

ФИЛЬТРАЦИЯ ДАННЫХ ГЛУБИНЫ С СЕНСОРА KINECT

Дусеев В.Р.¹, Мальчуков А.Н.¹, Мыцко Е.А.¹

¹Национальный исследовательский Томский политехнический университет, Томск, Россия (634050, Томск, пр. Ленина, 30), e-mail: vagiz.d@gmail.com

В связи с ограничениями технологии структурированного света, используемого сенсором Kinect, во время сканирования объектов образуется значительный шум. Данная статья предлагает решение, основанное на фильтре Калмана и алгоритмах закрашивания изображений с целью увеличить стабильность карты глубины в отдельном отрезке времени и закрасить сокрытые от сенсора области. Используемый алгоритм комбинирует данные предыдущего и текущего кадров. Коэффициент комбинации соответствует скорости изменения карты глубины. Отфильтрованные кадры итеративно включаются в процесс фильтрации. Области без данных заполняются с использованием алгоритмов закрашивания на основе методов FMM. Продемонстрирована энтропия каждого пикселя на карте глубины. Предлагаемый подход может быть использован в качестве предварительной обработки карты глубины перед её непосредственным использованием для целей распознавания изображений.

Ключевые слова: Kinect сенсор, глубина зондирования камеры, фильтр Калмана, быстро шагающий метод, метод закрашивания изображений, карта глубины предварительной обработки.

DEPTH DATA FILTERING WITH SENSOR KINECT

Duseev V.R.¹, Malchukov A.N.¹, Mytsko E.A.¹

¹National Research Tomsk Polytechnic University, Tomsk, Russia (634050, Tomsk, Lenin Avenue, 30), e-mail: vagiz.d@gmail.com

Due to limitations of structured light technology used sensor Kinect, during scanning of objects formed a substantial noise. This article proposes a solution based on the Kalman filter algorithms and painting images in order to increase the stability of the depth map in a separate period of time and paint hidden from the sensor area. The algorithm combines the data of the previous and current frames. Combination coefficient corresponds to the rate of change of the depth map. Iteratively filtered frames are included in the filtering process. Finally, the area with no data are filled using algorithms based on painting techniques FMM. The entropy of each pixel was presented. The proposed approach can be used as a preprocessing before using depth maps for image recognition.

Keywords: Kinect sensor, depth sensing camera, Kalman filter, fast marching method, in-painting techniques, depth map preprocessing.

Сканирующие сенсоры, подобные Microsoft Kinect [1,8], Asus X-tion и PrimeSense стали дешёвой альтернативой лазерным сканерам в самых разных областях, таких как робототехника, естественное взаимодействие с компьютером, ориентирование внутри помещения. Kinect был изначально разработан для распознавания движений тела человека и его жестов в компьютерных играх. Тем не менее, возможности Kinect завоевали сенсору популярность и в других областях [2]. Особое распространение Kinect получил среди инженеров-энтузиастов. Программисты и инженеры со всего мира получили возможность использовать в своей работе дешёвый трёхмерный сканер, что вылилось во множество реализованных проектов.

Сенсор Kinect считывает цветовые и глубинные данные одновременно на частоте до 30 кадров в секунду. Разрешение инфракрасного сенсора составляет $1,280 \times 1,024$ пикселей, однако максимальный размер выходного изображения не превышает 640×480 пикселей. Это

связано с привязкой к возможностям внутреннего процессора Kinect и пропускным возможностям шины USB.

В данной работе разрешение захватываемой для последующей обработки карты глубины является относительно небольшим и содержит 320×240 пикселей. Тем не менее, это означает, что каждый кадр содержит 76,800 точек с данными. Измеренные дистанции, предоставляемые сенсором, нормализованы и разбиты на дискретные значения от 0 до 2,047. Каждая точка карты глубины представляет собой 11-битное целое значение, в котором один бит отведён для того, чтобы пометить пустые пиксели, не содержащие данных. Общая стабильность полученных изображений низкая.

Принцип работы сенсора Kinect заключается в следующем: проецируется заранее известный инфракрасный шаблон на поверхность (рис. 1) с помощью встроенного инфракрасного проектора, измеряя разницу между полученной проекцией и сохранённым в памяти устройства шаблоном. Выходная карта глубины содержит шум и области, не содержащие данных, что обусловлено тенями, которые отбрасывают различные объекты при освещении их инфракрасным проектором. На рис. 1 представлен обильный шум на необработанной карте глубины, представленной в оттенках серого цвета.

