

УДК 681.51:007.52:631.324

Последовательная регистрация дальнометрических и телевизионных данных при построении трехмерной модели внешней среды

Загоруйко С.Н., студент

*Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана,
кафедра «Специальная робототехника и мехатроника»*

*Научный руководитель: Носков В.П., к.т.н.,
зав. сектором НИИСМ МГТУ им. Н.Э. Баумана
Россия, 105005, г. Москва, МГТУ им. Н.Э. Баумана
kafsm7@sm.bmstu.ru*

Традиционно в системах технического зрения (СТЗ) для обеспечения автономного движения используются отдельно или дальнометрические, или телевизионные изображения. Появление быстродействующих сканирующих дальномеров, работающих синхронно с телекамерами, и высокопроизводительной вычислительной техники позволяют на современном этапе строить визуально подобные цифровые модели внешней среды в реальном времени. В данной работе предлагается объединять (комплексировать) информацию от видеокамеры и дальномеров [1] и, получая для каждой отдельной точки обзора текстурированный фрагмент виртуального трёхмерного изображения окружающей среды, объединять эти фрагменты в одной системе координат.

В качестве сенсора был выбран Asus Xtion Pro Live (аналог Microsoft Kinect), имеющий датчик глубины, цветную видеокамеру и микрофонную решетку. Датчик глубины устройства состоит из инфракрасного проектора, объединенного с монохромной КМОП-матрицей, что позволяет датчику получать трёхмерное изображение при любом естественном освещении [2]. Параметры устройства приведены в таблице, внешний вид показан на рис.1.

Параметры устройства сканирования Asus Xtion Pro

Потребление питания	менее 2.5 Вт
Дистанция сканирования	от 0.8 м до 3.5 м
Углы обзора	58° H, 45° V, 70° D
Датчики	цветная видеокамера, сенсор глубины, 2 микрофона
Разрешение сенсора глубины	VGA (640x480) : 30 fps QVGA (320x240): 60 fps
Разрешение видеокамеры	SXGA (1280*1024)



Рис. 1. Внешний вид устройства сканирования Asus Xtion Pro

В разработанном алгоритме решение задачи регистрации ведется с помощью метода Iterative Closest Point (ICP) [3]. На каждой итерации алгоритма производится поиск таких матриц поворота R и перемещения t , при которых функционал

$$F(R, t) = \frac{1}{N} \sum_{i=1}^N \|Rx_i + t - y_i\|^2 \quad (1)$$

принимает наименьшее значение. Здесь x_i и y_i - пары точек в разных сканах, соответствие между которыми находится с помощью поиска дескрипторов на телевизионных изображениях и вычисления лучших совпадений.

Поиск минимума функционала (1) проводится на основе метода двойных кватернионов, описанного в [3].

Алгоритм выполняется в цикле до достижения сходимости, причем на каждом шаге сравниваемое облако точек переводится в новую систему координат. При последовательном во времени выполнении алгоритма с сохранением прошлых состояний строится трехмерная модель окружающей среды.

Библиотека OpenNI позволяет получать с датчика глубины данные, откалиброванные по данным с видеокамеры сенсора, и проводить преобразования из двумерной проекционной системы координат в трехмерную.

Библиотека OpenCV используется для обработки изображений, полученных с сенсора, и других вычислений. Для поиска пар соответствующих точек на двух изображениях используется следующий алгоритм:

- 1) поиск и вычисление локальных дескрипторов SIFT [4];
- 2) совмещение дескрипторов с помощью быстрого приближенного поиска ближайших соседей;
- 3) выбор пар дескрипторов по нескольким критериям.

Локальные дескрипторы представляют собой вектора признаков, построенные по отдельным фрагментам изображения. То есть они не описывают все изображение, а содержат информацию только о выбранных некоторым способом фрагментах. Самыми известными локальными дескрипторами являются SIFT (Scale Invariant Feature Transform), SURF [5] (Speeded Up Robust Features), PCA-SIFT [6] (PCA – Principal Component Analysis) и т.д. Метод SIFT [4] (Scale Invariant Feature Transform) обнаруживает и описывает

