

УДК 004.021

М.А. Калинин, В.Ю. Карпычев

МЕТОДЫ ВЫЧИСЛЕНИЯ ОПТИЧЕСКОГО ПОТОКА ДЛЯ ЗАДАЧ АНАЛИЗА ТРЕХМЕРНОЙ СУЩНОСТИ СЦЕНЫ

Нижегородский государственный технический университет им. Р.Е. Алексеева

Рассматривается задача восстановления трехмерной структуры сцены с помощью использования свойств оптического потока. Приводятся результаты применения различных методов вычисления оптического потока для данной задачи. Предоставляются математические основы дифференциальных методов вычисления оптического потока.

Ключевые слова: компьютерное зрение, оптический поток, алгоритм Лукаса-Канаде, алгоритм Фарнебака.

В современных компьютерных системах и системах обработки изображений часто возникает задача определения формы объектов, а также перемещения этих объектов в трехмерном пространстве. Для решения данной задачи существует множество подходов, и один из них – это использование свойств оптического потока.

Оптический поток - это изображение видимого движения объектов, поверхностей или краев сцены, получаемое в результате перемещения наблюдателя (глаз или камеры) относительно сцены. В настоящее время оптический поток находит все новые области своего применения: он интенсивно используется для создания видео и спецэффектов, сжатия видео, стабилизации видео, компьютерном зрении и многих других областях. В оптическом потоке также зашифрована полезная информация о трехмерной структуре сцены. В [3] приводится простой мысленный эксперимент, показывающий это: представьте себя наблюдающим из окна движущегося автомобиля, можно заметить, что удаленные объекты характеризуются гораздо более медленным кажущимся движением по сравнению с близкими объектами, следовательно, скорость кажущегося движения уже позволяет получить определенную информацию о расстоянии.

Таким образом, имея на входе некоторую последовательность кадров, полученную с помощью оптического сенсора или камеры, можно попытаться восстановить запечатленное на них трехмерное пространство. В настоящей статье рассматривается применение различных методов вычисления оптического потока к решению задачи анализа трехмерной сущности и структуры сцены.

Следует отметить, что до настоящего времени для данной задачи не проводилось каких-либо серьезных научных работ. Существует несколько программных продуктов, частично решающих данную задачу, но применяемые алгоритмы и методы решения, используемые в них, закрыты для всеобщего обозрения и дальнейшего исследования.

Варианты расчета оптического потока

Суть оптического потока в том, что для каждой точки некоторого изображения находится такой сдвиг (dx , dy), чтобы исходной точке соответствовала точка на втором, последующим изображении. В [1] оптический поток определяется как векторное поле, где каждый вектор показывает перемещение отдельной точки между двумя последующими кадрами (рис. 1). Для нахождения этого сдвига нужно выбрать какую-либо характеристику точки, которая не изменяется в результате смещения. Обычно такой характеристикой выступает интенсивность (яркость) или цвет (для цветных изображений). Такой подход естественно может давать сбои при сильном изменении освещения сцены или изменения угла падения света. Од-

нако если мы рассматриваем видеопоток, то в большинстве случаев сильного изменения освещения между двумя кадрами не происходит из-за небольшого временного интервала между двумя кадрами.



Рис. 1. Два последующих кадра и векторное поле перемещений

Всего существует два варианта расчета оптического потока: плотный (dense) и выборочный (sparse). Выборочный поток рассчитывает сдвиг отдельных заданных точек (например, некоторых характеристических точек, предварительно найденных одним из алгоритмов детектирования таких точек). Плотный поток считает сдвиги для всех точек изображения. Естественно, выборочный поток вычисляется быстрее (далее будет произведен сравнительный анализ двух существующих реализаций для нахождения плотного и выборочного потока), однако для некоторых алгоритмов разница во времени не такая и большая, а для некоторых задач требуется нахождение потока во всех точках изображения.

Обзор существующей реализации алгоритмов вычисления оптического потока

Библиотека алгоритмов компьютерного зрения и обработки изображений OpenCV предлагает реализацию нескольких популярных и наиболее производительных алгоритмов. Такими являются алгоритм Лукаса-Канаде, а также алгоритм Фарнебака.

