Robust focal length estimation by voting in multi-view scene reconstruction

Martin Bujnak[1], Zuzana Kukelova[2], and Tomas Pajdla[2]

[1] Bzovicka 24, 85107, Bratislava, Slovakia
[2] Center for Machine Perception, Czech Technical University in Prague

Abstract. We propose a new robust focal length estimation method in multi-view structure from motion from unordered data sets, e.g. downloaded from the Flickr database, where *jpeg-exif* headers are often incorrect or missing. The method is based on a combination of RANSAC with weighted kernel voting and can use any algorithm for estimating epipolar geometry and unknown focal lengths. We demonstrate by experiments with synthetic and real data that the method produces reliable focal length estimates which are better than estimates obtained using RANSAC or kernel voting alone and which are in most real situations very close to the ground truth. An important feature of this method is the ability to detect image pairs close to critical configurations or the cases when the focal length can't be reliably estimated [3].

Key words: focal length, epipolar geometry, 3D reconstruction

1 Introduction

Estimating the focal length of an unknown camera is an important computer vision problem with applications mainly in 3D reconstruction. Previously, uncalibrated cameras were used to create a projective 3D reconstruction of the observed scene which was then upgraded to a metric one by enforcing camera properties [8]. Another approach was to first calibrate cameras and then register cameras directly in Euclidean space. This was shown to produce better results even for large scale datasets [23, 18, 17, 24, 19]. Efficient solvers, e.g. the 5-pt relative pose solver for calibrated cameras [21, 26], also helped in developing such Structure from Motion (SFM) pipelines.

An interesting open problem appears with modern digital cameras when the internal parameters [8] except for the focal length are known. Sometimes, it is possible to extract focal lengths from the *jpeg-exif* headers. This was often done in the above mentioned SFM pipelines [23, 18, 17, 24]. Unfortunately, many images downloaded from photo-sharing websites do not contain *jpeg-exif* headers, or listed focal lengths are not correct due to image editing.

A number of algorithms for simultaneous estimation of camera motion and focal length have been invented: the 7-pt or 8-pt algorithm for uncalibrated

---

cameras [8] followed by the extraction of two focal lengths from the fundamental matrix [7, 2, 14], or by the extraction of one focal length common to both cameras [27], or by the extraction of one focal length assuming that the second focal length is known [30], the 6-pt algorithm for cameras with unknown but same focal length [25, 16, 15], the 6-pt algorithm for one unknown and one known focal length [3]. Although these algorithms are well understood and fast, they are rarely used in SFM pipelines. This has mainly the two following reasons.

First, all above mentioned algorithms suffer from some critical configurations, e.g. when optical axes of the cameras are parallel or intersecting [13], or if the scene is planar. In these situations, it is not possible to compute the focal lengths because there exist many Euclidean interpretations of images.

Secondly, every image is usually matched with many different images and therefore one obtains several (often many) candidates for the estimated camera focal length. Mostly these focal lengths are different and one can't select the best one easily. Selecting the focal length with the largest number of inliers or selecting the median or mean focal length do not always produce satisfactory results since estimated geometries may be wrong.

In this paper we propose a new multi-view method for robust focal length estimation based on a combination of RANSAC with weighted kernel voting. Our method can use any focal length extraction algorithm (6-pt, 7-pt, etc.). We follow the paradigm proposed in [16] where a simple kernel voting method was successfully used for estimating focal lengths by the 6-pt algorithm. This method draws 6-tuples of corresponding points, estimates unknown focal lengths and stores them into a vector. Kernel voting is used to smooth data and to select the best focal length after several trials.

A combination of kernel voting method with the RANSAC paradigm was used in [22, 28, 29] to estimate epipoles (resp. camera translations).

Work [22] introduced the idea of splitting the epiplar geometry estimation to first estimating the translation (epipole) and then the rest plus the global uncertainty of the epipolar geometry. A data driven sampling was used to estimate translation candidates. The best model was then selected in a secondary sampling process initialized by the translation candidates.

