) errors in pixels on the used datasets when using the provided ground truth correspondences and the corrected ones projected to the model manifold as reference inliers. The lowest and second lowest errors are highlighted in red and blue, respectively.