

Одеський національний університет імені І. І. Мечникова

Інститут математики, економіки і механіки

Кафедра обчислювальної математики

Дипломна робота

бакалавра

на тему: **Програмна стабілізація відео**

Program video stabilization

Виконав: студент денної форми навчання
напряму підготовки

6.040301 Прикладна математика

Максімов Майк Юрійович

Керівник: проф., к. т. н., Мороз В.В.

Рецензент: доц., к. ф-м. н., Таїрова М.С.

Рекомендовано до захисту:

Протокол засідання кафедри

№ ____ від «_____» _____ р.

Завідувач кафедри

Захищено на засіданні ЕК № _____

Протокол № ____ від «_____» ____ р.

Оцінка _____ / _____ / _____

Голова ЕК

Одеса — 2017 р.

ЗМІСТ

Глосарій		4
Вступ		5
1 Огляд методів оцінки руху		6
1.1 Прямі методи оцінки руху		6
1.2 Непрямі методи оцінки руху		8
1.3 Порівняння методів		11
2 Математичний опис проблеми та алгоритми пошуку ключових точок		13
2.1 Математичний опис проблеми		13
2.2 Оператор Харріса		14
2.2.1 Класифікація точок за матрицею M		16
2.2.2 Функція відгуку кута		17
2.2.3 Пригнічення не максимумів		18
2.2.4 Модифікація метода		20
2.3 Scale Invariant Feature Transform (SIFT)		20
2.3.1 Процес пошуку ключових точок		21
2.3.2 Обчислення орієнтації ключової точки		25
2.3.3 Побудова дескриптора ключової точки		25
2.4 Speeded Up Robust Features (SURF)		27
2.4.1 Відмінності від SIFT		27
3 Алгоритм стабілізації відео		30
3.1 Features from Accelerated Segment Test (FAST)		30
3.1.1 Machine Learning Approach		31
3.1.2 Пригнічення не максимумів		33
3.2 FREAK: Fast Retina Keypoint descriptor		33
3.2.1 Орієнтація		35
3.3 Алгоритм RANSAC		36
3.4 Фільтр Калмана		38

4	Аналіз отриманих результатів	40
	Висновки	52
	Список літератури	53

ГЛОСАРІЙ

Ключова точка - попри те, що точного або універсального визначення не існує (адже це залежить від задачі та типу застосування), найчастіше ключовою точкою зображення називається фрагмент зображення (наприклад, округ з деяким радіусом), який має деякі інваріантні відносно певних перетворень ознаки, тобто ці ознаки зберігаються при розтягу або стисканню, зміні кута огляду, освітлення, тощо.

Дескриптор - особливий ідентифікатор ключової точки, який дозволяє вирізнити та порівнювати ключові точки незалежно від їх розташування.

Матчинг(порівнянням дескрипторів, від англ. «matching») - наз. процес порівняння отриманих дескрипторів між двома зображеннями та пошук співпадаючих. Існують різні підходи для цього, які залежать від алгоритму пошуку ключових точок та формату отриманих дескрипторів.

Rolling shutter (ролінг-шаттер) - це метод отримання зображення, при якому кожен кадр записується не як моментальний знімок, виконаний в певну мить, а як, швидше, сканування, зчитування площини кадру, рядок за рядком — або вертикально, або горизонтально. Іншими словами, не всі частини зображення записуються в один і той самий момент, незважаючи на те, що вихідний знімок пізніше, під час перегляду, буде відображений як єдине ціле — в одній світлинці.

Optical flow (оптичний потік) - є картиною відомого руху об'єктів, поверхонь і країв у візуальній сцені, викликаній відносним рухом між спостерігачем і сценою. В області комп'ютерного зору, оптичний потік це обчислений на основі систем диференціальних рівнянь показник зміни інтенсивності світлового потоку.

ВСТУП

У сучасному світі актуальною є зйомка відео. Відео використовується майже у всіх сферах діяльності - навчання, розвага, спілкування, спорт. Однією з актуальних проблем зйомки відео є його стабільність. У разі екстремальної зйомки, наприклад спуску лижника з гори, коли камера прикріплена до шолома, відео виходить тремтячим, скакучим, розмитим і не приємним для перегляду. Для вирішення цієї проблеми використовують механічну або програмну стабілізацію. Механічна стабілізація використовує спеціальні пристосування для фіксації камери, такі як триноги, камеро візки або стедікама. Програмна стабілізація в свою чергу використовує математичні методи для обробки відео в режимі реального часу або пост процесорної стабілізації, коли відео вже знято. В цілому, стабілізація цифрового відео складається з трьох основних етапів: оцінки руху, звуження руху та синтез кадрів. В цьому документі основна увага приділяється першому етапу. Для оцінки руху використовують один з методів - зіставлення ключових точок або методи оптичного потоку.

