

## МЕТОД ОЦЕНКИ РАЗМЕРОВ БАЗОВЫХ ЭЛЕМЕНТОВ СТРУКТУР ОТСКАНИРОВАННЫХ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

В.А. Митекин<sup>1,2</sup>, В.А. Федосеев<sup>1,2</sup>

<sup>1</sup>Институт систем обработки изображений Российской академии наук, Самара, Россия,

<sup>2</sup>Самарский государственный аэрокосмический университет Самара, Россия

### Аннотация

В статье рассматривается алгоритм оценки размеров базовых элементов структур отсканированных цифровых изображений, основанный на подсчете длин серий и вычислении локального поля направлений. Приводятся результаты эксперимента, исследуется эффективность алгоритма.

### Введение

В ряде задач анализа изображений документов [1] используются образцы, полученные путем сканирования документов, напечатанных полиграфическим способом. Очевидно, такие образцы обладают рядом искажений, вызванных процессом печати-сканирования: зашумленность, геометрические искажения, размытие, погрешности дискретизации и квантования и др.

Как показано в [2], влияние этих искажений может быть уменьшено путем использования сканирующих устройств заведомо большего разрешения, нежели разрешение полиграфического устройства, которым были напечатаны используемые документы. Однако в этом случае неизбежно масштабирование базового элемента структуры анализируемых изображений, вследствие которого некоторые процедуры их обработки, в частности, обработка скользящим окном заданного размера, будут давать не вполне корректные результаты.

Для решения данной проблемы необходимо предварительно оценить коэффициент масштабирования базового элемента структуры в результате цикла печати-сканирования. Если известна линиятура (плотность полиграфического раstra) печатающего устройства, то решение этой задачи не представляет трудностей, однако в ряде случаев линиятура не известна, и задача требует специальных методов компьютерного анализа, один из которых и предлагается в настоящей работе.

Для удобства терминологии введем понятие *цифровой линиятуры* как коэффициента увеличения геометрических примитивов в результате процесса печати-сканирования. Заметим, что численно цифровая линиятура будет равна размеру минимальных базовых элементов изображения в пикселях. В дальнейшем мы будем понимать под линиятурой не физическую, а именно цифровую линиятуру. Пользуясь данной терминологией, можно сформулировать поставленную задачу как оценку линиятуры отсканированного цифрового изображения.

Существующие методы оценки размеров базовых элементов изображения основаны, большей частью, на концепции *обработки изображений с мультиразрешением* (Scale Space Image Processing). Наиболее детально разработанный алгоритм, основанный на данном подходе, описан в [3]. Однако, несмотря на универсальность и отсутствие серьез-

ных ограничений, накладываемых на тип анализируемого изображения, практическое применение данного алгоритма ограничено из-за высокой вычислительной сложности. Кроме того, данный алгоритм является итерационным, и время его работы линейно зависит от размера базовых элементов структуры анализируемого изображения.

Целью данной работы является разработка прямого однопроходного алгоритма, позволяющего оценивать величину линиятуры.

### Общая идея алгоритма

Большинство сканеров на выходе дают полноцветное RGB-изображение. Поскольку для решения задачи оценки линиятуры не представляет интереса не только цвет, но и яркость точек изображения, на предварительном этапе необходимо преобразовать изображение к бинарному.

Предлагаемый метод основан на подсчете статистики длин серий одинаковых (в бинарном случае – черных) пикселей при горизонтальном и вертикальном проходе изображения. В дальнейшем мы исходим из предположения, что максимуму этой статистики (наиболее часто встречающейся длине серии) будет соответствовать значение линиятуры, то есть фактически размера базового элемента структуры в пикселях. Однако данный подход влечет за собой некоторые трудности, вызванные, к примеру, попаданием в статистику «посторонних» серий, соответствующих краевым участкам объектов на изображении.

Рисунок 1 иллюстрирует описанную выше проблему. Светлым цветом выделены горизонтальные серии черных точек, положение которых дает основание полагать о близости их длин с искомой линиятурой.



Рис. 1. Горизонтальные серии, которые должны попадать в статистику

По приведенному на рисунке фрагменту можно сделать вывод о необходимости введения дополнительного условия, характеризующего пространственное положение серии относительно базовых элементов изображения, при соблюдении которого следует учитывать серию в статистике. В качестве такого условия можно использовать условие близости локальных полей направлений в точках начала и конца серии.

Под *локальным полем направлений* в точке будем понимать направление на область наибольшего скопления черных пикселей в некоторой окрестности данной точки. На рис. 2 стрелками обозначены направления для некоторых серий изображения, представленного на рис. 1 (большими стрелками обозначены близкие направления, а маленькими – существенно различающиеся).



Рис. 2. Локальные поля направлений на концах горизонтальных серий

Как видно из рисунка, дополнительное условие является достаточным для того, чтобы в статистику не попадали «лишние» серии. Таким образом, предлагаемый метод основывается на подсчете статистики длин серий с учетом вычисленных локальных полей направлений. В последующих разделах приводится детальное описание алгоритма.