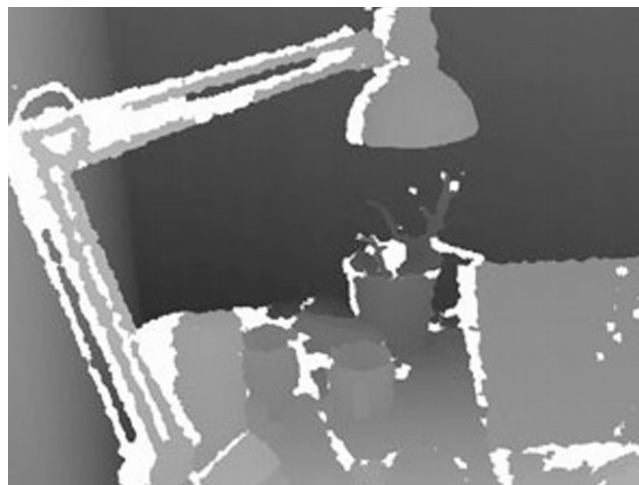


Рис. 1. Обильный шум на необработанной карте глубины, представленной в оттенках серого цвета

Для улучшения качества выходного изображения сенсора был предложен подход, позволяющий исключить шум на изображении и заполнить пустые области такими данными, которые бы с определённой точностью отражали действительные значения глубины в этих точках.

Поскольку Kinect был разработан с целью получения данных о глубине рабочей сцены, данная работа фокусируется на карте глубины. Предлагаемый подход обработки изображения представлен на рис. 2 и состоит из двух частей:

а. Закрашивание сокрытых областей. В данном разделе предложен алгоритм закрашивания пустых областей.

б. Шумоподавление состоит из предотвращения колебаний отдельных пикселей и целых областей на изображении.

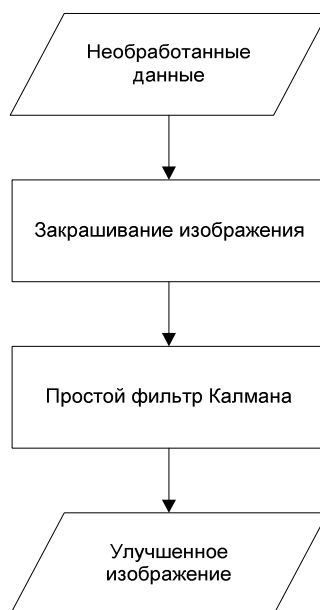


Рис. 2. Предлагаемый подход для уточнения исходных данных

Технология работы Kinect

Kinect испускает инфракрасную сетку и получает результат при помощи CMOS камеры с инфракрасным фильтром. Внутренний процессор, разработанный компанией PrimeSense, вычисляет разницу между полученным снимком проекции инфракрасной сетки и зашитым в память контроллера шаблоном. Когда отдельная точка спроецирована на объект, расположенный ближе или дальше, чем расчётная плоскость проекции, её позиция, очевидно, будет смещена. Измерение смещения каждого пикселя посредством процедуры корреляции изображений позволяет вычислить матрицу несоответствия, которая позволяет Kinect определить истинную дистанцию от плоскости устройства до любой точки на сканируемой поверхности.

Тени в глубине изображения

Небольшое расстояние между инфракрасным проектором и инфракрасной камерой может привести к образованию тени на той стороне объекта, которая находится ближе к камере, чем к проектору. Данная проблема представлена на рис. 3.

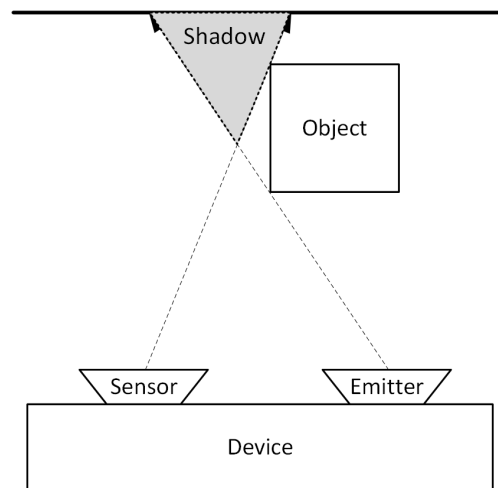


Рис. 3. Схема лазерного пути, в случае если бы она была перекрыта объектом

Простой метод закрашивания

Для закрашивания затенённой области, могут быть использованы простейшие алгоритмы. Например, для каждого пикселя внутри области, которая не содержит в себе никаких данных о глубине, запускается простая процедура. Рассматриваются значения ближайших значимых пикселей с каждой из четырёх сторон. Затем среднеарифметическое значение этих четырёх пикселей присваивается текущей рассматриваемой точке, не содержащей данных. Данный подход представлен на рис. 4

