

ACCELERATED RANSAC FOR 2D HOMOGRAPHY ESTIMATION BASED ON GLOBAL BRIGHTNESS CONSISTENCY

Gaku Nakano

Central Research Labs, NEC Corporation, Japan
g-nakano@cq.jp.nec.com

ABSTRACT

This paper proposes a novel sampling method for accelerating RANSAC family on 2D homography estimation. From the initial set of matched points, the proposed method generates a promising reduced subset having higher inlier ratio than the initial set by utilizing pixel values of the matches. Regarding pairs of the pixel value as two dimensional scattered points, we estimate the global brightness consistency of the pixel values. Then, points that violate the global brightness consistency are removed from the initial point set. Incorporating the proposed method with RANSAC and USAC, we demonstrate that the number of iterations and computational time are both significantly reduced by orders of magnitude while maintaining accuracy of homography estimation.

Index Terms— RANSAC, robust estimation, outlier rejection, 2D homography

1. INTRODUCTION

Estimating a 2D homography parameter between two images is a classical but fundamental problem for computer vision applications such as object recognition [1, 2], panorama stitching [3, 4]. Recent trend in those applications is to handle thousands or millions of images, therefore, fast algorithms have been studied for each processing step: feature point detection and description, descriptor matching, and parameter fitting. For the first and second steps, many efficient feature point detectors and descriptors have been proposed [5]. However, irrespective of the method we use, mismatched points are inevitable due to changes of viewpoint, illuminations, and occlusions. Therefore, the final parameter estimation step must be robust against mismatched points, called outliers.

The random sample consensus (RANSAC) [6] is one of the standard approaches for estimating a parameter from a set of data points contaminated by outliers. As the name implies, RANSAC repeatedly performs generating a hypothesis of the parameter from randomly sampled points and verifying its correctness by counting the number of inliers, for which error against the hypothesis is less than a predefined threshold. The stopping criterion of RANSAC, the minimum number of iterations to obtain the best parameter, can be given by

$N = \log(1 - p) / \log(1 - w^s)$, where w is inlier ratio in the whole data, s is the sample size to generate a hypothesis, and p is the probability that at least one of the random points has no outliers, which is typically set to $p = 0.99$. As the equation implies, RANSAC becomes exponentially slower for lower inlier ratio and larger sampling size.

Since it is theoretically impossible to reduce N by sampling less than four ($s < 4$) points for 2D homography estimation in general cases, the motivation of this paper is to generate a small subset from the initial data points, which is expected to have higher inlier ratio than the initial data. As shown in Fig.1, generating such subset is particularly effective for severely contaminated data.

For generating promising subsets, some methods have been proposed in the literature [7, 8, 9], which utilize a prior information on image content. The progressive sample consensus (PROSAC) [7] divides the initial point set into several subsets based on sorted matching scores of feature points. The hypothesis-testing step starts by the top-ranked subset, then, shifts to lower quality subsets until convergence. Although this strategy enables to reduce the number of iterations, solutions returned by PROSAC are often trapped into local minima, which fit a few high quality points concentrated in a small area on the input image space. Assuming that a large image area contains more inliers than smaller areas, GroupSAC [8] applies image segmentation or optical flow on the initial point set to find subsets. While GroupSAC experimentally showed an advantage of the grouping strategy, a trade-off problem has not been discussed well yet between grouping accuracy, grouping efficiency, and sampling improvement. Universal RANSAC (USAC) [9], the state-of-the-art of RANSAC family, efficiently delivers stable solutions by integrating various techniques proposed over the years; the PROSAC sampling, sample validation before parameter estimation, SPRT test [10] for model validation, DEGENSAC [11] for checking parameter degeneracy, LO-RANSAC [12] for faster convergence. Those works showed that utilization of a priori information on image content has a large potential to accelerates RANSAC.

