АЛГОРИТМ СЛУЧАЙНОЙ ВЫБОРКИ В ЗАДАЧЕ ПОИСКА ОБРАЗЦА В ИЗОБРАЖЕНИИ

ВАХИТОВ А.Т., ПАВЛЕНКО Д.В.

УДК 004.932.2

Вахитов А.Т., Павленко Д.В. **Алгоритм случайной выборки в задаче поиска образца в изображении.**

Аннотация. Задача поиска образца в изображении в той или иной форме является одной из базовых задач в области компьютерного зрения. Она возникает в задачах определения оптического потока, стереозрения, сопровождения. Один из классических подходов, предложенный Б. Лукасом и Т. Канаде, основывается на оптимизации некоторого функционала качества. В работе предлагается алгоритм поиска образца в изображении (алгоритм случайной выборки), основанный на подходе Лукаса–Канаде и показывающий высокие результаты по производительности.

Ключевые слова: компьютерное зрение, подход Лукаса–Канаде, оптический поток, алгоритм стохастической аппроксимации с пробным одновременным возмущением на входе.

Vakhitov A.T., Pavlenko D.V. **Random sampling algorithm in the image registration problem.**

Abstract. Image registration is one of the basic problems of computer vision. It arises in optical flow estimation, stereo vision, and tracking problems. One of the classical approaches proposed by B. Lucas and T. Kanade is based on optimization of some cost function. In this article image registration algorithm based on Lucas–Kanade approach is proposed (random sampling algorithm). It shows high performance results.

Keywords: computer vision, Lucas–Kanade approach, optical flow, simultaneous perturbation stochastic approximation (SPSA).

1. Введение. Задача поиска образца в изображении — одна из важнейших задач компьютерного зрения. В настоящей статье предложен алгоритм для решения этой задачи, основанный на оптимизации некоторого функционала среднего риска и сходящийся за малое в сравнении с традиционным подходом Лукаса–Канаде [6] количеством его вычислений.

С ростом вычислительных мощностей увеличиваются потребности пользователя. Задачи компьютерного зрения становятся всё более и более актуальными и, в частности, задача поиска образца в изображении. С развитием робототехники возникают задачи быстрого автоматического построения трёхмерных сцен для определения препятствий на пути робота и навигации вообще. С появлением и массовым распространением цифровых видеокамер появляется задача эффективного сжатия видеопотока. От современных систем видеонаблюдения требуется не просто запись материала, но и выделение и анализ дви-

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

жущихся объектов, вычисление расстояний до них, их траекторий и внешних параметров. Распознавание жестов позволяет вывести взаимодействия человека и компьютера на новый уровень [7]. При съёмке дрожащей камерой актуальна задача стабилизации изображения и компенсации дрожания. Для медицинских целей очень важно совмещать несколько снимков (возможно, деформированных, например, из-за дыхания) в один для получения представления о структуре организма [4]. Аналогичная задача возникает при построении панорамы по нескольким изображениям. Применение предупредительных систем в автомобилях помогает предотвратить аварии. На этапе тестирования и усовершенствования находятся и системы автономного управления автомобилями.

Решаемая задача заключается в поиске положения образца в некотором изображении. В общем случае образец содержится в изображении с точностью до деформации и преобразований изменения освещения, однако самый актуальный вариант задачи — поиск смещения образца при условии, что образец в изображении не повёрнут. Для решения задачи в том или ином варианте предложены различные подходы.

Простейший подход — перебор всех возможных смещений образца. Другой — для каждой текущей оценки вектора смещения перебрать все соседние целочисленные смещения, например, в окрестности 3×3, и выбрать лучший вариант следующей оценкой [6]. Подход на основе корреляции фаз описан в [3, 4].