Об'єктом дослідження є процес стабілізації відео. Предметом дослідження є методи оцінки руху, та їх вплив на якість стабілізації. Мета роботи є порівняльний аналіз методів оцінки руху. Порівняльний аналіз виконувався на серії тестових відео з видимими артефактами наступного типу: тремтіння, коливання, скручування (*rolling shutter*).

ВИСНОВКИ

У ході цієї роботи було виявлено, що оцінка руху на основі зіставлення ключових точок дає кращій результат, ніж методом оптичного потоку *Pyramidal Lukas & Kanade*, якщо оцінювати рух відео яке є досить тремтячим та має сильні дисторсії. Якщо оцінювати рух відео, яке не має сильного тремтіння, оцінка руху майже однакова з результату трьох методів пошуку ключових точок - *HARRIS*, *FAST* та *SURF*.

За часом стабілізації, найшвидший був алгоритм *FAST*. Він не потребує велику кількість операцій для своєї реалізації та оцінка руху на основі точок цього алгоритму майже не відрізняється від інших алгоритмів.

За якістю стабілізації кращі були відео, у яких використовувалися точки *Harris* та *SURF*. Вони дали набагато стабільніші відео, у порівнянні з *FAST*, в яких кількість чорних країв що потрапляють в кадр більше.

Для стабілізації відео у режимі реального часу краще з усіх підходить алгоритм *FAST* з використанням оптичного потоку, але для гарного результату оцінки руху та стабілізації, треба використовувати додаткову регуляризацию та більше ніж чотири ступені свободи.

Стабілізація відео застосовуючи тільки чотири ступені свободи може дуже гарно себе проявляти при застосуванні до не сильно скакучого чи тремтячого відео. В іншому випадку треба буде вирішувати проблему *роулінг-шаттера* та використовувати всі вісім ступенів свободи.

На майбутнє планується реалізувати повноцінну власну стабілізацію відео з видаленням *роулінг шаттера*.

СПИСОК ЛІТЕРАТУРИ

1. C. Je and H.-M. Park. Optimized Hierarchical Block Matching for Fast and Accurate Image Registration. *Signal Processing: Image Communication*, Volume 28, Issue 7, pp. 779-791, August, 2013.
2. Li, Renxiang; Zeng, Bing; Liou, Ming (August 1994). "A New Three-Step Search Algorithm for Block Motion Estimation". *IEEE Trans. Circuits And Systems For Video Technology*. 4 (4): 438–442.
3. H. Foroosh (Shekarforoush), J.B. Zerubia, and M. Berthod, "Extension of Phase Correlation to Subpixel Registration," *IEEE Transactions on Image Processing*, V. 11, No. 3, Mar. 2002, pp. 188-200.
4. B. S Reddy and B. N. Chatterji, "An FFT-based technique for translation, rotation, and scale-invariant image registration", *IEEE Transactions on Image Processing* 5, no. 8 (1996): 1266–1271.
5. Bruce D. Lucas and Takeo Kanade, An Iterative Image Registration Technique with an Application to Stereo Vision, *Proceedings DARPA Image Understanding Workshop*, April 1981, pp.121-130
6. B.K.P. Horn and B.G. Schunck, "Determining optical flow." *Artificial Intelligence*, vol 17, pp 185–203, 1981.
7. Glyn W. Humphreys and Vicki Bruce (1989). *Visual Cognition*. Psychology Press. ISBN 0-86377-124-6.
8. S. S. Beauchemin; J. L. Barron (1995). *The computation of optical flow*. ACM New York, USA.
9. B. Glocker; N. Komodakis; G. Tziritas; N. Navab; N. Paragios (2008). *Dense Image Registration through MRFs and Efficient Linear Programming*. *Medical Image Analysis Journal*.
10. Moravec, H.: *Obstacle avoidance and navigation in the real world by a seeing robot rover*. In: tech. report CMU-RI-TR-80-03, Robotics Institute, Carnegie Mellon University & doctoral dissertation, Stanford University. Carnegie Mellon University (1980) Available as Stanford AIM-340, CS-80-813 and republished as a Carnegie Mellon University Robotics Institute Technical Report to increase availability.
11. C. Harris and M. Stephens (1988). "A combined corner and edge detector".