локальные особенности изображения. Получаемые с помощью него признаки инвариантны относительно масштаба и поворота, устойчивы к ряду аффинных преобразований, шуму, изменению в освещении. Данный алгоритм можно разделить на две части: определение «точек интереса» и построение дескрипторов окрестностей данных точек. Существует несколько способов определения точек интереса. Алгоритм, предложенный в рамках SIFT, один из самых известных. Он заключается в использовании пирамиды Гаусса, которая строится для изображения. Далее изображения приводятся к одному размеру, и вычисляется их разность (DoG, difference-of-Gaussian images). Причем в качестве кандидатов точек интереса выбираются только те пиксели, которые сильно отличаются от остальных, это делается, например, путем сравнения каждого пикселя изображения с несколькими соседними данного масштаба, с несколькими соответствующими соседями в большем и меньшем масштабе. Пиксель выбирается как точка интереса только в том случае, если его яркость является экстремумом. Далее для каждой такой точки интереса вычисляется локальный дескриптор, характеризующий направление градиентов в пикселях некоторой окрестности. Главным минусом SIFT дескрипторов является их высокая размерность и большое количество на изображении. PCA-SIFT дескриптор – одна из вариаций SIFT, в которой уменьшается размерность дескриптора с помощью анализа главных компонент. Это достигается с помощью нахождения пространства собственных векторов, на которое впоследствии проецируются вектора признаков. SIFT менее чувствителен к изменению освещения или угла, под которым был сделан снимок.

Совмещение дескрипторов с помощью быстрого приближенного поиска ближайших соседей производится с помощью библиотеки FLANN, интегрированной в OpenCV.

Найденные дескрипторы фильтруются по расстоянию и из соображений возможного перемещения сенсора. Для фильтрации по расстоянию отбрасываются все пары с расстоянием больше среднего расстояния среди всех пар. По возможному перемещению сенсора накладывается ограничение на относительное перемещение между координатами пикселей дескрипторов. Для этого используется двумерный алгоритм ICP по контурам изображения.

Поиск контуров на изображениях производится с помощью оператора Собеля [7]. Перед этим цветность изображения переводится в градации серого. Пиксели, значения интенсивности которых после применения оператора Собеля больше некоторого порога, формируют вектора x и y - вход двумерного алгоритма ICP.

Для поиска матриц R и t в двумерном алгоритме используется линейаризация компонент R :

$$R = \begin{vmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{vmatrix} \approx \begin{vmatrix} 1 & -\gamma \\ \gamma & 1 \end{vmatrix} \quad (2)$$

И производится следующая оптимизация:

$$\hat{M}_{opt} = \arg \min \sum_{i=1}^n \left((\hat{M}x_i - y_{f(i)})n_i \right) \quad (3)$$

где $f(i)$ - функция, возвращающая индекс ближайшей точки к x_i из вектора y , n_i - нормаль в этой точке, построенная по k соседним к ней точкам, а матрица M задана как:

$$M = \begin{vmatrix} 1 & -\gamma & t_x \\ \gamma & 1 & t_y \\ 0 & 0 & 1 \end{vmatrix} \quad (4)$$

Матрица (3) вычисляется с помощью SVD разложения [8]. По найденной матрице M из (4) находятся R и t . Алгоритм выполняется итеративно до достижения сходимости, причем во время выполнения куммулятивно заполняется вектор полного перемещения s . Вектор s используется для сравнения с векторами перемещений найденных пар дескрипторов. Пары отбрасываются, если углы между векторами не равны или их длины различаются.

Пример изображений с камеры Asus Xtion Pro с найденными описанным алгоритмом парами отсеянных дескрипторов показан на рис.2.

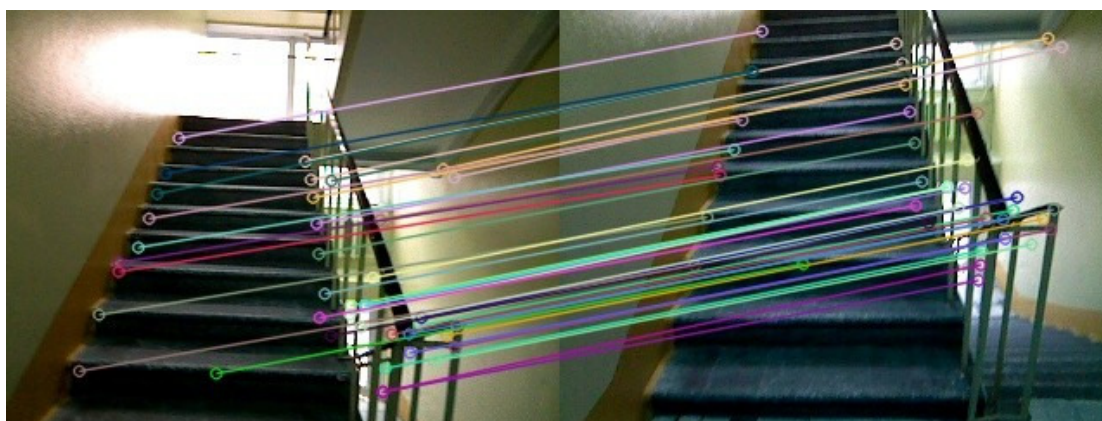


Рис. 2. Пример пар дескрипторов SIFT, найденных на изображениях с камеры Asus Xtion Pro

Координаты найденных точек на обоих изображениях переводятся в трехмерную систему координат с центром в точке наблюдения (позиции камеры) с помощью Молодежный научно-технический вестник ФС77-51038

библиотеки OpenNI, Ось Z выбранной системы координат направлена вверх от сенсора, ось Y совпадает с оптической осью сенсора. Таким образом, для каждого скана получаются идущие на вход алгоритма ICP вектора точек описанного выше.