Классический вариант алгоритма Лукаса-Канаде реализован в OpenCV в виде функции *calcOpticalFlowPyrLK*. Данный алгоритм использует выборочный поток, и для некоторого определенного массива точек (для поиска таких точек библиотекой рекомендуется детектор углов Shi-Tomasi) на первом кадре оценивает их расположение на последующем. Входные и выходные параметры данной функции следующие: два изображения(или кадра), следующие друг за другом, массив начального набора характеристических точек, массив результирующих точек, массив который показывает статус поиска каждой точки, массив погрешностей для каждой операции поиска, размер окна по которому производится гауссово усреднение, дополнительные флаги, которые можно использовать для установки начального приближения потока или выбрать метод оценки погрешности, а также пороговое значение градиента.

В результате работы эта реализация стабильно и с хорошей точностью способна обнаруживать как маленькие, так и большие сдвиги. Время работы алгоритма – около 10 мс. для 400 точек (тестирование проводилось на процессоре Intelcore i7 950).

Алгоритм Фарнебака реализован в OpenCV в виде функции *calcOpticalFlowFarneback*. Данная реализация использует плотный поток, поэтому считает сдвиг для каждой точки и нет никакой необходимости использовать дополнительные детекторы точек, как в случае с алгоритмом Лукаса-Канаде. Входные и выходные параметры данной функции: два изобраа-

жения (или кадра), следующие друг за другом, выходной поток в формате двухканальной матрицы (матрица векторов), дополнительные флаги, которые можно использовать для установки начального приближения потока или выбрать метод оценки погрешности и пороговое значение градиента.

Этот алгоритм менее стабилен и плохо работает на больших сдвигах. Время работы алгоритма-около 600 мс. для изображения 512x512 пикселей (тестирование проводилось на процессоре Intelcore i7 950).

Восстановление трехмерной структуры сцены

После вычисления оптического потока мы имеем набор точек на исходном кадре и набор соответствующих им точек на последующем кадре. Взяв пару соответствующих точек, можно выполнить триангуляцию и определить координаты их прообраза в трехмерном пространстве (рис. 2).

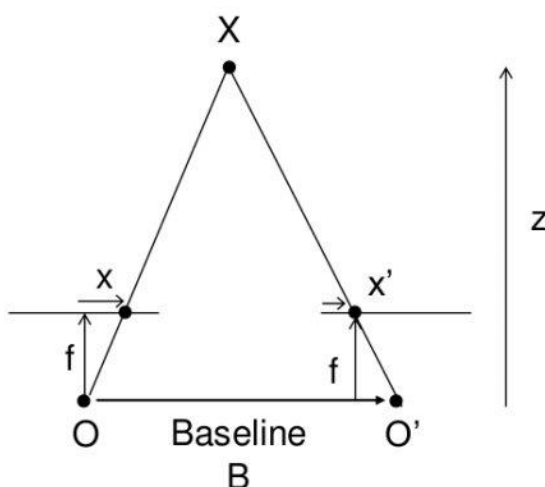


Рис. 2. Триангуляция пары соответствующих точек

Рассмотрим подробнее рис. 2. Здесь расстояние до объекта обозначено буквой z и вычисляется, используя метод подобных треугольников по следующей формуле:

$$Z = \frac{Bf}{\Delta x}$$

где B – это шаг камеры (расстояние на которое переместилась камера в реальном мире), f – фокусное расстояние камеры (расстояние от центра камеры до поверхности ее экрана), Δx – расстояние между соответствующими точками. Если у нас нет встроенного шагомера и мы не знаем, на какую дистанцию переместилась камера в реальном мире, можно немного упростить задачу и предположить, что движение камеры происходит плавно, расстояние между двумя соседними кадрами примерно одинаковое и принять B за единицу.

Используя данный метод, применим его для соответствующих точек, полученных с помощью OpenCV реализации алгоритма Лукаса-Канаде и алгоритма Фарнебака.