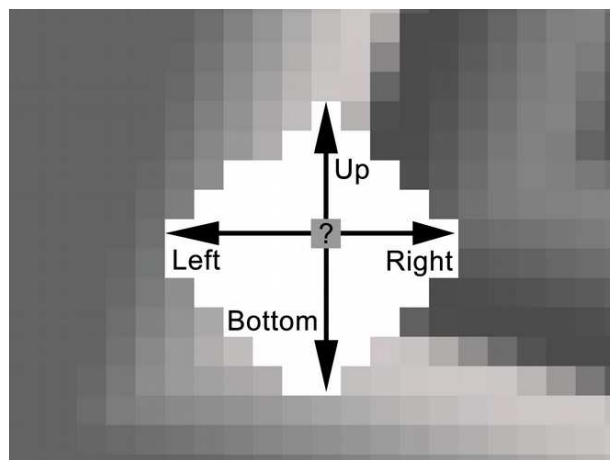


Рис. 4. Схема простого алгоритма закрашивания пустых пикселей

Однако данный алгоритм позволяет получить только результаты низкого качества. Указанный алгоритм использует лишь информацию о ближайших пикселях и вовсе не рассматривает какое-либо значение градиента в данной области.

Быстро шагающий алгоритм

Более сложный алгоритм заполнения пустой области описан в статье [10]. Данный алгоритм основан на быстро шагающем методе (FMM, Fast Marching Method). Алгоритм начинает работу на границе пустой области и продвигается внутрь, в первую очередь, заполняя пиксели на границе области. Процедура заполнения принимает в расчёт небольшую

область пикселей вокруг точки на границе, не содержащей данных. Данная точка приравнивается значению нормализованной взвешенной суммы всех рассматриваемых пикселей во взятой области. Учёт весов пикселей является важной деталью данного алгоритма. Каждый соседствующий пиксель получает свой вес в зависимости от его расстояния до рассматриваемой точки. Наибольший вес получают пиксели, лежащие непосредственно рядом с точкой, рядом с нормалью к границе пустой области и те, что находятся на контурах данной границы. Как только значение текущего пикселя определено, процедура перемещается к следующей неизвестной точке, используя FMM, который обеспечивает заполнение пустых областей рядом с уже известными пикселями в первую очередь [9]. Рис. 5 иллюстрирует закрашиваемую маску карты глубины, представленной на рис. 1. Белые области отмечены также внутри пикселей в реализации алгоритмов. Границы областей белого помечены как BAND.

Результат применения второго подхода, основанного на быстро шагающем методе, продемонстрирован на рис. 6. Метод FMM позволяет получить результаты, схожие с описанными в статьях [5] и методе Olivera [6]. Тем не менее, время работы FMM во многих случаях меньше, чем, например, у алгоритма BSCB [9]. Более того, представленный метод достаточно легко реализовать.



Рис. 5. Закрашиваемая маска карты глубины, представленной на рис. 1.



Рис. 6. Результат работы алгоритма закрашивания, основанный на FMM

Понимание проблемы

Даже в случае статичного положения устройства, значение пикселя будет колебаться в небольшом диапазоне. В результате (из-за колебания значения пикселя) изображение становится крайне нестабильным. Очевидно, что во временном отрезке, содержащем 10,000 кадров, приблизительно 2,5% сохраняют одно значение, 20% колеблются между двумя значениями, 38% перескакивают между тремя различными уровнями и остальные пиксели имеют 4 и более уровней. Естественно, среднее число уровней у пикселей увеличивается с ростом длины рассматриваемого временного отрезка [7].

Рис.7 демонстрирует энтропию каждого пикселя на карте глубины, вычисленную на временном отрезке, содержащем 1000 кадров с расстоянием в 1,0 м от сканируемой плоскости. Ярко выраженные полосы увеличенной энтропии вызваны одним из адаптивных алгоритмов, который используется процессором PrimeSense. Вне полос разброс значений энтропии приблизительно одинаков.

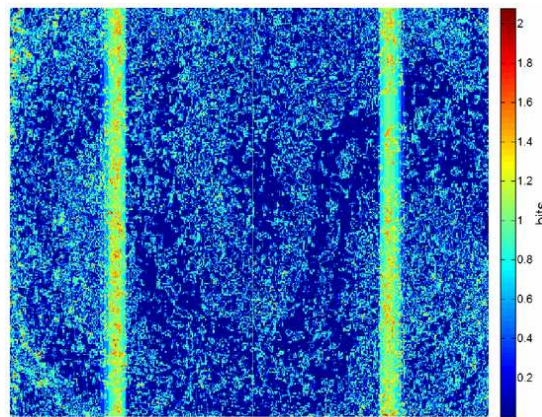


Рис. 7. Энтропия каждого пикселя вычисляется с помощью 1000 кадров на расстоянии 1,0 м.