Пример совмещения сканов описанным алгоритмом показан на рис.3.

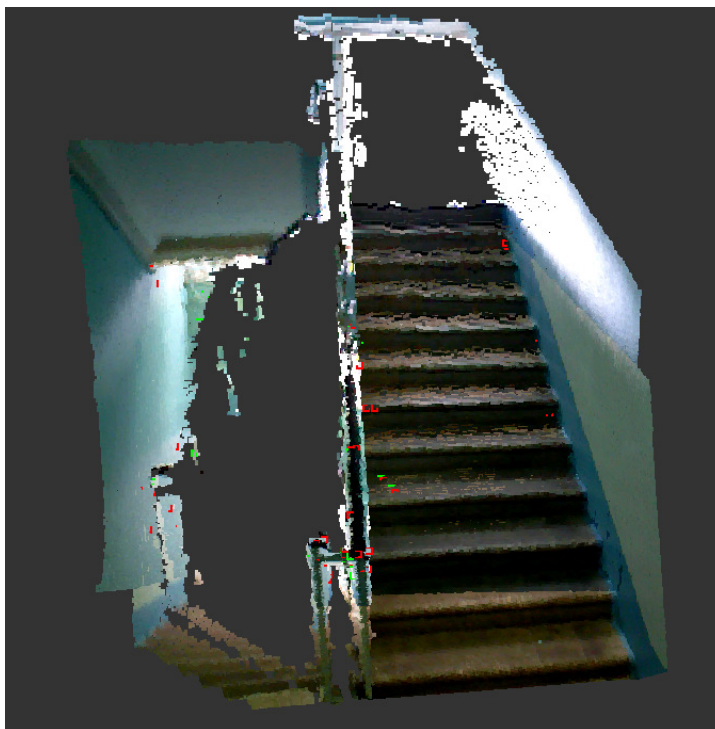


Рис. 3. Пример регистрации пары комплексированных изображений

Тестирование реализованного алгоритма показало эффективность методов обработки изображений при построении трехмерных моделей и высокую вычислительную эффективность метода двойных кватернионов при регистрации. Таким образом, разработанный алгоритм позволяет проводить последовательную регистрацию в режиме реального времени в условиях ограниченной вычислительной мощности. Это особенно важно при создании мобильных устройств сканирования, где вычислительная мощность ограничена временем автономной работы, массой и другими факторами. В отличие от [1], где для решения задачи регистрации использовался алгоритм двумерной экстремальной навигации, в разработанном методе используется трехмерный алгоритм, позволяющий перемещение и вращение сенсора в любых направлениях при построении объемной модели.

Разработанный алгоритм позволяет получить построенную по серии снимков объемную цветовую модель окружающей среды, которая может быть использована в контуре дистанционного управления мобильными роботами и манипуляторами. Другим важным приложением объемной цветной модели внешней среды является ее

использование ее для предварительного виртуального проигрывания различных сценариев будущих работ, операций или боевых действий в экстремальных условиях (ликвидация последствий аварий и катастроф, разминирование, борьба с террористами и т.п.).

Список литературы

1. В.П. Носков, И.В. Рубцов, А.Ю. Романов. Формирование объединенной модели внешней среды на основе видеокамеры и дальномера. // Мехатроника, автоматизация, управление. 2007. №8, С. 2-5.
2. Paul J. Besl, Neil D. McKay. A method for registration of 3-D shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 14(2), pp. 239-256, 1992.
3. Zhengyou Zhang. Iterative point matching for registration of free-form curves. 1992. // URL <http://research.microsoft.com/~zhang/Papers/IJCV-94-ICP.pdf>.
4. L. David, Object recognition from local scale-invariant features // Proceedings of the International Conference on Computer Vision.
5. H. Bay, T. Tuytelaars, L. V. Gool, SURF: Speeded Up Robust Features // Computer Vision and image Understanding, 2008.
6. Y. Ke, R. Sukthankar, PCA-SIFT: A More Distinctive Representation for Local Image Descriptors // URL <http://www.cs.cmu.edu/~rahuls/pub/cvpr2004-keypoint-rahuls.pdf>
7. URL http://en.wikipedia.org/wiki/Sobel_operator.
8. William H. Press, Saul A. Teukolsky, William T. Vetterling, Brian P. Flannery. 2.6 Singular Value Decomposition // Numerical Recipes in C. — 2nd edition. — Cambridge: Cambridge University Press.