Evaluating Blur Detection Methods for Factory SLAM Applications

Taggart Bonham
Corporate R&D - Magna International
Technical White Paper
August 2017

ABSTRACT: Motion blur, one of the biggest problems for feature-based SLAM (Simultaneous Localization and Mapping) systems, causes inaccuracies and location losses during tracking. This paper evaluates the runtime and accuracy of top blur detection methods during use with SLAM in factory conditions. It was found that Modified Sum of Laplacian method performs the best overall. Based on its low computational requirement and high precision, it is the best solution for real-time, onboard SLAM applications.

INTRODUCTION

Indoor robotic vision systems operating in warehouses must navigate through dynamic, consistently-lit, feature-dense, enclosed areas (Gadd and Newman 2015, Andreasson 2008). In this use case, feature-based SLAM is more efficient compared to direct SLAM and more accurate in tracking dynamic scenes (Mur et al. 2015, Davison et al. 2007). The biggest limitation is in tracking on uniform-surfaces and dealing with motion blur (Klein 2008, Valencia et al. 2008). In this paper, the top blur detection methods are evaluated for use in real-time SLAM in factories.

METHOD

Based on the many literature surveys, a list on the top blur detection methods and their properties (including processing time and relative quality) was compiled (Rusiñol 2014, Mir 2014). The relative quality is a number given as a comparison to An et al. 2008's 3D Laplacian method, and the processing times come from Pertuz et al. 2013 [Table.1]. This SLAM system would be running a 30fps camera, with a 2.2GHz Intel i7 processor with 4GB of RAM, thus speed was a greater concern than accuracy. As such, any method slower than 20ms or with accuracy less than 70% was removed. After these criteria were met, four methods remained[1]—

1. *Modified Sum of Laplacian:*

$$\phi(x,y) = \sum_{(i,j)\in\Omega(x,y)} \Delta_m I(i,j),$$

where $\Delta_m I$ is the modified Laplacian of I, computed as

$$\Delta_m I = |I * \mathcal{L}_X| + |I * \mathcal{L}_Y|.$$

The convolution masks used to compute the modified Laplacian are

$$\mathcal{L}_X = [-1\ \ 2\ \ -1],$$

and $\mathcal{L}_Y = \mathcal{L}_X^T$.

[Eqn.1: LAP2 (Nayar 1994)]

2. *Diagonal Laplacian:*

$$\Delta_m I = |I * \mathcal{L}_X| + |I * \mathcal{L}_Y| + |I * \mathcal{L}_{X1}| + |I * \mathcal{L}_{X2}|,$$

$$\mathcal{L}_{X1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 & 0 & 1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathcal{L}_{X2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

[Eqn.2: LAP3 (Thelen et al. 2009)]

3. *Gray-Level Variance:*

$$\phi_{x,y} = \sum_{(i,j)\in\Omega(x,y)} (I(i,j) - \mu)^2,$$

where μ is the mean gray-level of pixels within $\Omega(x,y)$.

[Eqn.3: STA3 (Firestone et al. 1991)]

4. *Variance of Laplacian:*

$$\phi_{i,j} = \sum_{(i,j)\in\Omega(x,y)} (\Delta I(i,j) - \overline{\Delta I})^2,$$

where $\overline{\Delta I}$ is the mean value of the image Laplacian within $\Omega(x,y)$.

[Eqn.4: LAP4 (Pech-Pacheco et al. 2000)]

Since Pertuz et al. 2013 showed that the relative quality of the detection methods should remain proportional across datasets, we first tested these methods for speed. Since the initial data on these methods was generated on images with depths of less than .5m, I rewrote these methods and tested them on real-time on using the image data generated by the Xbox Kinect RGB-D sensor from a TurtleBot[2] and several applicable RGB-D SLAM image sets from the Technical University of Munich (Strum 2014). These following measures were tested on four datasets with varying amounts of motion blur, translation, and rotation filmed in both warehouse and office settings in order to better understand their speed and accuracy

[1] Equations from appendix of Pertuz et al. 2013

[2] A mobile robot development platform. https://www.clearpathrobotics.com/turtlebot-2-open-source-robot/

in more realistic situations [Table.2]. The average speeds were recorded in [Table.3].