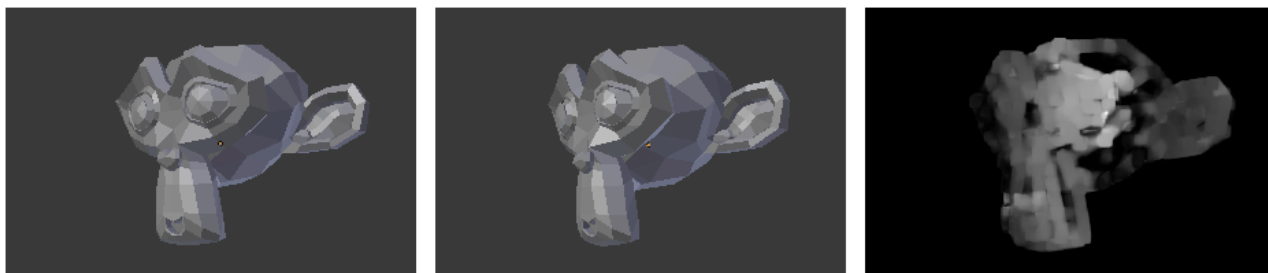


Рис. 3. 1 – первый кадр. 2 – второй кадр. 3 – карта глубин, полученная с помощью оптического потока Фарнебака, более светлым цветом выделены области, расположенные ближе к наблюдателю

На рис. 3 показан результат работы для алгоритма Фарнебака. Для алгоритма Лукаса-Канаде такой карты глубины получить не удалось, ввиду маленького количества характеристических точек. Однако для отдельных характеристических точек на сцене, этот алгоритм работает намного точнее и способен искать совпадения на больших сдвигах, поэтому сейчас ведется работа над объединением свойств этих алгоритмов для корректировки результатов, полученных с помощью алгоритма Фарнебака.

Математические основы дифференциальных методов вычисления оптического потока

В основе всех дифференциальных методов вычисления оптического потока (таких как алгоритм Лукаса-Канаде и алгоритм Фарнебака) лежит одно важное предположение. Предположим, что значения пикселей переходят из одного кадра в следующий кадр без изменений. Таким образом, мы делаем допущение, что пиксели, относящиеся к одному и тому же объекту, могут сместиться в какую-либо сторону, но их значение останется неизменным. Глобальные условия освещения и освещенность самого движущегося объекта могут меняться от кадра к кадру, но это допущение достаточно хорошо работает на практике, особенно если речь идет о видеопотоке [2].

На математическом языке это предположение можно выразить следующим образом:

$$I(x, y, t) = I(x + u_x, y + u_y, t + 1),$$

где I - это функция яркости пикселей от положения на кадре и времени. Таким образом, x и y - это координаты пикселя в плоскости кадра, вектор u - смещение, а t - номер кадра в последовательности. Условимся, что между двумя соседними кадрами проходит единичный отрезок времени.

Для знакомства с этим методом сначала рассмотрим одномерный случай. Например, возьмем два одномерных кадра, каждый кадр имеет 1 пиксель в высоту и 20 пикселей в ширину (рис. 4). На втором кадре изображение немного смещено вправо. Именно это смещение мы и хотим найти.



Рис. 4. Два одномерных кадра

Для более удобной работы с этими кадрами мы можем представить их в виде двух функций. В качестве параметра функции будет выступать положение пикселя в кадре, а в качестве выходного значения - его яркость (или интенсивность).

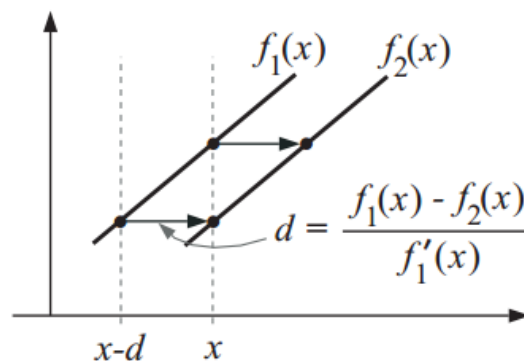


Рис. 5. Два одномерных кадра в виде функций

На таком графике очень хорошо видно искомое смещение (на рис. 5 обозначено как d). Если мы внимательно посмотрим на наши графики и исходные кадры, то видно, что

функция f_2 – это просто смещенная функция f_1 . Выражая предыдущее предложение математическим языком, получим

$$f_2(x) = f_1(x - d).$$

Также мы можем записать функции f_1 и f_2 в общем виде:

$$\begin{aligned} f_1(x) &= I(x, y, t), \\ f_2(x) &= I(x, y, t + 1). \end{aligned}$$

В данном одномерном случае x и t равны нулю.

