International Scientific Conference 20 UNITECH 26

INTERNATIONAL SCIENTIFIC CONFERENCE
19-21 November 2026
GABROVO

МНОГОКАДРОВА РЕГИСТРАЦИЯ И ЛОКАЛНА КОРЕКЦИЯ НА 3D ОБЛАЦИ ОТ ТОЧКИ ОТ TIME-OF-FLIGHT КАМЕРА В ЗОНИТЕ НА ДЪЛБОЧИННИ ПРЕКЪСВАНИЯ

Александър Любенов

Технически университет – Габрово, ул. „Хаджи Димитър“ № 4, 5300 Габрово, България
**кореспондиращ автор: alyubenov10@gmail.com*

MULTI-FRAME REGISTRATION AND LOCAL CORRECTION OF 3D POINT CLOUDS FROM A TIME-OF-FLIGHT CAMERA IN DEPTH-DISCONTINUITY REGIONS

Alexander Lyubenov

Technical University of Gabrovo, 4 Hadzhi Dimitar St., 5300 Gabrovo, Bulgaria
**Corresponding author: alyubenov10@gmail.com*

Abstract
Time-of-Flight (ToF) cameras provide dense depth measurements for rapid three-dimensional scene acquisition, but the resulting data are affected by depth-edge artifacts and flying pixels. This paper presents a MATLAB-based multi-frame methodology for registration and localized correction of 3D point clouds acquired with a PMD CamBoard Pico Flexx camera. Three organized XYZ frames were captured from different viewpoints at approximately 0°, 10° and 20°. Neighboring point clouds were registered sequentially using coarse Normal Distributions Transform (NDT) alignment followed by Iterative Closest Point (ICP) refinement. The PC2-to-PC1 and PC3-to-PC2 registrations achieved trimmed RMSE values of 3.033 mm and 2.129 mm with overlap ratios of 82.1% and 88.6%, respectively; the composed PC3-to-PC1 alignment achieved 5.528 mm with 69.2% overlap. Flying-pixel candidates were detected only near depth discontinuities using a local two-surface analysis of the organized depth maps. In the reference frame, 121 intermediate-depth candidates were identified within a 494-point edge region. Of these, 112 unsupported candidates were removed and nine were reconstructed from clean measurements in the registered neighboring frames. The final cloud retained 36,990 of 37,102 reference points (99.70%). The results demonstrate localized correction without global smoothing or aggressive point-cloud reduction.

Keywords: Time-of-Flight camera; point cloud; NDT; ICP; flying pixels; depth discontinuity; multi-frame correction.

ВЪВЕДЕНИЕ

Time-of-Flight (ToF) камерите са активни оптични сензори, които определят разстоянието до наблюдаваната сцена чрез измерване на времезакъснението или фазовото отместване на модулирано инфрачервено излъчване. Основното им предимство е възможността да получават дълбочинна информация за цялото зрително поле при сравнително висока честота на кадрите и в компактна сензорна конфигурация, което ги прави подходящи за роботика, машинно зрение и триизмерна реконструкция [1]. Наред с това ToF измерванията се влияят от систематични и несистематични грешки, свързани с разстоянието, отражателната способност на повърхнините, осветеността и геометрията на сцената [2].

Грешките около прекъсванията по

Print ISSN: 1313-230X, Online ISSN: 3033-1404

дълбочина са особено съществени за 3D реконструкцията. В такива области една измервателна клетка може да получи смесен оптичен принос от повърхнини на преден и заден план и да формира междинна дълбочина, която не съответства на реална физическа повърхнина. Тези артефакти обикновено се означават като летящи пиксели (flying pixels) и възникват предимно около контурите на обектите и ръбове, получени от разлика в дълбочината на измерванията[3]. При директно преобразуване на измерените данни в облак от точки, летящите пиксели се проявяват като разпръснати 3D точки между действителните повърхнини на предния и задния план и могат да влошат последващата регистрация, сегментация и реконструкция.