This paper proposes a novel sampling method for generating a promising subset of a planar object matching based on correlation analysis of corresponding pixel values. Assuming

Sample size s	Inlier ratio w					
	90%	...	40%	30%	20%	10%
3	4	...	70	168	573	4603
			Theoretically impossible			
4	4	...	178	566	2876	46049
			Proposed approach			

Fig. 1. The number of iterations N and the concept of the proposed method. Since a three-point method does not exist for 2D homography estimation, our approach aims to generate a reduced subset, which has higher inlier ratio than the original point set.

that a Lambertian plane is lit by a single dominant light, we find out that pixel values of inlier points are globally correlated. Then, we generate a subset by removing points that do not satisfy the global brightness consistency. We call this new sampling strategy based on the global brightness consistency as *GBC sampling*. The process to estimate the global brightness consistency is equivalent to a two dimensional correlation analysis, therefore, an additional computational time is negligibly small in the overall runtime. Experiments on real data show that the proposed method significantly accelerates RANSAC family by orders of magnitude, in spite of a quite simple algorithm.

2. GLOBAL BRIGHTNESS CONSISTENCY

In this paper, we assume that a plane to be matched has Lambertian surface and is lit by a single dominant light. This assumption leads to the following predictions about the relation between pixel values of inliers.

Let I_i and J_i be pixel values of i -th inlier feature points on images I and J .

- If the light and the object did not move, I_i and J_i are nearly equal. ($J_i \approx I_i$)
- If the angle or the strength of the light changed, I_i and J_i are in linear or log-linear. ($J_i = \alpha I_i + \beta$ or $J_i = \alpha I_i^\beta$)
- If camera response functions of the images are highly non-linear, the relation of I_i and J_i can be approximated by a polynomial equation. ($J_i = \sum_{k=0}^n c_k I_i^k$)

Due to the Lambertian assumption, the above brightness variations are globally consistent, i.e., the coefficients α , β , c_k are common for all inliers and not dependent on each point.

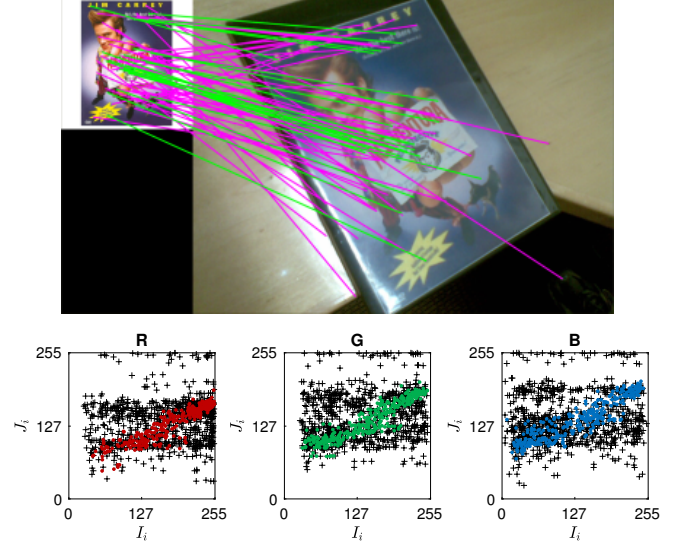


Fig. 2. (top) Line drawing of matched feature points: inliers in green and outliers in magenta. Only 20% of matches are shown for ease of viewing. (bottom) Scatter plot of RGB values of all matched points in the top images. A pixel value I_i of the i -th point in the left image corresponds to a pixel value J_i of the i -th point in the right image. Red, green, and blue dots show inlier colors on each channel and small black pluses show outlier colors.

In other words, points that do not satisfy the global brightness consistency are potentially outliers. Moreover, we can expect that the dominant light assumption holds in various kinds of scene, e.g., outdoor sequences under the sunlight and indoor sequences with gradual shading.

Figure 2 visually shows this concept. Inlier pixel values, RGB dots in the bottom figure, are distributed in a certain area, which seems to be on a line or curve with some distance. On the other hand, outlier pixel values, black pluses, seem to be uniformly distributed around the inlier area. This result indicates that it is possible to distinguish inliers and outliers if the promising area can be determined by analyzing the distribution of pixel values.