Б. Лукас и Т. Канаде предложили решать задачу поиска образца в виде поиска минимума функционала

$$F(h) = \sum_{p \in \boldsymbol{P}} K(p)(I(p+h) - J(p))^2,$$

где $\boldsymbol{P}$ — множество пикселей образца, I — изображение — отображение, ставящее в соответствие каждому пикселю из $\boldsymbol{P}$ его интенсивность, J — содержащийся в изображении I образец, h — искомое смещение. Если использовать предположение о том, что изображение — гладкая функция от двух переменных, можно приближённо записать

$$I(p+h) \approx I(p) + \nabla I(p) h,$$

$$F(h) = \sum_{p \in \boldsymbol{P}} K(p)\big(I(p) + \nabla I(p) h - J(p)\big)^2.$$

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

Приравнивая к нулю производную $\frac{dF}{dh}$, после преобразований имеем:

$$h \approx \left[\sum_{p\,\in \boldsymbol{P}} K(p)(\nabla I(p))^T(\nabla I(p))\right]^{-1}\left[\sum_{p\,\in \boldsymbol{P}} K(p)\nabla I(p)(I(p)-J(p))\right].$$

Точное равенство имеет место, только если $I(p)$ — линейная функция. На практике, как правило, это не так. Сдвинув образец J на вектор h, мы можем повторить вычисления. Лукас и Канаде в [6] предложили такой алгоритм построения оценок h:

$$h_{n+1} = {H_n}^{-1}\left[\sum_{p\,\in \boldsymbol{P}} K(p)\nabla I(p+h_n)(I(p+h_n)-J(p+h_n))\right],$$

где

$$H_n = \left[\sum_{p\,\in \boldsymbol{P}} K(p)\big(\nabla I(p+h_n)\big)^T\big(\nabla I(p+h_n)\big)\right].$$

Значение h_n может быть дробным. В этом случае Б. Лукас [5] рекомендует интерполировать I.

Алгоритм может быть обобщён на случай, если образец в изображении получен не только смещением, но и произвольным аффинным преобразованием, а также с помощью изменения параметров освещения. В этом случае функционал может быть записан в виде:

$$F(\theta) = \textstyle\sum_{p\,\in \boldsymbol{P}} K(p)(Q(p,\theta)-J(p))^2,$$

где $Q(p,\theta)$ — некоторое преобразование интенсивности пикселя образца, например, с учётом поворота, растяжения или изменения освещения. Например, θ и $Q(p,\theta)$ могут иметь вид:

$$\theta = (A, h_x, h_y, a, b)^T,$$

$$Q(p,\theta) = aI(Ap+h)+b,$$

где a, b задают параметры освещения, матрица A размерности 2×2 и вектор $h = (h_x, h_y)^T$ определяют аффинное преобразование. Стоит

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

отметить, что вычисление $Q(p,\theta)$ в общем виде сравнительно трудоёмко (несколько операций умножения, интерполяция), особенно, если поворот параметризован не с помощью элементов матрицы, а с помощью тригонометрических функций, которые отнимают сравнительно много машинного времени при вычислении на современных процессорах.

В [2] рассмотрено применение различных методов оптимизации функционала $F(\theta)$ в случае $a = 1$, $b = 0$: метод градиентного спуска, Гаусса-Ньютона, Ньютона и Левенберга–Марквардта.

Классическое решение задачи оптимизации функционала с применением псевдоградиентных методов в -мерном пространстве заключается в итеративном движении в сторону, противоположную направлению градиента. В задачах оптимизации, где аналитическая форма оптимизируемого функционала недоступна, значение градиента приходится аппроксимировать конечно-разностным отношением. Для оценки градиента $\frac{dF}{d\theta}$ необходимо производить как минимум $2r$ вычислений функционала качества $F(\theta)$ при использовании формулы

$$\frac{dF(\theta)}{d\theta} \approx \begin{pmatrix} \dfrac{F(\theta + \beta_1 e_1) - F(\theta + \beta_1 e_1)}{2\beta_1} \\ \cdots \\ \dfrac{F(\theta + \beta_r e_r) - F(\theta + \beta_r e_r)}{2\beta_r} \end{pmatrix}.$$

Далее мы покажем, как применение рандомизированных алгоритмов может существенно сократить количество вычислений Q, требуемых для оптимизации. В этой работе предлагается применить рандомизированный алгоритм стохастической оптимизации с пробным одновременным возмущением на входе, описанный в [1], позволяющий сократить количество вычислений Q на итерацию.