In [28, 29], votes were not casted directly by each sampled epipolar geometry but by the best epipolar geometries recovered by ordered sampling of PROSAC [11]. The PROSAC with 5000 cycles was run 50 times and its results were collected by the kernel voting. This lead up to 25000 samples but usually terminated much sooner. Here we use a more complex, hybrid, sampling strategy, which turns out to be more efficient than the approach of [28, 29]. In our method, statistics are collected either directly inside a RANSAC loop or in separate sampling process executed on an inliers set returned by a robust RANSAC estimator like DEGENSAC [10]. This is followed by the kernel voting weighted using weights derived from the number of inliers of each particular vote. All reliable votes are accumulated in camera accumulators and contribute to the camera focal length estimation. Finally, camera focal lengths are obtained by kernel voting on votes obtained from all pairwise matchings.
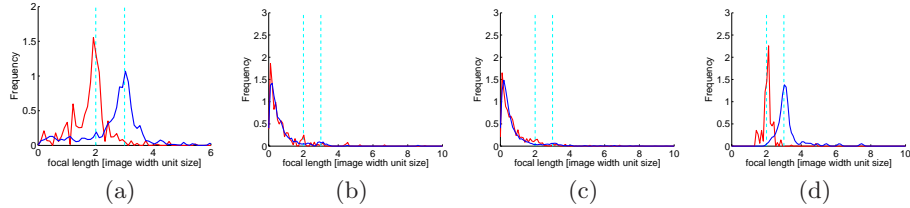
Fig. 1. (a) Standard kernel voting with 500 trials on general outliers free scene. Results for scene with 40% of outliers and (b) 500 trials resp. (c) 5000 trials. (d) Kernel voting weighted by the number of inliers, 5000 trials and 40% of outliers. The results for the left focal length are in blue, for the right focal length in red and the ground truth focal lengths are displayed as cyan vertical lines.

2 Problems in focal length estimation

It is known that RANSAC [5], RANSAC voting [28, 29] and standard kernel voting [16], produce good and reliable estimates of focal lengths for a single image pair in general configuration and under small contaminations by outliers and noise. However, problems occur when we have image pairs close to critical configurations, degenerate scenes, higher numbers of outliers and large noise or when we need to select the camera focal length from several candidates obtained by matching one image with many different images.

Next we describe each from these issues in more detail, show how they affect existing methods and proposed some solutions which will lead to our new method for robust focal length estimation.

2.1 Outliers

RANSAC is robust to outliers since after sufficiently many cycles we get at least one "outlier free" sample which results in a model with the greatest support. The correctness of the best model is, however, not guaranteed.

Large contamination by outliers causes major problems in the standard kernel voting method, see Figure 1 (b). This is because the probability of drawing a good sample dramatically decreases with increasing the number of outliers. Even increasing the number of voting cycles does not solve the problem, since false peaks remain or new appear, see Figure 1 (c).

On the other hand a model estimated from an outlier contaminated sample usually does not have high support. Therefore we weight the vote generated by a sample by the number of inliers supporting the model of the sample. This reduces the influence of outliers and false peaks disappear, Figure 1 (d).

2.2 Noise

Kernel voting as well as RANSAC are immune to contamination by small noise. However, for higher noise levels both methods may deliver wrong focal length estimates. In RANSAC it is not possible to use the size of the support to determine
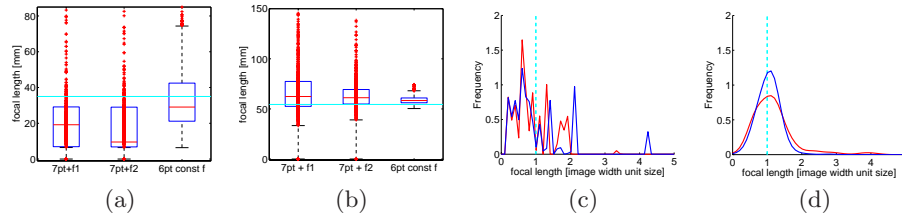
Fig. 2. Real scene in close to critical (a,c) and non-critical (b,d) configurations. (a,b) show boxplots from 1000 runs of DEGENSAC algorithm with focal length extractions. (c,d) show results of weighted kernel voting. Cyan lines are the ground truth values.