Многокадровото 3D картографиране с ToF камери изисква надеждно оценяване на относителната поза между последователните наблюдения. Подходите за 3D измервания с ToF сензори комбинират филтриране и регистрация на облаци от точки, за да компенсират ограничената точност и устойчивост на отделните измервания [4]. Класическият Iterative Closest Point (ICP) алгоритъм остава основен метод за локално подравняване на два 3D набора [5], докато триизмерният Normal Distributions Transform (3D-NDT) алгоритъм предоставя подходящ модел за груба регистрация на частично припокриващи се пространствени данни [7]. Различните подходи могат допълнително да използват пространствената повторяемост между няколко ToF наблюдения за корекция на данните [6].

Целта на настоящата работа е да се разработи MATLAB-базирана последователност за организирано получаване на ToF данни, последователна регистрация на облаци от точки и локализирана корекция на артефакти при прекъсвания по дълбочина. Основната хипотеза е, че един летящ пиксел може първо да бъде идентифициран като измерване с междинна дълбочина между две локално наблюдавани дълбочинни повърхнини и след това да бъде валидиран чрез регистрирани измервания от съседни гледни точки. Стабилните измервания от референтния кадър умишлено остават непроменени, така че корекцията да бъде ограничена само до зоните на ръбовете, определени от разликите в дълбочината на измерване.

ЕКСПЕРИМЕНТАЛНА СИСТЕМА И ПОЛУЧАВАНЕ НА 3D ДАННИ

Експерименталната система е базирана на ToF камера PMD CamBoard Pico Flexx с максимална резолюция 224×171 точки, Royale SDK 4.24.0.1201 и MATLAB R2023a. Чрез MATLAB интерфейса на Royale за всеки заснет кадър директно се извличат организираните матрици X, Y и Z. За разлика от подход, при който X и Y се заменят с пикселни индекси, тук и трите координати са физически координати, изразени в метри, и могат непосредствено да се използват за построяване на реален 3D облак от точки.

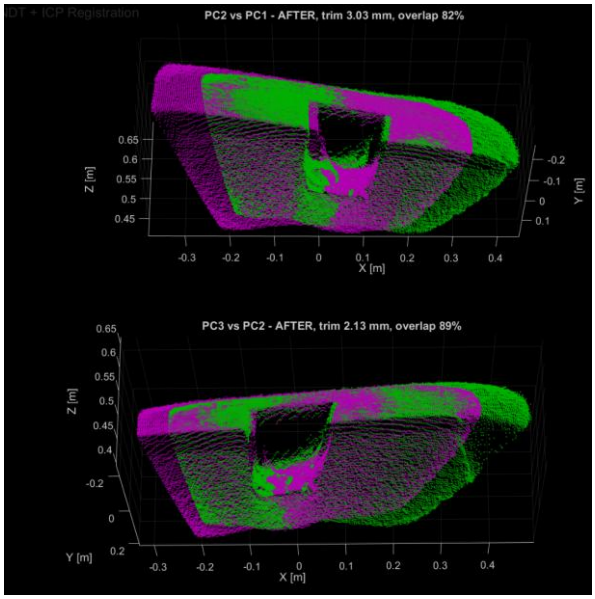
Измерванията с невалидни стойности, както и точките извън предварително зададен работен интервал по Z, се отстраняват чрез логическа маска. В реализацията е използван диапазон 0,25-3,50 m. Липсващите стойности за дълбочината не се запълват, тъй като интерполацията би могла да въведе изкуствени 3D точки, които не са реално измерени от сензора. За всяко наблюдение се запазват както списъкът с валидни XYZ координати, така и оригиналната организирана Z матрица с 171×224 измервания. Финалната тестова сцена съдържа обект на преден план пред по-отдалечен фон, с което по контура на обекта се създава изразено прекъсване по дълбочина.