2. Постановка задачи. Рассмотрим обобщённую задачу поиска образца в изображении. Имеется образец, описываемый функцией интенсивности $J(p)$, $p \in \boldsymbol{P}$, где $\boldsymbol{P}$ — множество пикселей образца. Задано -мерное пространство параметров Θ, и на множестве $\boldsymbol{P} \times \Theta$ задано параметрическое преобразование $Q(p,\theta)$ пикселя исходного изображения (учитывающее, например, изменение координат и освещения), определяемое r-мерным вектором θ. На $\boldsymbol{P}$ задана весовая функция $K(p) > 0$, определяющая влияние пикселей образца на оценку качества сопоставления, причём

$$\textstyle\sum_{p\in \boldsymbol{P}} K(p) = 1.$$

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

Известно некоторое приближение решения θ.

Требуется найти вектор θ, который минимизирует функционал оценки качества сопоставления:

$$F(\theta) = \sum_{p \in \boldsymbol{P}} K(p)(Q(p,\theta) - J(p))^2.$$

Мерой оценки качества метода решения будем считать количество вычислений Q.

3. Алгоритм случайной выборки. Рассмотрим новый алгоритм решения поставленной задачи, названный алгоритмом случайной выборки и являющийся логичным применением идей стохастической оптимизации. Одна из идей заключается в том, что для решения задачи не нужна точная аппроксимация градиента, а важно как можно быстрее найти решение.

Рассмотрим величину:

$$Y(p,\theta) = \ (Q(p,\theta) \ - \ J(p))^2.$$

Предлагаемый алгоритм использует оценку слагаемого $Y(p,\theta)$ взвешенной суммы в одном пикселе как зашумлённую оценку $F(\theta)$. При этом сами координаты p пикселя рассматриваются как некоторая неконтролируемая помеха в вычислении этой величины. Эта искусственно привносимая помеха может компенсироваться алгоритмами стохастической оптимизации. В частности, используется *рандомизированный алгоритм с пробным одновременным возмущением на входе с двумя измерениями* [1]. Использование этого алгоритма позволяет произвоить только 2 измерения Q на итерацию.

Если рассмотреть случайный двумерный вектор w с реализациями в $\boldsymbol{P}$, такой что

$$\forall p \ \in \boldsymbol{P} \quad P\{w \ = \ p\} = \ K(p),$$

и поставить задачу оптимизации математическое ожидания $EY(w,\theta)$, применение рандомизированного алгоритма стохастической оптимизации с двумя измерениями даёт следующую последовательность оценок:

$$\theta_n = \theta_{n-1} - \ \alpha_n \frac{Y\left(w_n,\theta_{n-1} + \beta_n\Delta_n\right) - Y\left(w_n,\theta_{n-1} - \beta_n\Delta_n\right)}{2\beta_n}\Delta_n,$$

где w_n — реализация случайной величины w, то есть *случайно выбранный пиксель* p с вероятностью $K(p)$; $\Delta_n = (\pm 1, \pm 1, \dots, \pm 1)$ — реализация случайного бернуллиевского вектора, α_n, β_n — некоторые последовательности. Поскольку $EY(w,\, \theta\,) \, = \, F(\,\theta\,)$, построенные оценки можно использовать как оценки решения поставленной задачи.