3. PROPOSED METHOD

This section describes a method for determining a promising area of scattered color points (I_i, J_i) , where most inliers lie and follow the global brightness consistency. We will explain the proposed method, the GBC sampling, step-by-step with referring to Fig. 3, which illustrates the algorithm.

Step 1: Pick pixel values (I_i, J_i) of all initial matches on RGB channels.

Step 2: Regarding pixel values (I_i, J_i) as scattered two dimensional points, calculate a covariance ellipse on each

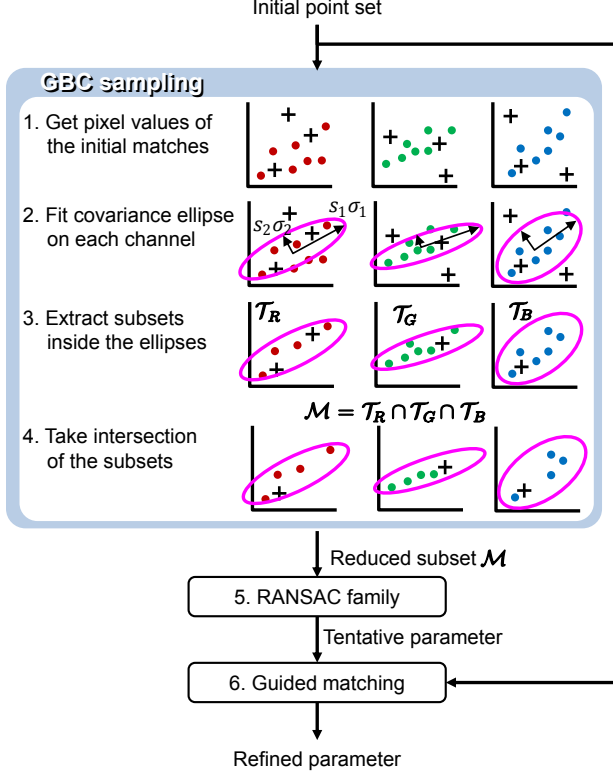


Fig. 3. Algorithm overview of RANSAC integrated with the proposed method, the GBC sampling. Note that inliers (color dots) and outliers (black pluses) are not separated before the GBC sampling but just shown for ease of understanding.

channel by performing principal component analysis. The length of major and minor axes can be given $s_1\sigma_1$ and $s_2\sigma_2$, where σ_1 and σ_2 represent the square root of the eigenvalues of the covariance matrix, and s_1 and s_2 are scale factors.

Step 3: Extract subsets $\mathcal{T}_R, \mathcal{T}_G, \mathcal{T}_B$, whose points are inside the covariance ellipse on each channel.

Step 4: Calculate a reduced subset \mathcal{M} as an intersection of the three subsets, $\mathcal{M} = \mathcal{T}_R \cap \mathcal{T}_G \cap \mathcal{T}_B$. This step is omitted for grayscale images.

Step 5: Execute RANSAC or its variants on the reduced subset \mathcal{M} . We obtain inliers of the subset and a tentative parameter fitting on the inliers.

Step 6: Apply guided matching [13] on the initial point set. Finally we obtain more inliers and a refined parameter.

The key of the GBC sampling is the **Step 2** for determining the size of the subset on each channel. As described in Section 2, we expect that inlier and outlier colors follow a normal and uniform distribution, respectively. According to the central limit theorem, the union of both color points

are assumed to be follow a heavy-tailed normal distribution. Thus, regarding the corresponding color points (I_i, J_i) as scattered 2D points, we fit a confidence ellipse on the color points by two dimensional correlation analysis in order to extract a promising area. Under a normal distribution, the scale factors s_1 and s_2 should be determined so that the ellipse covers $\geq 90\%$ and $\leq 50\%$ points for the major and minor axes, respectively. For instance, $(s_1, s_2) = (3.0, 0.5)$ has 99% and 38% confidence intervals for each axes.