МНОГОКАДРОВА РЕГИСТРАЦИЯ НА ОБЛАЦИТЕ ОТ ТОЧКИ

Алгоритъмът използва три последователни наблюдения на неподвижна сцена. Първият облак P1 се приема за референтен. Камерата се премества за получаване на второ и трето наблюдение при приблизително 10° и 20° спрямо началната гледна точка, докато сцената остава неподвижна. За да се намали разликата в гледните точки, която трябва да бъде обработена при всяка регистрационна стъпка, PC2 се подравнява към PC1, а PC3 - към PC2; след това двете твърди трансформации се композират, за да се представи PC3 в координатната система на PC1.

$$p' = R \cdot p + t \qquad (1)$$

Всяка съседна двойка се регистрира чрез грубо търсене с начални ориентации, последвано от 3D-NDT и фино ICP доуточняване. Прореждане на точките се използва само вътре в регистрационната процедура, а крайната твърда трансформация се прилага върху оригиналните неподредени XYZ данни. Кандидат-трансформациите се оценяват чрез средноквадратичната грешка на най-близкия съсед (nearest-neighbor RMSE) заедно с коефициента на припокриване, което е по-устойчиво от RMSE (Root Mean Square Error) върху целия облак при частично припокриване на получените от ToF данни[7]. Последователната стратегия намалява ефективното ъглово преместване от приблизително 20° до две съседни регистрационни задачи от приблизително 10°.



Фиг. 1. *Последователна регистрация на съседните кадри след NDT + ICP доуточняване: PC2 към PC1 - RMSE 3,033 mm и 82,1% припокриване; PC3 към PC2 - 2,129 mm и 88,6% припокриване.*

ЛОКАЛНО ОТКРИВАНЕ НА ДВЕ ПОВЪРХНИНИ И МНОГОКАДРОВА КОРЕКЦИЯ

Flying-pixel кандидатите се откриват в оригиналните данни за дълбочината преди прилагането на каквото и да е глобално филтриране върху облака от точки. За всяко валидно измерване се формира локален прозорец 7×7, а валидните Z стойности в него се сортират. Ако най-голямата разлика между две последователни сортирани стойности е поне 30 mm и от двете страни на разделянето има достатъчен брой измервания, локалната околност се приема като съдържаща две физически дълбочинни повърхнини. Представителните им дълбочини се изчисляват чрез медианата съответно на близката и далечната група.

Измерването се класифицира като кандидат с междинна дълбочина, когато неговата Z координата е на повече от 12 mm от медианата и на двете локални повърхнини. След това кандидатите от референтния кадър се проверяват спрямо последователно регистрираните PC2 и PC3. Подкрепящо измерване трябва да удовлетворява зададения пространствен допуск и самото то да не е класифицирано като междинен кандидат в оригиналния си кадър.

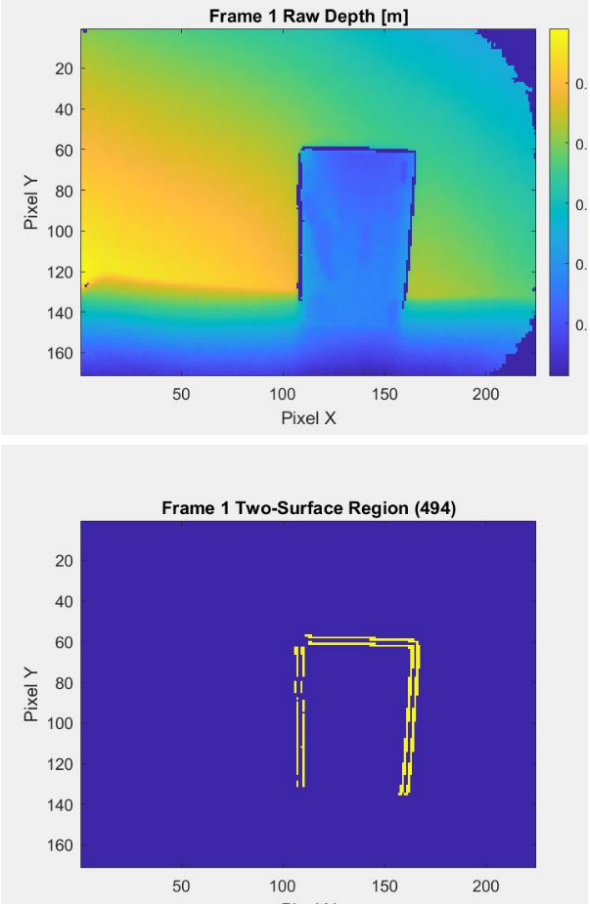
$$d_s = \min(|z - z_near|, |z - z_far|) > T_s \quad (2)$$