Finally, the accuracy of these methods was evaluated. The focus measures for each image across the three databases were statistically compared for correlation against the most accurate method as given by Pertuz et al. 2013. The accuracy measures are recorded in [Table.3].

RESULTS

Method Name	Code	Time (ms)	Relative Quality
Modified Sum of Laplacian	LAP2	7.2	0.92
Diagonal Laplacian	LAP3	10	0.83
Gray-level Variance	STA3	18	0.74
Variance of Laplacian	LAP4	15	0.72

[Table.1: Preliminary Data from Literature Surveys]

Method Name	TurtleBot Motion (ms)	Desk 2 (ms)	Walking 360 (ms)	Pioneer 360 (ms)
LAP2	1.37	1.40	1.30	1.38
LAP3	2.71	2.66	2.72	2.74
STA3	1.18	1.22	1.28	1.23
LAP4	4.45	4.47	4.90	4.20

[Table.2: Processing Times by Method and Dataset]

Method Name	Correlation (%)	Avg. Runtime (ms)
LAP2	100	1.36
LAP3	99.5	2.71
STA3	45.7	1.23
LAP4	94.8	4.51

[Table.3: Accuracy by Method]

DISCUSSION

Based on data from the TurtleBot and TUM datasets, there were tradeoffs between speed and accuracy. STA3, averaging 1.23ms, barely outperformed the second fastest measure, LAP2 at a close 1.36ms [Table.2]. Noting that Pertuz et al. 2013 found LAP2 20% more precise than STA3, the accuracy of the results was then computed [Table.1]. In this paper, precision was determined by correlation to the most accurate method as given by Pertuz et al. 2013 in [Table.1]. The results showed that STA3 was only 45.7% correlated to the best measure [Table.3]. As these numbers have a questionably larger range than current literature in [Table.1], the method chosen to calculate accuracy may not have been the optimal choice. Therefore, we should accept the trends, rather than the exact numbers, and agree with Pertuz et al. 2013 that STA3 is not as precise at blur detection as it is not strongly correlated to the most accurate measure. Given its accuracy and relative speed, the Sum of Modified Laplacian method is the ideal focus measure for warehouse-based SLAM systems [Eqn.1].

Given that the camera would run at 30 fps, or 33.3ms/image, LAP2 is fast enough at 1.36ms to be able to run onboard in real-time, making it practical for non-theoretical use. Moving forward, this blur detection method could be used to determine when blur-correction should be applied in the SLAM process. Reducing motion-blur would enable situation-agnostic improvement, enabling the robot to move faster and more accurately.

In the future, it would be interesting to build a salience detector into these focus measures. The biggest problem with these methods is that a sharply-focused, uniform white wall would score similarly to a blurry image. By factoring in the presence of features (using common detection methods such as SIFT, SURF, ORB etc.), a more robust score would be generated. By considering the number of features, this measure could better handle the boundary cases of uniform surfaces. Additionally, future work should look into the effect of varied lighting to better understand how the accuracy of these measures might change throughout the day based on available sunlight.

REFERENCES

An, Y., Kang, G., Kim, I. J., Chung, H. S., & Park, J. (2008). Shape from Focus Through Laplacian Using 3D Window. In *Future Generation Communication and Networking, 2008. FGCN'08. Second International Conference on* (Vol. 2, pp. 46-50). IEEE. http://ieeexplore.ieee.org/abstract/document/4734170/

Andreasson, H. (2008). *Local Visual Feature-Based Localization and Mapping by Mobile Robots* (Doctoral Dissertation, Örebro Universitet). https://pdfs.semanticscholar.org/783c/118a0443e5c7fa69b67de403a9d756897f50.pdf

Davison, A. J., Reid, I. D., Molton, N. D., & Stasse, O. (2007). MonoSLAM: Real-Time Single Camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(6), 1052-1067.
http://ieeexplore.ieee.org/abstract/document/4160954/

Firestone, L., Cook, K., Culp, K., Talsania, N., & Preston, K. (1991). Comparison of Autofocus Methods for Automated Microscopy. *Cytometry*, *12*(3), 195-206. http://onlinelibrary.wiley.com/doi/10.1002/cyto.990120302/full

Gadd, M., & Newman, P. (2015). A Framework for Infrastructure-Free Warehouse Navigation. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on* (pp. 3271-3278).
IEEE. http://www.robots.ox.ac.uk/~mobile/Papers/2015ICRA_gadd.pdf