To fit the covariance ellipses at the **Step 2**, we implicitly assume an independence between the color channels. To be exact, RGB values at an image point are actually not unique on spectral space. However, the dependency or the overlapped wavelength is very slight in general. Thus, we ignore it in this paper and will experimentally validate this assumption in Section 4.

The **Step 6**, guided matching on the initial point set, is carried out to obtain more inliers and a more accurate parameter since the subset \mathcal{M} does not contain all inliers.

4. EXPERIMENTS

4.1. Dataset

We have tested the performance of the proposed method on publicly available image datasets [9, 14, 15]. We chose image pairs taken under different illumination conditions and performed feature point matching by using SIFT implemented on VLFeat [16] if the point information is not provided. To generate the groundtruth data, we performed RANSAC with fixed 10^7 iterations and 2.0-pixel threshold on the feature point pairs. Due to limitations of space, we selected four sequences from the datasets, which are adequate to demonstrate effectiveness of the proposed method in realistic situations. We integrated the GBC sampling with RANSAC and USAC, then, evaluated performance improvements. In the experiment, optional parameters of RANSAC and USAC were configured to the default values as provided by the USAC-1.0 source code. All evaluations were executed on a Core i7-6700 PC and MATLAB 2016b.


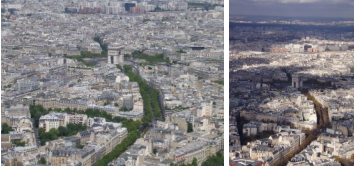
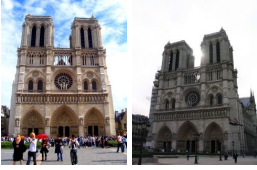
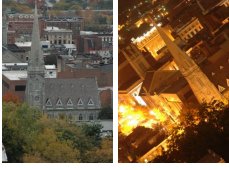
4.2. Results

Table 1 summaries experimental results that are mean values of 100 trials for each sequence.

4.2.1. Inlier ratio

The Dataset column of Table 1 clearly shows that inlier ratio of subsets generated by the GBC sampling is significantly increased by nearly or more than 10% in all sequences. From this, we see that the global brightness consistency holds in real situations.

Table 1. Results on real images. The number of points (m), inlier ratio (w [%]), the number of iterations ($iter$), computational time in milliseconds ($time$), acceleration factor of runtime against the original RANSAC ($speed-up$), and relative error of estimated parameter against the groundtruth in percentage (err).

Dataset				Results					
Images		Point set	m	w	Method	$iter$	$time$	$speed-up$	err
		Initial	717	43.0	RANSAC	282	4.6	1	2.07
		Reduced	141	71.6	USAC	24	5.0	0.92	1.23
				+28.6	RANSAC+GBC	44	0.6	8.57	1.26
					USAC+GBC	3	1.4	3.20	1.24
		Initial	191	69.6	RANSAC	44	0.5	1	0.97
		Reduced	36	86.1	USAC	5	1.5	0.36	0.07
				+16.5	RANSAC+GBC	21	0.3	2.30	0.10
					USAC+GBC	5	0.6	0.86	0.08
		Initial	286	14.3	RANSAC	25208	262	1	3.53
		Reduced	103	24.3	USAC	4113	10	54.6	5.42
				+10.0	RANSAC+GBC	2629	22	12.4	4.92
					USAC+GBC	37	1	317.6	1.31
		Initial	979	12.5	RANSAC	44675	883	1	0.55
		Reduced	206	29.6	USAC	267047	243	3.8	0.08
				+17.1	RANSAC+GBC	1161	13	67.9	0.40
					USAC+GBC	5	1	883	0.08

4.2.2. Efficiency

The results with and without the GBC sampling are listed in the Results column of Table 1. Although the runtime includes an additional time spent by the proposed method, it was negligibly small.