Стабилните точки от PC1 се запазват точно във вида, в който са измерени. Ако даден кандидат има едно чисто подкрепящо наблюдение, референтният кандидат се заменя с това измерване. Ако има чиста подкрепа и в двата съседни кадъра, двете подкрепящи XYZ координати се усредняват. Кандидат без чиста многокадрова подкрепа се отстранява. По този начин остатъчните регистрационни грешки не се пренасят върху цялата сцена и корекцията остава ограничена до измерванията с междинна дълбочина около физическите depth edges [6].

РЕЗУЛТАТИ И ДИСКУСИЯ

Финалният експеримент дава 37 102, 37 132 и 37 030 валидни точки съответно в PC1, PC2 и PC3. Последователната регистрация постига чрез т.нар. отсечена средноквадратична грешка (trimmed RMSE) 3,033 mm при 82,1% припокриване за PC2 към PC1 и 2,129 mm при 88,6% припокриване за PC3 към PC2. След композирането на съседните трансформации подравняването PC3 към PC1 има trimmed RMSE 5,528 mm и 69,2% припокриване. Медианните nearest-neighbor разстояния от PC1 са 3,576 mm за регистрирания PC2 и 5,151 mm за регистрирания PC3.

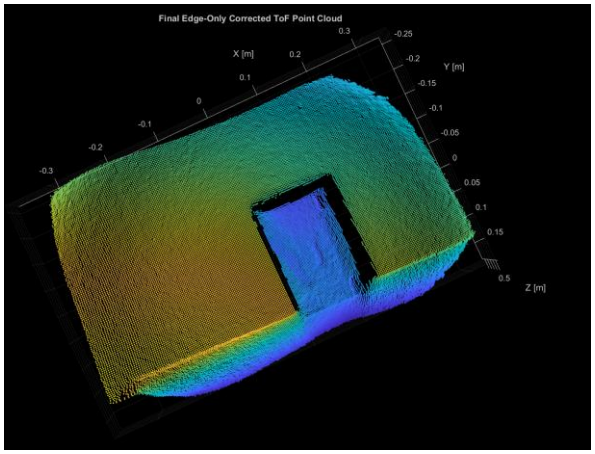
Локалният детектор идентифицира 494, 546 и 495 измервания в зони на depth edges съответно за кадри 1, 2 и 3. В тези зони 121, 143 и 145 измервания са класифицирани като кандидати с междинна дълбочина. В референтния кадър 121-те кандидата представляват 24,5% от откритите измервания в зоната на ръба, а сходният брой кандидати при трите гледни точки показва повторяемо локализиране на един и същ тип артефакти при прекъсванията по дълбочина.



Фиг. 2. *Организирано depth изображение на кадър 1 и локалната two-surface зона, открита с околност 7×7 и минимална разлика между двете повърхнини 30 mm*

Многокадровата валидация на 121-те кандидата от референтния кадър показва, че 112 нямат чисто подкрепящо измерване в нито един от регистрираните съседни кадри и са отстранени. Пет кандидата са реконструирани от едно чисто съседно наблюдение, а четири - от двете

съседни наблюдения, или общо девет реконструирани измервания. Така 92,56% от откритите кандидати с междинна дълбочина са без подкрепа и са премахнати, докато 7,44% са реконструирани. Крайният облак съдържа 36 990 точки, което съответства на 99,70% запазване и само 0,30% общо отстраняване. Тъй като 36 981-те точки от PC1, които не са кандидати, остават непроменени, процедурата действа локално, вместо да изглажда или прорежда целия облак. Тези проценти се отнасят до алгоритмично открити кандидати; при липса на независим ground-truth модел те не се интерпретират като абсолютна точност на откриване на всички летящи пиксели.