- Proceedings of the 4th Alvey Vision Conference. pp. 147–151.
12. Shi and C. Tomasi. Good Features to Track. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 593-600, June 1994.
 13. Zheng, Z., Wang, H., Teoh, E.K.: Analysis of gray level corner detection. *Pattern Recognition Letters* 20 (1999) 149–162
 14. Lowe, David G. "Distinctive Image Features From Scale-Invariant Keypoints". *International Journal of Computer Vision* 60.2 (2004): 91-110.
 15. Bay, Herbert et al. "Speeded-Up Robust Features (SURF)". *Computer Vision and Image Understanding* 110.3 (2008): 346-359.
 16. Mikolajczyk, K., Schmid, C.: Indexing based on scale invariant interest points. In: 8th IEEE International Conference on Computer Vision, Vancouver, Canada, Springer (2001) 525–531
 17. E. Rosten and T. Drummond, "Machine learning for high speed corner detection," in 9th European Conference on Computer Vision, vol. 1, 2006, pp. 430–443.
 18. Guiducci, A.: Corner characterization by differential geometry techniques. *Pattern Recognition Letters* 8 (1988) 311–318
 19. Calonder M., Lepetit V., Strecha C., Fua P.: BRIEF: Binary Robust Independent Elementary Features, CVLab, EPFL, Lausanne, Switzerland
 20. E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. Orb: an efficient alternative to sift or surf. 2011. 1, 2, 3, 4
 21. Alahi, Alexandre, Ortiz, Raphael, and Pierre Vandergheynst, "FREAK: Fast Retina Keypoint IEEE Conference on Computer Vision and Pattern Recognition, 2012.
 22. Marco Zuliani, RANSAC for Dummies, July4, 2014
 23. Matthias Grundmann, Vivek Kwatra, Daniel Castro, Irfan Essa, "Calibration-Free Rolling Shutter Removal" in Proceedings of IEEE Conference on Computational Photography
 24. Olivier Saurer, Kevin Koser, Jean-Yves Bouguet, Marc Pollefeys: Rolling Shutter Stereo, ICCV 2013 IEEE International Conference, 03 March 2014
 25. Rolling Shutter Motion Deblurring. Shuochen Su^{1,2}. Wolfgang Heidrich^{2,1}.

- 1University of British Columbia. 2KAUST.
26. Ja-Ling Wu, "Motion Estimation for Video Coding Standards", National Taiwan University.
 27. Lindeberg, T. 1994. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*, 21(2):224-270.
 28. Kalman, R. E. (1960). "A New Approach to Linear Filtering and Prediction Problems". *Journal of Basic Engineering*. 82: 35.
 29. Tuan Q Pham, Non-maximum Suppression Using Fewer than Two Comparisons per Pixel, ACIVS 2010 December 13-16, Part 1
 30. Mikolajczyk, K., and Schmid, C. 2002. An affine invariant interest point detector. In *European Conference on Computer Vision (ECCV)*, Copenhagen, Denmark, pp. 128-142.
 31. Panchal P M, Panchal S R, Shah S K. A Comparison of SIFT and SURF[J]. *International Journal of Innovative Research in Computer and Communication Engineering*, 2013, 1(2320): 9798.
 32. Edward Rosten, Reid Porter and Tom Drummond, "FASTER and better: A machine learning approach to corner detection" in *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2010, vol 32, pp. 105-119.
 33. Edward Rosten and Tom Drummond, "Fusing points and lines for high performance tracking" in *Proc. Int. Conf. Computer Vision*, October 2005, vol. 2, pp. 1508–1511.
 34. Bruce D. Lucas "Image Matching by the Method of Differences", Carnegie Mellon University, year 1984.
 35. Justin graham, "target tracking with Lucas kanade optical flow ", University of Texas, year 2010.
 36. J. Y. Bouguet, (2001) . Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm. Intel Corporation, 5.
 37. Quinlan, J.R.: Induction of decision trees. *Machine Learning* 1 (1986) 81–106
 38. S. Leutenegger, M. Chli, and R. Siegwart. Brisk: Binary robust invariant scalable keypoints. 2011. 1, 2, 3, 5, 6
 39. Тестове відео - <https://www.youtube.com/watch?v=kf1h78MWwPo>