if the estimated focal length is reliable or not. For example, critical configurations may result in wrong epipolar geometries with large supports. Hence another methods need to be used to measure the reliability of the result [8].

Kernel voting, on the other hand, provides information about the reliability of the estimated result. It either produces the result as a dominant peak or noise level is too high, which serves as a certificate that the camera pair is not reliable.

Based on these observations we incorporate a detection of cases when the focal length can't be reliably estimated to our method, e.g. due to large noise contamination. We use kernel voting and the estimated focal length is considered reliable only if the highest peak is sufficiently higher than the second highest peak.

2.3   Critical configurations

It is known that critical configurations cause major problems in focal length estimation [13]. If a critical configuration appears, it is not possible to estimate the epipolar geometry and focal length because there exist several (infinite number of) Euclidean interpretations of the structure and camera parameters. Hence we need to detect and reject camera pairs in critical configurations.

Unfortunately, in real situations many critical configurations can't be easily detected. When the camera pair is near the critical configuration, which can't be easily detected, the estimated focal lengths are almost random and the support is usually high. Therefore, RANSAC often returns some result with high support which is however far from the ground truth value.

This can be seen in Figure 2 (a) and (b) which shows boxplots of focal lengths obtained by 1000 runs of the DEGENSAC [10], where we extracted focal lengths using the Bougnoux equations [2] inside the DEGENSAC loop. In each run of the DEGENSAC the real focal length with the highest support was returned. Figure 2 (a) shows results for the real scene where camera optical axes were almost intersecting. This is the critical configuration for a pair of cameras with constant or varying focal lengths [13]. Because in this case the configuration was not perfectly critical, i.e. principal points did not perfectly matched, the DEGENSAC always returned some focal lengths and epipolar geometry with a good support. However, the focal lengths were wrong.
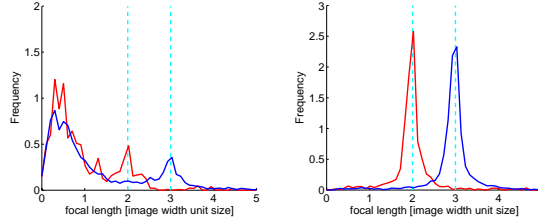
Fig. 3. The kernel voting on the scene with a dominant plane with only 10% off the plane points. Standard kernel voting (left), proposed algorithm with dominant plane detection (right). Left (right) focal length is blue (red), ground truth is cyan.

Unfortunately it is not possible to determine whether the estimated focal length is correct from one result of the DEGENSAC. This is also not completely clear by comparing results of multiple runs of the DEGENSAC, as it can be seen from the Figure 2. Here the variations of the focal lengths estimated from 1000 runs of the DEGENSAC are very similar for scenes close to critical configuration Figure 2 (a) and for non-critical configuration Figure 2 (b).

Again this is not a problem for kernel voting as it can be seen from the Figure 2 (c) and (d) where the results for the same sequences and the weighted kernel voting on the data collected during a single execution of the DEGENSAC are shown. More peaks in Figure 2 (c) are results of the model instability near the critical motion. The plot looks crisp too, since many votes were dropped due to the detected epipolar geometry degeneracy or because extracted focal lengths were complex. On the other hand the result for the general scene (Figure 2 (d)) is nicely smooth with only a single peak. Therefore it is meaningful to consider the estimated focal length reliable only if the highest peak is sufficiently higher and more consistent than the remaining data.