Фиг. 3. *Краен ToF облак след локална edge-only корекция. Стабилната геометрия на PC1 е запазена, неподкрепените кандидати с междинна дълбочина са отстранени, а подкрепените кандидати са реконструирани*

Получените резултати потвърждават търсеното поведение на метода: корекцията е концентрирана в зоните на прекъсвания по дълбочина, като същевременно се запазват 99,70% от референтния облак. Подходящи следващи стъпки са отделно изследване на чувствителността към параметрите и оценяване спрямо калибрирана истинна (ground-truth) геометрия, което ще позволи количествено определяне на точността на откриване и избор на праговете независимо от конкретната сцена.

ЗАКЛЮЧЕНИЕ

Разработена и експериментално оценена е MATLAB-базирана последователност за организирано получаване на ToF данни, последователна 3D регистрация и локализирана корекция на артефакти при прекъсвания по дълбочина. Последователната NDT + ICP регистрация осигурява trimmed RMSE 3,033 mm и 2,129 mm за съседните двойки кадри, а композираната регистрация при приблизително 20° промяна на гледната точка достига 5,528 mm. Локалният „two-surface“ детектор идентифицира 121 кандидата с междинна дълбочина в референтния кадър; 112 неподкрепени кандидата са отстранени, а девет са реконструирани от чисти регистрирани измервания. Крайният облак запазва 36 990 от 37 102 точки (99,70%), което показва, че предложената стратегия избягва глобално изглаждане и съхранява оригиналната референтна геометрия извън кандидатната зона. Тъй като не е използвана независима истинна (ground truth) геометрия, отчетената стойност 92,56% представлява дела на алгоритмично откритите кандидати, отстранени чрез многокадрова валидация, а не абсолютния дял на всички летящи пиксели в сцената. Бъдещата работа ще включва анализ на чувствителността към параметрите, оценяване спрямо калибрирана истинна геометрия и изпитвания с допълнителни материали и геометрии на сцената.

ЛИТЕРАТУРА

[1] Foix S., Alenyà G., Torras C. Lock-in Time-of-Flight (ToF) Cameras: A Survey. IEEE Sensors Journal 2011;11(9):1917-1926. doi:10.1109/JSEN.2010.2101060.

[2] Rapp H., Frank M., Hamprecht F.A., Jähne B. A theoretical and experimental

investigation of the systematic errors and statistical uncertainties of Time-of-Flight-cameras. International Journal of Intelligent Systems Technologies and Applications 2008;5(3/4):402-413. doi:10.1504/IJISTA.2008.021303.

[3] Chugunov I., Baek S.-H., Fu Q., Heidrich W., Heide F. Mask-ToF: Learning Microlens Masks for Flying Pixel Correction in Time-of-Flight Imaging. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, p. 9116-9126.

[4] May S., Droeschel D., Fuchs S., Holz D., Nüchter A. Robust 3D-Mapping with Time-of-Flight Cameras. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2009, p. 1673-1678. doi:10.1109/IROS.2009.5354684.

[5] Besl P.J., McKay N.D. A Method for Registration of 3-D Shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 1992;14(2):239-256. doi:10.1109/34.121791.

[6] He Y., Chen S. Error correction of depth images for multiview time-of-flight vision sensors. International Journal of Advanced Robotic Systems 2020;17(4). doi:10.1177/1729881420942379.

[7] Magnusson M., Lilienthal A.J., Duckett T. Scan registration for autonomous mining vehicles using 3D-NDT. Journal of Field Robotics 2007;24(10):803-827. doi:10.1002/rob.20204.