Упрощенный фильтр Калмана

Описанные колебания и мерцания пикселей могут быть устранены с помощью фильтра Калмана [3]. Данный фильтр представляет собой аккумулирующий буфер, в котором текущий кадр комбинируется с предыдущим через определённый коэффициент. Вес коэффициента, который приписывается кадру Z_{k+1} определяется коэффициентом Калмана K , который вычисляется итеративно:

$$d_{k+1}^{opt} = K \cdot z_{k+1} + (1 - K) \cdot (d_k^{opt} + f_k),$$

где d_k^{opt} представляет аккумулирующий буфер, $f(k)$ – выражение, которое описывает поведение каждого пикселя во времени.

Тем не менее, оригинальный подход не может быть использован, так как мы не в состоянии обеспечить $F(K)$ выражение. Таким образом, итерационное выражение для расчета коэффициента веса не имеет смысла, поэтому, мы используем упрощенный буфер

накопления. Весовой коэффициент K ($0,0 \leq \lambda \leq 1,0$) представляет собой количество нестабильных пикселей в кадре, а не итеративный коэффициент. Количество нестабильных точек вычисляется во время сравнения текущих и предыдущих кадров, где каждая изменяемая точка рассматривается в качестве нестабильной.

Буфер накопления не только уточняет края объектов, но и повышает стабильность всего изображения.

Заключение

Существуют многие параметры, которые влияют на датчик в данной ситуации. Например, датчик в состоянии обнаружить какие-либо отражающие материалы. Кроме того, любые тени от наблюдаемого объекта увеличивает количество не выявленных и пустых мест. Предложен подход, основанный на алгоритме закрашивания Александра TELEA и упрощенном фильтре Калмана. Описанный в работе подход улучшает общую стабильность изображения и заполняет красками пустые области. В дальнейших работах могут быть реализованы более сложные алгоритмы, например, совместный двусторонний фильтр или алгоритм восстановления глубины, как представлено в [4], который использует сегментацию изображения наравне с голосованием для того, чтобы восстановить недостающую информацию о глубине.

Список литературы

1. Дусеев В.Р., Рудь М.Н., Мальчуков А.Н. Исследование методов фильтрации карты глубины с сенсора KINECT [Электронный ресурс] // Молодежь и современные информационные технологии: сборник трудов XI Международной научно-практической конференции студентов, аспирантов и молодых ученых, Томск, 13-16 Ноября 2013. - Томск: ТПУ, 2013 - С. 143-145. - Режим доступа: [tp://www.lib.tpu.ru/fulltext/c/2013/C04/C04.pdf](http://www.lib.tpu.ru/fulltext/c/2013/C04/C04.pdf)
2. Andrew D.W. Using a depth camera as a touch sensor. In Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, Saarbrücken, Germany, 7–10 November 2010.
3. Kalman R.E. "A new approach to linear filtering and prediction problems". Journal of Basic Engineering 82 (1): 1960. - pp. 35-45.
4. Kokaram A.C., Morris R.D., Fitzgerald W.J., Rayner P.J.W.; "Interpolation of missing data in image sequences".
5. M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. "Image Inpainting." In Proceedings SIGGRAPH 2000, Computer Graphics Proceedings, Annual Conference Series, edited by Kurt Akeley, pp. 417—424, Reading, MA: Addison-Wesley, 2000.

6. M. Oliveira, B. Bowen, R. McKenna, and Y. -S. Chang. “Fast Digital Image Inpainting.” In Proc. VIIP 2001, pp. 261—266, 2001.
7. M.R. Andersen, T. Jensen, P. Lisouski, A.K. Mortensen, M.K. Hansen, T. Gregersen and P. Ahrendt. Kinect depth sensor evaluation for computer vision applications. Electrical and Computer Engineering Technical Report ECE-TR-6. Aarhus University, 2012.
8. Microsoft Kinect. [Электронный ресурс]: <http://www.xbox.com/en-us/kinect/> (дата обращения: 17 марта 2014).
9. OpenCV 3.0.0 documentation. Image Inpainting Available online: http://docs.opencv.org/trunk/doc/py_tutorials/py_photo/py_inpainting/py_inpainting.html (дата обращения 20 июня 2014).
10. Telea, Alexandru. “An image inpainting technique based on the fast marching method.” Journal of graphics tools 9.1., 2004.

Рецензенты:

Авдеева Д.К., д.т.н., профессор, директор ООО «Медприбор», г. Томск;

Ким В.Л., д.т.н., профессор кафедры вычислительной техники Института кибернетики, ФГАОУ ВО НИ ТПУ, г. Томск.