2.4   Degenerate scenes

Degenerate scenes produce results with high support but usually with incorrect focal lengths. For example, in scenes with dominant planar structure it often happens that all points from the sample are on the plane. Thus the epipolar geometry is degenerate but all points on the plane match this epipolar geometry perfectly [8] and the standard as well as weighted kernel voting and RANSAC fail to estimate the correct focal length.

Therefore we combine our kernel voting method with degeneracy tests. Note that in scenes containing dominant planes we can't use the number of inliers as weights since degenerated focal lengths have high support on points from the plane. Therefore we use weights estimated only from the points off the plane.

Figure 3 left shows result for the standard kernel voting without a test on planar scene degeneracy. The right plot in Figure 3 shows the result of our kernel voting where the planarity is taken into account.

2.5   Multiple focal length candidates

It often happens that we have several candidates for camera focal length obtained by running RANSAC several times for one image pair or by running RANSAC for several image pairs with common cameras. Mostly, these focal lengths are different as it can be seen in Figure 2 (b) and it is difficult to select the correct one. Strategies like selecting the focal length with the largest number of inliers, selecting the median or mean focal length, or running standard kernel voting on results from RANSAC [28, 29] do not always produce satisfactory results.

To solve this problem we collect reliable candidate focal lengths with their weights for each camera pair, respectively each run of the RANSAC. Then we use weighted kernel voting to select the best focal length from these candidates.

3   The robust method for focal length estimation

Unlike previous works [21, 22, 28, 29], we execute single RANSAC algorithm and then postprocess obtained inliers. The idea is the following: If we executed a RANSAC based algorithm on all $\binom{N}{7}$ 7-tuples chosen out of N tentative matches, then we would obtain all maximal inlier sets. In general, each of the maximal inlier sets can be obtained from many different sampled 7-tuples. Each 7-tuple generating a maximal inlier set may, however, result in a different epipolar geometry and different focal length. For reliable estimates, the distribution of these focal lengths should have a clear dominant peak.

To speed-up the process, we run RANSAC only once to obtain an inlier set. Then, we study the distribution of the focal lengths which result from 7-tuples sampled from and generating this inlier set (or its similarly sized subset). In this way one can determine if the estimated focal lengths are reliable and also select the "best" focal length as the value corresponding to the highest peak in the distribution.

Our weighted kernel voting algorithm is a cascade consisting of four phases. Block diagram of the algorithm is presented in Figure 4 and the pseudo code of this algorithm can be found in [4].

3.1   Phase 1 - Matches selection

The main goal of the first phase is to achieve computational efficiency by quickly rejecting easy mismatches and thus not wasting time and effort in the next phase.

We run DEGENSAC [10] which returns a set of matches in which the proportion of mismatches is greatly reduced and most of correct matches are preserved. In other words the decision process of the first phase generates a negligible number of falsely rejected good matches (false negatives) but a non-negligible number of correctly rejected false matches (true negatives).

It is important to use DEGENSAC [10] or a similar algorithm which is capable of detecting panoramas and pure planar scene configurations and obtaining inliers which are not affected by presence of dominant planes. Panoramas and planar scenes are rejected.
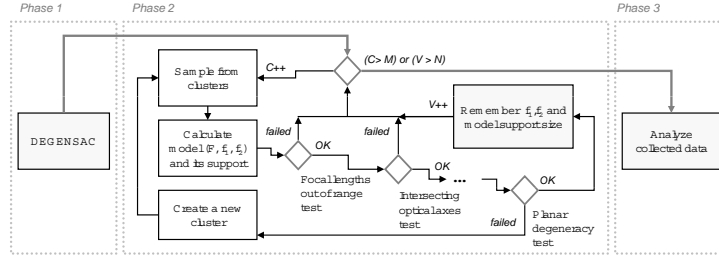
Fig. 4. Block diagram of the weighted kernel voting algorithm. See text for description.