The increase of inlier ratio, which is mentioned in Section 4.2.1, leads to a significant reduction of the number of iterations by orders of magnitude. Consequently, the computational time is accelerated by ranging from $\times 2.31$ to $\times 883$.

An interesting result is that RANSAC+GBC is faster than the naive USAC and USAC+GBC in the two sequences. Implementing USAC requires many techniques; however, this result suggests that a simple RANSAC with the GBC sampling can be efficient enough depending on a scene.

4.2.3. Parameter accuracy

The estimation error of homography transformation was slightly improved or almost comparable. What we would like to emphasize is that the proposed method achieved the comparable accuracy with significantly saving computational time, as described in Section 4.2.2.

5. CONCLUSIONS

In this paper, we have presented a simple and powerful method, called *GBC sampling*, for accelerating RANSAC family on 2D homography estimation by using pixel values as a priori information. We made an assumption on a scene illumination and a planar object so that pixel values of inlier points follow the global brightness consistency. Under the assumption, we proposed a method for generating a promising reduced subset of which inlier ratio is higher than the initial point set. Real image experiments showed that the proposed method significantly reduces the number of iterations and computational time by orders of magnitude while keeping parameter estimation accuracy. The experiments in this paper are small enough to verify the concept; however, an objective evaluation on a large dataset is required to show more effectiveness and applicability. Furthermore, expanding the proposed approach to other parameter estimation problems, e.g., fundamental matrix estimation, is also an interesting topic as a future work.

6. REFERENCES

- [1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, IEEE, 2009, pp. 248–255.
- [2] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman, “The pascal visual object classes (voc) challenge,” *International journal of computer vision*, vol. 88, no. 2, pp. 303–338, 2010.
- [3] Matthew Brown and David G Lowe, “Automatic panoramic image stitching using invariant features,” *International journal of computer vision*, vol. 74, no. 1, pp. 59–73, 2007.
- [4] Noah Snavely, Steven M Seitz, and Richard Szeliski, “Modeling the world from internet photo collections,” *International Journal of Computer Vision*, vol. 80, no. 2, pp. 189–210, 2008.
- [5] Ondrej Miksik and Krystian Mikolajczyk, “Evaluation of local detectors and descriptors for fast feature matching,” in *Pattern Recognition (ICPR), 2012 21st International Conference on*, IEEE, 2012, pp. 2681–2684.
- [6] Martin A Fischler and Robert C Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
- [7] Ondřej Chum and Jiří Matas, “Matching with prosac-progressive sample consensus,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, IEEE, 2005, vol. 1, pp. 220–226.
- [8] Kai Ni, Hailin Jin, and Frank Dellaert, “Groupsac: Efficient consensus in the presence of groupings,” in *Computer Vision, 2009 IEEE 12th International Conference on*, IEEE, 2009, pp. 2193–2200.
- [9] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jose Matas, and Jens Frahm, “Usac: A universal framework for random sample consensus,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, no. 8, pp. 2022–2038, 2013.
- [10] Jiří Matas and Ondřej Chum, “Randomized ransac with sequential probability ratio test,” in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, IEEE, 2005, vol. 2, pp. 1727–1732.
- [11] Ondřej Chum, Tomáš Werner, and Jiří Matas, “Two-view geometry estimation unaffected by a dominant plane,” in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, IEEE, 2005, vol. 1, pp. 772–779.
- [12] Ondřej Chum, Jiří Matas, and Josef Kittler, “Locally optimized ransac,” in *Pattern recognition*, pp. 236–243. Springer, 2003.
- [13] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, ISBN: 0521540518, second edition, 2004.
- [14] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, “Lost in quantization: Improving particular object retrieval in large scale image databases,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2008.
- [15] Ling-Yu Duan, Jie Lin, Jie Chen, Tiejun Huang, and Wen Gao, “Compact descriptors for visual search,” *IEEE MultiMedia*, vol. 21, no. 3, pp. 30–40, 2014.
- [16] A. Vedaldi and B. Fulkerson, “VLFeat: An open and portable library of computer vision algorithms,” 2008.