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

Табл. **Результаты экспериментов на изображениях Tank, Tsukuba и Moon**

Алгоритм	Вычислений Q	Значение F	% сошлось к ответу по 15 запускам
Тестовое изображение Tank			
Лукаса–Канаде	—	—	0
Кифера–Вольфовицаа	—	—	0
Случайной выборки	$1{,}1 \cdot 10^4$	$7{,}6 \cdot 10^6$	46,7
Тестовое изображение Tsukuba			
Лукаса–Канаде	$1{,}4 \cdot 10^6$	$3{,}4 \cdot 10^6$	100
Кифера–Вольфовицаа	$6{,}8 \cdot 10^5$	$6{,}8 \cdot 10^6$	100
Случайной выборки	$4{,}9 \cdot 10^3$	$4{,}7 \cdot 10^6$	80
Тестовое изображение Moon			
Лукаса–Канаде	$5{,}5 \cdot 10^4$	$2{,}3 \cdot 10^5$	100
Кифера–Вольфовицаа	—	—	0
Случайной выборки	$3 \cdot 10^3$	$3{,}8 \cdot 10^5$	53,3

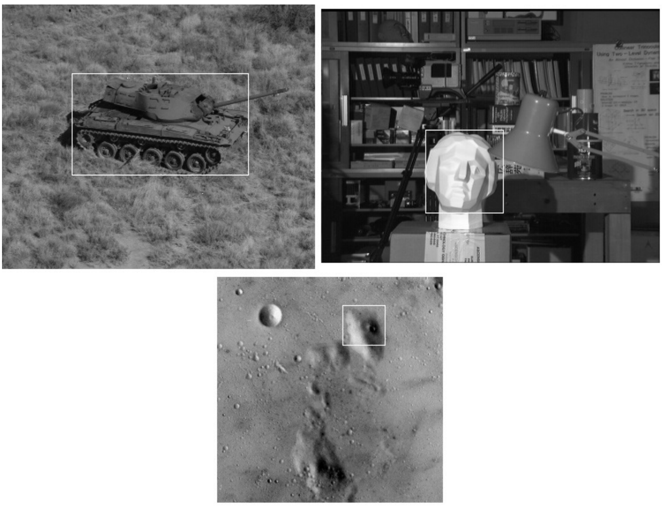

Рис. Тестовые изображения с правильным положением образца в них: Tank, Tsukuba, Moon.

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

4. Экспериментальные результаты. Предложенный алгоритм был опробован на ряде тестовых изображений (рис.). Для сравнения были реализованы алгоритмы Лукаса–Канаде и Кифера–Вольфовица. Результаты представлены в табл. Прочерки означают, что алгоритм не сошёлся к правильному положению образца на изображении. Видно, что предложенный алгоритм значительно превосходит существующие алгоритмы в производительности, несколько уступая им в точности. Кроме того, он далеко не всегда сходится к искомому решению, но, как правило, хотя бы несколько раз из 15 запусков, в отличие от детерминированных алгоритмов.

5. Заключение. В работе был предложен новый алгоритм поиска образца в изображении, основанный на случайной выборке пикселей. Были проведены эксперименты по сравнению типичных поведений алгоритмов Лукаса–Канаде, Кифера–Вольфовица и предложенного в работе, показывающие, что последнему требуется на порядок меньшее количество вычислений значения функции Q, чем алгоритму Лукаса–Канаде, на основе идей которого он и был разработан, что может быть полезно, когда вычисление Q — трудоёмкая операция. Эксперименты также выявили, что по точности предложенный алгоритм уступает алгоритму Лукаса–Канаде и зачастую сходится к нежелательным локальным минимумам, однако лучше ведёт себя в условиях нерегулярных помех.

В дальнейшем планируется исследовать свойства модификации алгоритма, использующей выборку не из одного, а из нескольких пикселей. Ожидается, что такой алгоритм будет обладать положительными свойствами как алгоритма Лукаса–Канаде (большая точность, меньшая вероятность попасть в нежелательный локальный минимум), так и рассмотренного алгоритма случайной выборки (производительность, работа в условиях помех).