Для любой произвольной координаты x , нам известно значение обеих функций, а также мы можем вычислить производные этих функций в этих координатах.

Для выражения неизвестной величины d через известные значения функций f_1 и f_2 запишем разложение в ряд Тейлора для $f_1(x-d)$:

$$f_1(x - d) = f_1(x) - df_1'(x) + O(d^2 f_1'')$$

На данном этапе нам нужно сделать еще одно предположение. Предположим, что функция $f_1(x-d)$ хорошо аппроксимируется первой производной и мы можем отбросить все младшие члены ряда Тейлора. После того как мы сделали второе предположение, можно переписать полученную формулу

$$f_1(x - d) = f_1(x) - df_1'(x)$$

Теперь следует проанализировать второе предположение. Отбрасывая все младшие члены ряда Тейлора после первой производной, можно сильно потерять в точности вычислений, но как показывает практика, такой подход сильно упрощает метод, однако если понадобится более высокая точность, можно сделать последовательное приближение.

Из полученной формулы легко выразить искомую величину d , сделав одну простую замену. Как уже было отмечено:

$$f_2(x) = f_1(x - d).$$

Можно переписать последнее полученное выражение следующим образом:

$$f_2(x) = f_1(x) - df_1'(x).$$

Отсюда выражаем искомую величину

$$d = \frac{f_1(x) - f_2(x)}{f_1'(x)}.$$

Применяя данный метод, мы также можем найти смещение пикселей между двумя соседними кадрами и для двумерного случая.

Выводы

В данной статье было рассмотрено применение различных вариантов вычисления оптического потока для восстановления трехмерной структуры сцены и получен устойчивый результат для алгоритма Фарнебака для всех точек на сцене. Также была поставлена задача о возможности улучшения этого результата с помощью алгоритма Лукаса-Канаде.

На сегодняшний день существует несколько других подходов к решению проблемы вычисления оптического потока и многие из них также дают неплохие результаты. Например, в последнее время можно заметить повышение интереса к дискретным методам оптимизации блочным методам.

Проблема нахождения оптического потока решается уже не один десяток лет, и существующие методы и реализации продолжают постоянно улучшаться и совершенствоваться. Такая работа продолжается также ввиду того, что от качества и стабильности нахождения оптического потока зависит производительность многих других алгоритмов и систем компьютерного зрения.

Библиографический список

1. **Форсайт, Д.** Компьютерное зрение. Современный подход / Д.Форсайт, Ж. Понж. – М.: Издательский дом «Вильямс». 2004.
2. **Fleet, J.D.** OpticalFlowEstimation / J.D. Fleet, Y. Weiss. 2006.
3. Искусственный интеллект. Системы и модели. – Интернет ресурс: www.rriai.org.ru. 2003.

*Дата поступления
в редакцию 06.05.2016*

M.A. Kalinin, V.Y. Karpychev

METHODS OF OPTICAL FLOW ESTIMATION FOR 3D SCENE ANALYSIS

Nizhny Novgorod state technical university n. a. R. E. Alexeev

Purpose: This article considers the problem of restoring 3D structure of the scene by using the properties of optical flow. We also discuss the application of different methods of optical flow estimation for given problem.

Design/methodology/approach: This article provides an introduction to mathematical basis of differential methods of estimating optical flow.

Research limitations/implications: The present study provides a starting-point for investigation of possibility of improving this result by using of Lucas–Kanade method.

Originality/value: In scope of this study, we have obtained a pretty much stable result for Farneback algorithm for all points on the scene

Key words: computer vision, optical flow, Lucas–Kanade method, Farneback method.