3.2 Phase 2 - Votes collecting

The second phase of our algorithm is used to collect "focal length votes". Each vote, i.e. each estimated focal length, is weighted by the support of the epipolar geometry corresponding to the focal length. The higher the support of the model, the higher the weight of the estimated focal length. It is important to filter degenerate models since they usually have good supports but incorrect focal lengths.

The algorithm tries to collect N votes (non-degenerate epipolar geometries with their focal lengths) in less than M ($M > N$) trials. Since input data are already inliers, we cannot use ordinary statistics developed for the RANSAC to estimate M because it would be too small. In our experiments we set $N = 50$ and $M = 100$. We rejected a camera pair if it was not possible to collect 50 non-degenerated votes in 100 trials, or in other words if $V/C < 0.5$, where $V \leq N$ is the number of collected non-degenerated votes in $C \leq M$ trials.

Note that this phase is as computationally expensive as at most M additional RANSAC cycles. For $M = 100$ this is amounts to 10's of milliseconds.

Clusters To avoid computation of degenerate epipolar geometries we divide all matches into several clusters. We distinguish planar clusters and the remaining data (the "Zero cluster"). Each planar cluster represents a set of points laying on a non-negligible plane. Clusters smaller than five points and all remaining matches are stored in the Zero cluster. The algorithm starts by putting all matches to the Zero cluster. Then, the clusters corresponding to planes in the scene are automatically created during the algorithm runtime as will be explained next.

Model calculation Computation of epipolar geometries is done using a small (often minimal) number of point correspondences required to estimate the model. Correspondences are drawn from different clusters to avoid selection of points laying in one plane. Since the Zero cluster contains points in general position, we also allow sampling all correspondences from this cluster.

Various solvers can be used to calculate fundamental matrices and focal lengths. It is better to use information about cameras whenever available since this yields more stable parameter estimation [8, 25, 15, 27, 3, 30].

Degeneracy tests Several degeneracy tests are executed to avoid voting of degenerate samples/models. First, models with focal lengths that are outside a reasonable interval are ignored. These may be products of too noisy data or mismatches. Similarly, votes resulting from cameras with intersecting optical axes, i.e. $(0, 0, 1)\mathtt{F}(0, 0, 1)^T = 0$, are rejected. For the plane degeneracy test we are using test developed in DEGENSAC [10].

If at least 6 points are on the plane or sample was drawn from the zero cluster and 5 points are on the plane then we create a new cluster. First, a plane is calculated from 5 or 6 points. Then, points on the plane are removed from clusters and a new cluster using on plane points is created. Finally, clusters with less than 5 points are relabeled to Zero cluster.

Although the plane degenerate samples are ignored, each such sample creates a new cluster with points laying on the plane. Since samples are drawn from different clusters then the probability of sampling a new plane degenerate sample is gradually decreasing as more and more dominant planes are removed from the Zero cluster.

3.3   Phase 3 - Votes analysis

After the votes are collected, the algorithm determines whether the estimated focal lengths are consistent and reliable by analyzing collected data.

First, if a camera pair was close to some critical configuration [13], then almost all votes were rejected by degeneracy tests (see above) and hence the number of trials C required for obtaining V votes was high. If the fraction of non-degenerated votes and the number of cycles is small, i.e. $V/C < 0.5$, then we reject such a camera pair.

Next, the weighted kernel method with weights estimated from the support of each focal length is used to estimate the kernel density approximation of the probability density function of collected focal lengths. If the distribution produces a dominant peak, i.e. the highest peak is at least 20% above the remaining data, we extract the focal length as the argument of its maximum. Otherwise we ignore the camera pair.

We consider the estimated focal length as reliable only if both these criteria are met.

3.4   Phase 4 - Multi-view voting

For the multi-view voting process we create an accumulator for each camera where the results from camera pairwise estimations are collected. Each accumulator is a vector covering the range from $20mm$ to $150mm$ with $1mm$ tessellation. Given the result of a pairwise estimating we first analyze if the result is reliable.

