

А. Ю. Чуриков

МЕТОД АДАПТИВНОЙ БИНАРИЗАЦИИ В РЕАЛЬНОМ ВРЕМЕНИ ИЗОБРАЖЕНИЙ СО СЛАБОЙ КОНТРАСТНОСТЬЮ

Рассмотрена задача бинаризации изображений со слабой контрастностью методом адаптивной бинаризации и изучены особенности эффективной реализации данного метода с использованием подхода вычислений на графических картах.

A problem of low contrast image binarization by the adaptive binarization method and an effective realization of this method based on GPGPU computing paradigm is described.

Ключевые слова: обработка изображений, компьютерное зрение, вычисления на графических платах, отслеживание объектов.

Key words: image processing, computer vision, GPGPU, objects tracking.

Бинаризация изображения является одной из самых распространенных задач компьютерной графики. С ней сталкиваются при разработке алгоритмов компьютерного зрения для робототехники, в системах слежения за перемещением объектов в пространстве, при разработке систем анализа изображений, получаемых с различного рода медицинского оборудования, например при триплексном сканировании или МРТ. Последние две области применения представляют особый интерес: изначально метод адаптивной бинаризации использовался для отслеживания и анализа объектов в реальном времени в рамках патентованной технологии гибридной мультисенсорной панели [1], а в данный момент он находит применение в проекте медицинского оборудования для ортопедии — сканере стопы.

Существует огромное количество различных реализаций бинаризации изображений, которые описаны, например, в работах [2–4]. В данной же статье на примере проекта мультисенсорной панели анализируются основные существующие методы бинаризации, дается оценка их эффективности в рамках решаемой задачи и описываются основные моменты, касающиеся эффективной реализации данных методов при вычислениях на графических платах.

Бинаризация изображения заключается в отнесении каждого пикселя изображения в один из двух непересекающихся классов. Вопрос о принадлежности пикселя определенному классу решается с учетом специфики решаемой задачи, но в любом случае учитывает характеристики как этого пикселя, так и его соседей. В рассматриваемых в данной статье задачах требуется выделить область на изображении, которая по некоторым признакам отличается от других, причем данные признаки не всегда четко выражены.



Рис. 1. Изображение
пальцев руки, получаемое
с мультисенсорной панели

Например, показанное на рисунке 1 изображение пальцев руки, получаемое с мультисенсорной панели [1], не обладает высоким уровнем контраста по сравнению с окружающей областью, что привносит дополнительные трудности при отслеживании пальцев.

Формализуем задачу бинаризации. Введем понятие изображения.

Изображение — нормализованная функция $f: D \rightarrow [0, 1]$, где $D \subset M^2$, $M \subset \mathbb{N}$ — область определения изображения.

Требуется найти функцию $g(x, O(x), f, p) \rightarrow [0, 1]$, где $x \in D$; $O(x) \in D$ — некоторая окрестность x ; f — изображение; p — некоторый параметр или группа параметров.

Один из самых простых методов бинаризации — так называемая *глобальная пороговая бинаризация* — задается соотношением

$$g(x, x, f, p) = \begin{cases} 0 & \text{при } f(x) < p, \\ 1 & \text{в противном случае,} \end{cases} \quad p \in [0, 1].$$

Очевидно, что такой метод учитывает лишь характеристики текущей точки и использует глобальный для всего изображения параметр, что делает его непригодным для применения на практике. Этот факт проиллюстрирован на рисунке 2. Из-за неравномерной освещенности исходного изображения (крайнее слева) данный метод показывает неудовлетворительный результат (по центру $p = 0,5$). Ожидаемый же результат (справа) дает метод адаптивной бинаризации.



Рис. 2. Неприменимость глобальной пороговой бинаризации на практике

Основное отличие адаптивной бинаризации от предложенной выше глобальной пороговой состоит в том, что порог отсечения в первом случае определяется индивидуально для каждого пикселя с учетом его окрестности, то есть изначально заданный порог отсечения «адаптируется» для каждого отдельно взятого пикселя (отсюда и название метода):

$$g(x, O(x), f, p) = \begin{cases} 0 & \text{при } f(x) < p \cdot h(O(x), f), \\ 1 & \text{в противном случае,} \end{cases} \quad p \in [0, 1], h: (D, f) \rightarrow \mathbb{R}. \quad (1)$$

Ясно, что глобальная пороговая бинаризация является частным случаем адаптивной бинаризации. Из выражения (1) видно, что качество работы метода зависит от выбора области $O(x)$ и функции h .

В качестве h обычно выбирают функцию вида $h(O(x), f) = \frac{\sum_{y \in O(x)} f(y)}{\|O(x)\|}$. Данная функция показывает, насколько область вокруг выбранного пикселя отличается от его цвета и корректирует заранее заданный порог отсечения.