Литература

1. *Граничин О. Н.* Рандомизированные алгоритмы стохастической аппроксимации при произвольных помехах // Автоматика и телемеханика. 2002. № 2. С. 44–55.
2. *Baker S., Matthews I.* Lucas–Kanade 20 Years On: A Unifying Framework: Part 1 // International Journal of Computer Vision. 2004. Vol. 56, No. 3. P. 221–255.
3. *Eturk S.* Digital Image Stabilization with Sub-Image Phase Correlation Based Global Motion Estimation // IEEE Transactions on Consumer Electronics. 2003. Vol. 49, No. 4. P. 1320–1325.
4. *Himanshu A., Anoop M. N., Jawahar C. V.* Accurate Image Registration from Local Phase Information // Proceedings of 13th National Conference on Communications. 2007. P. 37–41.
5. *Lucas B.* Generalized Image Matching by the Method of Differences // Doctoral dissertation, Tech. Report. Carnegie Mellon University Pittsburgh. 1985. 167 p.

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru

6. *Lucas B., Kanade T.* An Iterative Image Registration Technique with an Application to Stereo Vision // Proceedings of Imaging Understanding Workshop. 1981. P. 121–130.
7. *K. Nickel, R. Stiefelhagen.* Visual recognition of pointing gestures for human–robot interaction // 2007. vol.25. P.1833–1835.

Граничин Олег Николаевич — д.ф.-м.н., проф.; профессор кафедры системного программирования математико-механического факультета Санкт-Петербургского государственного университета (СПбГУ). Область научных интересов: рандомизированные алгоритмы оптимизации и оценивания, стохастическая оптимизация в информатике, адаптивное и оптимальное управление, распознавание образов. Число научных публикаций — 70. oleg_granichin@mail.ru, www.math.spbu.ru/user/gran; СПбГУ, Университетский пр., д.28, Санкт-Петербург, 198504, РФ; р.т. +7(812)428-4910, факс +7(848)428-7109.

Granichin Oleg Nikolaevich — Doc. Sci., Ph. D., Prof. of department of software engineering of Mathematics and Mechanics Faculty of St. Petersburg State University (SPbSU). Research interests: randomized optimization and estimation algorithms, stochastic optimization in computer science, adaptive and optimcal control, pattern recognition. The number of publications — 70. oleg_granichin@mail.ru, www.math.spbu.ru/user/gran; SPbSU, Universitetsky prospekt, 28, Saint-Petersburg, 198504, Russia; office phone +7(812)428-4910, fax +7(848)428-7109.

Вахитов Александр Тимурович — аспирант математико-механического факультета Санкт-Петербургского государственного университета (СПбГУ). Область научных интересов: стохастическая оптимизация, распределенные вычисления, обработка изображений. Число научных публикаций — 10. av38@yandex.ru; СПбГУ, Университетский пр., д.28, Санкт-Петербург, 198504, РФ; office phone +7(812)428-7109. Научный руководитель — О.Н. Граничин.

Vakhitov Alexander Timurovich — graduate student of Mathematics and Mechanics Faculty of St. Petersburg State University (SPbSU). Research interests: stochastic optimization, distributed computing, image processing. The number of publications — 10. av38@yandex.ru; SPbSU, Universitetsky prospekt, 28, Saint-Petersburg, 198504, Russia; office phone +7(812)428-7109. Scientific advisor — O.N. Granichin.

Павленко Дмитрий Валентинович — студент математико-механического факультета Санкт-Петербургского государственного университета (СПбГУ). Область научных интересов: компьютерное зрение. Число научных публикаций — 2. dmit10@mail.ru; СПбГУ, Университетский пр., д.28, Санкт-Петербург, 198504, РФ; р.т. +7(812)428-7109. Научный руководитель — О.Н. Граничин

Pavlenko Dmitry Valentinovich — student of Mathematics and Mechanics Faculty of St. Petersburg State University (SPbSU). Research interests: computer vision. The number of publications — 2. dmit10@mail.ru; SPbSU, Universitetsky prospekt, 28, Saint-Petersburg, 198504, Russia; office phone +7(812)428-7109. Scientific advisor — O.N. Granichin.

Труды СПИИРАН. 2009. Вып. 9. ISSN 2078-9181 (печ.), ISSN 2078-9599 (онлайн)
SPIIRAS Proceedings. 2009. Issue 9. ISSN 2078-9181 (print), ISSN 2078-9599 (online)
www.proceedings.spiiras.nw.ru