(a)   (b)   (c)   (d)
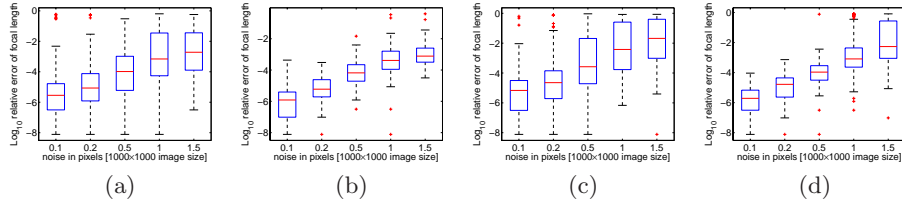
Fig. 5. Deviation of estimated focal length of the first camera using proposed voting approach in general scene (a,b) and scene with a dominant plane (c,d). Individual camera pairs are displayed in (a,c), grouped votes from 5-pairs in (b,d).

We do this using the two criteria described in Section 3.3. If both these conditions are satisfied, then we add votes with their weights to the camera accumulator otherwise we reject the camera pair. After all data are collected we run the final kernel voting for accumulator data.

4   Experiments

4.1   Synthetic data set

We study the performance of the method on synthetically generated ground-truth general 3D scenes as well as on the scenes with dominant planes. Scenes were generated as random points on a plane or in a 3D cube depending on the testing configuration or using a combination of both to get a planar scene with minor 3D structure. Each 3D point was projected by several cameras, where each camera orientation, position and focal length was selected randomly. Gaussian noise with a standard deviation σ was added to each image point.

Noise free data set Behavior of the standard kernel voting on noise free general 3D scenes was studied for the 6-pt algorithm with equal focal lengths already in [16]. The results are similar for the 7-pt algorithm followed by a focal length extraction. There is no reason for this algorithm to fail.

The behavior on planar scenes and for cameras near a critical configuration is different. Omitting degeneracy test causes that the standard kernel voting completely fails. This is shown in Figure 3. Since our algorithm samples points from different clusters, i.e. points from different planes, it rarely tests a 7-tuple of points laying on the plane. If it happens, i.e. when a planar sample is drawn from the Zero cluster, degeneracy test detects it and a new cluster is created. Adding outliers to the data does not affect the result since outlying votes are weak due to small support and hence weight. This is shown in Figure 1.

Data affected by noise It was demonstrated in [16] that the kernel voting is able to pick values close to the ground true value even for data affected by noise. In our experiments we fixed the focal length of the first camera to $35mm$

and generated 1000 random scene setups as described above. For each setup we executed 100 cycles of voting. We did the same for each selected noise level. Figure 5 (a) summarizes the results and shows that focal lengths estimated using our kernel voting method are accurate.

Figure 5 (b) shows the results, where we generated six cameras in each scene setup. Then votes from all five camera pairs between the first and the ith camera were used to vote for the focal length of the first camera. Obtained estimates are even more accurate.

Next, we repeated the above tests for a scene where 80% of all points are on a plane. Results are summarized in Figure 5 (c,d). It can be seen that results for planar scenes are slightly less accurate than the results for general scenes (a,b). This may be caused by the fact that it is harder to fit a good model to such data due to smaller amount of good matches.

Adding outliers to the tests did not affect the result too much. This is because the RANSAC and weighting with model support inside the voting algorithm can cope with outliers after if a sufficient number of trials are executed. We omit these results here, since they look similar to the ones obtained for outliers free scenes.

4.2   Real data set

To evaluate our voting approach on a real data we downloaded 2500 images from the Flickr [6] database using *"Di trevi"* keywords. In every such image we extracted SURF [1] feature points and descriptors. Tentative correspondences between each two calibrated images were obtained as points where the best descriptor dominates by 20% over the second best descriptor [20]. Then we used the DEGENSAC [10] algorithm to estimate inlaying correspondences and 100 cycles of our voting algorithm to analyze the quality of the estimated geometry of the pair. Each reliable geometry (see Section 3) was then added to the camera accumulators. We created accumulators with the range from $20mm$ to $200mm$ with one millimeter tessellation.