Groen, F. C., Young, I. T., & Ligthart, G. (1985). A Comparison of Different Focus Functions for use in Autofocus Algorithms. *Cytometry*, *6*(2), 81-91. http://onlinelibrary.wiley.com/doi/10.1002/cyto.990060202/full

Işık, Ş. (2014). A Comparative Evaluation of Well-Known Feature Detectors and Descriptors. International Journal of Applied Mathematics, Electronics and Computers, 3(1), 1-6. https://www.atscience.org/IJAMEC/article/view/135

Klein, G., & Murray, D. (2008). Improving the Agility of Keyframe-Based SLAM. *Computer Vision– ECCV 2008*, 802-815.
https://link.springer.com/chapter/10.1007%2F978-3-540-88688-4_59?LI=true

Mir, H., Xu, P., & Van Beek, P. (2014). An Extensive Empirical Evaluation of Focus Measures for Digital Photography. In *Digital Photography* (p. 90230I). https://www.spiedigitallibrary.org/conference-proceedings-of-spie/9023/90230I/An-extensive-empirical-evaluation-of-focus-measures-for-digital photography/10.1117/12.2042350.short?fileName=90230I

Mur-Artal, R., & Tardós, J. D. (2015). Probabilistic Semi-Dense Mapping from Highly Accurate Feature-Based Monocular SLAM. In *Robotics: Science and Systems*. http://www.roboticsproceedings.org/rss11/p41.pdf

Nayar, S. K., & Nakagawa, Y. (1994). Shape From Focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *16*(8), 824-831. http://ieeexplore.ieee.org/abstract/document/308479/

Pech-Pacheco, J. L., Cristóbal, G., Chamorro-Martinez, J., & Fernández-Valdivia, J. (2000). Diatom Autofocusing in Brightfield Microscopy: a Comparative Study. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on* (Vol. 3, pp. 314-317). IEEE. http://ieeexplore.ieee.org/abstract/document/903548/

Pertuz, S., Puig, D., & Garcia, M. A. (2013). Analysis of Focus Measure Operators for Shape-From- Focus. *Pattern Recognition*, *46*(5), 1415-1432. https://www.researchgate.net/publication/234073157_Analysis_of_focus_measure_operators_in_shape-from-focus

Rusiñol, M., Chazalon, J., & Ogier, J. M. (2014). Combining Focus Measure Operators to Predict OCR Accuracy in Mobile-Captured Document Images. In *Document Analysis Systems (DAS), 2014, 11th IAPR International Workshop on* (pp. 181-185). IEEE. http://ieeexplore.ieee.org/abstract/document/6830994/

Santos, A., Ortiz de Solórzano, C., Vaquero, J. J., Pena, J. M., Malpica, N., & Del Pozo, F. (1997). Evaluation of Autofocus Functions in Molecular Cytogenetic Analysis. *Journal of Microscopy*, *188*(3), 264-272. http://onlinelibrary.wiley.com/doi/10.1046/j.1365-2818.1997.2630819.x/full

Sturm, J., Engelhard, N., Endres, F., Burgard, W., & Cremers, D. (2012)ß. A Benchmark For the Evaluation of RGB-D SLAM systems. In *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on* (pp. 573-580). IEEE. https://vision.in.tum.de/data/datasets/rgbd-dataset/

Sun, Y., Duthaler, S., & Nelson, B. J. (2004). Autofocusing in Computer Microscopy: Selecting the Optimal Focus Algorithm. *Microscopy research and technique*, *65*(3), 139-149. http://onlinelibrary.wiley.com/doi/10.1002/jemt.20118/full

Thelen, A., Frey, S., Hirsch, S., & Hering, P. (2009). Improvements in Shape-From-Focus for Holographic Reconstructions with Regard to Focus Operators, Neighborhood-Size, and Height Value Interpolation. *IEEE Transactions on Image Processing*, *18*(1), 151-157. http://ieeexplore.ieee.org/abstract/document/4685626/

Valencia, R., & Andrade-Cetto, J. (2008). Introduction. In *Mapping, Planning and Exploration With Pose SLAM* (pp. 1-6). Springer, Cham. http://ieeexplore.ieee.org/abstract/document/4652565/