Этот метод при удачно выбранной окрестности $O(x)$ и параметре p дает очень хорошие результаты, которые можно использовать на практике. На рисунке 3 показан пример использования этой функции в технологии мультисенсорного дисплея [1].



Рис. 3. Использование $h(O(x), f) = \frac{\sum_{y \in O(x)} f(y)}{\|O(x)\|}$ в технологии мультисенсорного дисплея

На практике цифровое изображение представляется в виде двумерного массива f . Не умаляя общности, считаем, что изображение нецветное и каждый пиксель может принимать одно из 256

градаций серого. Введем понятие *интегральной карты изображения* как двумерной префиксной суммы для данного изображения:

$$I(x, y) = \begin{cases} x < 0 \vee y < 0, & 0, \\ \text{иначе} & f(x, y) + f(x-1, y) + f(x, y-1) - I(x-1, y-1), \end{cases}$$

$$x \in [0, w-1], y \in [0, h-1].$$

где w, h — ширина и высота изображения в пикселях соответственно.

Очевидно, что сложность нахождения интегральной карты составляет $\theta(wh)$ операций. Используя заранее вычисленную префиксную сумму можно за $O(1)$ найти значение функции

$$\text{sum}(x_1, x_2, y_1, y_2) = \sum_{i=x_1}^{x_2} \sum_{j=y_1}^{y_2} f(i, j) =$$

$$= I(x_2, y_2) - I(x_2, y_1 - 1) - I(x_1 - 1, y_2) + I(x_1 - 1, y_1 - 1).$$

Следовательно, выбрав согласно выражению (1) в качестве окрестности x квадрат со стороной в n пикселей (n — нечетное) с центром в x и полагая, что

$$g = \frac{\text{sum}(x - \lfloor n/2 \rfloor, x + \lfloor n/2 \rfloor, y - \lfloor n/2 \rfloor, y + \lfloor n/2 \rfloor)}{\| [x - \lfloor n/2 \rfloor, x + \lfloor n/2 \rfloor] \times [y - \lfloor n/2 \rfloor, y + \lfloor n/2 \rfloor] \cap [0, w-1] \times [0, h-1] \|},$$

можно достичь желаемых результатов. Таким образом, при грамотной реализации функции g и найденной интегральной карте сложность алгоритма адаптивной бинаризации составит те же $\theta(wh)$ операций.

Принцип *SIMD* (Single Instruction, Multiple Data) вычислений, используемый в графических платах, является идеальным для алгоритма адаптивной бинаризации. Действительно, при наличии готовой интегральной карты процесс бинаризации каждого пикселя проходит независимо от других пикселей — наблюдается классический случай распараллеливания на уровне данных. Главную сложность в данном случае представляет процесс нахождения интегральной карты.

Рассмотрим алгоритм нахождения одномерной интегральной карты (префиксной суммы), адаптированный под принцип SIMD вычислений, подробно описанный, например, в работе [5]. На первом этапе происходит суммирование элементов, расстояние между которыми равно 1, на каждом этапе это расстояние увеличивается в два раза до тех пор, пока шаг не станет равным или превысит размер массива. В статье [5] приводится оценка времени работы данного алгоритма $O(\log n)$, конечно, при учете того, что количество свободных ядер сопоставимо с размерами массива (количество операций суммирования равно $n \log n$).

Расширим данный алгоритм на двумерный случай. Рассмотрим матрицу $w \times h$. Согласно нашему алгоритму найдем префиксную сумму для каждой строки данной матрицы и запишем результат в виде новой матрицы. Данная задача выполнится за $O(h \log w)$ шагов. Повторяя процедуру с полученной матрицей для столбцов, получаем искомую интегральную карту. Поэтому в рамках изложенных выше предположений процесс вычисления префиксной суммы займет $O(h \log w + w \log h)$ времени, что на порядок меньше, чем при использовании традиционных методов вычислений префиксной суммы.

Список литературы

1. Исманов А. Т., Чуриков А. Ю., Валов С. В. Мультисенсорное дисплейное устройство: патент на полезную модель RU95142U1. 2010.
2. Bernsen J. Dynamic thresholding of gray-level images // Int. Conf. Pattern Recognition. 1986. Vol. 2. P. 1251 — 1255.
3. Chan F. H. Y., Lam F. K., Zhu H. Adaptive thresholding by variational method // IEEE Transactions on Image Processing 7. 1998. 3 (March). P. 468 — 473.
4. Shapiro L. G., Stockman G. C. Computer Vision. Prentice Hall, 2002.
5. Harris M. Parallel Prefix Sum (Scan) with CUDA // nVidia CUDA Reference manual. 2009.

Об авторе

Анатолий Юрьевич Чуриков — студ., РГУ им. И. Канта, e-mail: achurikov89@ gmail.com.