From the 2500 images we found only 240 images where focal length could be extracted from the *jpeg-exif* headers. About 130 of them were either showing something different or could not be matched, i.e. the number of correct tentative correspondences was less then 30. Algorithm marked about 30 images as unreliable. The *jpeg-exif* focal lengths of the remaining 80 images were compared with results of our algorithm, see Figure 6 (top left).

It can be seen that the estimated focal lengths (red dots) are in most cases very close to the focal lengths extracted from the *jpeg-exif* headers (green crosses). Examples of votes coming from individual camera pairs are shown in Figure 6 (bottom). The top right plot in the figure shows the final accumulator after applying KDE. Result of our method is displayed in red, standard kernel voting on inliers in black and vertical lines represent *jpeg-exif* focal length (cyan), mean (green), median (red) and result with max support (blue) from several DEGENSAC runs.

Figure 2 shows the results of standard DEGENSAC (a,b) and our method (c,d) for real images taken with known camera in close to a critical configuration
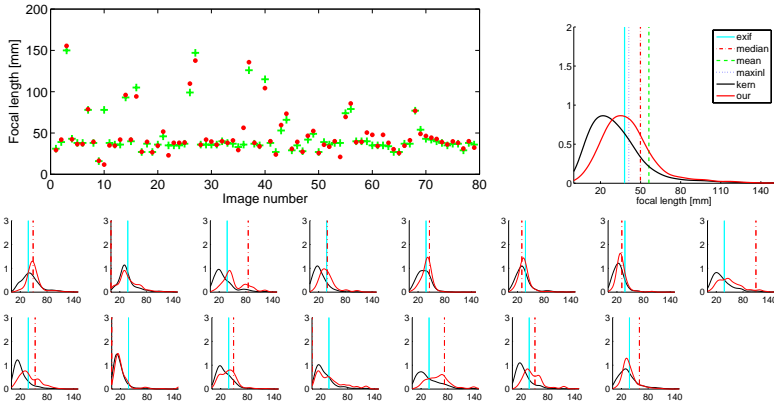
Fig. 6. Estimated focal lengths (red dots) with ground truth values (green crosses) extracted from the *jpeg-exif* (top-left). Distribution of votes for selected camera (bottom) and result for final multi-view voting (top-right). Our method is displayed in red, standard kernel voting on inliers in black. Vertical lines represent *jpeg-exif* focal(cyan), mean(green), median(red) and result with max support(blue) from DEGENSAC.

(a,c) and general configuration (b,d). As it can be seen DEGENSAC returned inaccurate estimates many times even for general scene. However, our method was able to detect scene closet to critical configuration and to estimate focal lengths close to the ground truth value for general configurations.

5   Conclusion

We have proposed a new, fast, multi-view method for robust focal length estimation. This method can be used with any focal length extraction algorithm (e.g. 6-pt, 7-pt, etc.), combines the RANSAC paradigm with the weighted kernel voting using weights derived from the number of inliers, contains detection of planar scenes and some critical configurations, thanks to which it can detect "bad" pairs and handle dominant planes.

This method produces reliable focal length estimates which are better then estimates obtained using plain RANSAC or kernel voting, and which are in most real situation very close to the ground truth values. This method is useful in SFM especially from unordered data sets downloaded from the Internet.

References

1. H. Bay, and A. Ess, and T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). *CVIU, 110:346–359, 2008*
2. S. Bougnoux. From projective to Euclidean space under and practical situation, a criticism of self-calibration. *ICCV 1998*

3. M. Bujnak, Z. Kukelova, and T. Pajdla. 3D reconstruction from image collections with a single known focal length. In *ICCV* 2009.
4. M. Bujnak, Z. Kukelova, and T. Pajdla. Robust focal length estimation by voting in multi-view scene reconstruction. *Research Report CTU-CMP-2009-09*, 2009
5. M. A. Fischler and R. C. Bolles. Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM*, 24(6):381–395, 1981.
6. *Flickr.* http://www.flickr.com/
7. R. Hartley Estimation of relative camera positions for uncalibrated cameras *ECCV 1992*, Italy, pp. 579-587, May 1992.
8. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2003.
9. O. Chum, J. Matas., J. Kittler.   Locally Optimized RANSAC. DAGM 2003.
10. O. Chum, T. Werner, and J. Matas. Two-View Geometry Estimation Unaffected by a Dominant Plane. *CVPR 2005* pp.772-779.
11. O. Chum,and J. Matas, Matching with PROSAC - Progressive Sample Consensus . *CVPR 2005.*
12. M. C. Jones, J. S. Marron, and S. J. Sheather. A brief survey of bandwidth selection for density estimation. *J. Amer. Stat. Assoc.*, 91(433):401-407, March 1996.
13. F. Kahl and B. Triggs.   Critical Motions in Euclidean Structure from Motion. *CVPR 1999*, pp. 2366.
14. K. Kanatani and C. Matsunaga. Closed-form expression for focal lengths from the fundamental matrix. *ACCV 2000*, Taipei, Taiwan, vol. 1, pp. 128-133
15. Z. Kukelova, M. Bujnak, T. Pajdla, Polynomial eigenvalue solutions to the 5-pt and 6-pt relative pose problems. *BMVC 2008.*
16. H. Li. A simple solution to the six-point two-view focal-length problem. *ECCV 2006*, pp. 200–213.
17. X. Li, C. Wu, C. Zach, S. Lazebnik, and J. Frahm. Modeling and recognition of landmark image collections using iconic scene graphs. In *ECCV* 2008.
18. D. Martinec and T. Pajdla. Robust Rotation and Translation Estimation in multi-view Reconstruction In *CVPR* 2007.
19. *Microsoft PhotoSynth.* http://www.photosynth.net
20. M. Muja, and D. Lowe. Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration. *Preprint, University of British Columbia, 2008.*
21. D. Nister. An efficient solution to the five-point relative pose. *IEEE PAMI*, 26(6):756–770, 2004.
22. D. Nister and C. Engels. Visually Estimated Motion of Vehicle-Mounted Cameras with Global Uncertainty. *SPIE* Defense and Security Symposium, Unmanned Systems Technology VIII, April 2006.
23. N. Snavely, S.M. Seitz, R. S. Szeliski. Photo Tourism: Exploring image collections in 3D. In *SIGGRAPH 2006*, pp. 835-846.
24. N. Snavely, S. Seitz, and R. Szeliski. Skeletal graphs for efficient structure from motion. In *CVPR* 2008.
25. H. Stewenius, D. Nister, F. Kahl, and F. Schaffalitzky. A minimal solution for relative pose with unknown focal length. *CVPR 2005*, pp. 789–794.
26. H. Stewenius, C. Engels, and D. Nister. Recent developments on direct relative orientation. *ISPRS J. of Photogrammetry and Remote Sensing*, 60:284–294, 2006.
27. P. Sturm. On Focal Length Calibration from Two Views. *CVPR 2001.*
28. A. Torii, M. Havlena, T. Pajdla, and B. Leibe. Measuring Camera Translation by the Dominant Apical Angle. *CVPR 2008*, Anchorage, Alaska, USA, 2008.
29. A. Torii and T. Pajdla. Omnidirectional camera motion estimation. *VISAPP 2008*.
30. M. Urbanek, R. Horaud, and P. Sturm. Combining Off- and On-line Calibration of a Digital Camera. *Third Int. Conf. on 3-D Digital Imaging and Modeling, 2001.*