**Метод оценки направления в окне**

Известно множество методов построения поля направлений [4, 5]. Однако для оценки линиатуры необходимо посчитать не все поле направлений для изображения, а лишь локальное поле направлений в точках начала и конца серий, в окрестности которых, как правило, находится примерно одинаковое количество черных и белых пикселей. Поэтому в данном случае оно определяется как направление из центра окна на центр масс черных точек в локальном окне. Размер окна выбирается равным

$$n = \min \{ L_{white}, L_{black} \},$$

где  $L_{black}$  – длина текущей черной серии, а  $L_{white}$  – длина ближайшей белой серии. На  $n$  накладываются дополнительное ограничение сверху и снизу, чтобы избежать с одной стороны излишне грубых оценок,

а с другой – лишних вычислений. В данной работе использовалось ограничение  $n \in [7, 30]$ .

**Оценка влияния погрешностей на вычисление направления**

В предыдущих разделах говорилось о том, что серия учитывается в статистике в случае близости направлений в начале и конце серии. Очевидно, критерий близости направлений будет определяться погрешностью оценки локального поля направлений  $\varepsilon_{ij}$  в точке с координатами  $(i, j)$  (где  $i$  – номер строки, а  $j$  – номер столбца), которая зависит от размера окна  $n_{ij}$  и распределения шума  $\rho$  в этом окне:

$$\varepsilon_{ij} = \varepsilon(i, j, n_{ij}, \rho(i, j, n_{ij})). \tag{1}$$

Тогда критерий близости направлений в точках  $(i, j)$  – начала серии – и  $(i, j + L_{black})$  – конца серии (при горизонтальном проходе изображения) – выглядит следующим образом:

$$|d_{ij} - d_{ij+L_{black}}| < \varepsilon_{ij} + \varepsilon_{ij+L_{black}}, \tag{2}$$

где  $d_{ij}$  – локальное поле направлений в точке  $(i, j)$ .

Определим приближенно зависимость (1) для модели печати-сканирования. Предположим, что исходное изображение в результате этого процесса было подвергнуто зашумлению. Далее в расчетах будем исходить из модели импульсного шума, а также будем считать, что

$$\varepsilon_{ij} = \varepsilon(n_{ij}) = \varepsilon_n,$$

т.е., исходя из предположения о слабой зависимости погрешности от положения центра окна, заменим значение интенсивности шума в окне его средней интенсивностью на всем изображении.

При большом размере окна импульсный шум не искажает поле направлений, так как он равномерно распределен по всему окну. Оценим возможность неравномерного распределения пикселей шума внутри окна. Пусть  $\lambda$  – базовая вероятность в модели импульсных помех. Далее будем рассматривать распределение величины  $1/L_k$ , где  $L_k$  – расстояние между  $k$ -ым и  $(k+1)$ -ым шумовыми пикселями в окне.

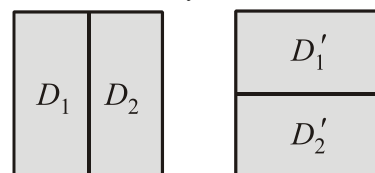


Рис. 3. Взаимное расположение областей в окне

Рассмотрим выборки данной величины для областей  $D_1, D_2, D_1', D_2'$  внутри окна (рис. 3). Объем выборки примем равным  $\frac{1}{2}n^2\lambda$ , где  $n$  – размер окна. Так как характеристики импульсного шума одинаковы для указанных областей, то должна выполняться статистическая гипотеза о равенстве матема-

тических ожиданий  $M_{D1} = M_{D2} \wedge M_{D1'} = M_{D2'}$ , при заданном  $\alpha$ , что равнозначно выполнению гипотезы  $M_{D1} = M_{D2}$  при  $\alpha' = \sqrt{\alpha}$ .

Необходимо определить максимальное  $\Delta M_{noise} = |M_{D1} - M_{D2}|$ , которое не вызовет отклонения описанной выше статистической гипотезы. Оценка ожидаемого количества «шумовых» пикселей может быть определена как

$$L_{noise} = \frac{1}{2} n^2 \cdot \Delta M_{noise}.$$

На основании вычисленного значения  $L_{noise}$  становится возможным оценить ошибку, возникающую при вычислении центра масс. Полагая, что связь между  $L_{noise}$  и ошибкой при вычислении центра масс является линейной, получим относительную погрешность  $\varepsilon'_n = L_{noise} / N_{black}$ , где  $N_{black}$  – число черных пикселей в окне, которое может либо вычисляться для каждого конкретного окна, либо приниматься равным среднему значению  $N_{black} = \frac{1}{2} n^2$ . В этом случае

$$\varepsilon'_n = \frac{L_{noise}}{N_{black}} = \frac{2L_{noise}}{n^2}, \quad (3)$$

$$L_{noise} = \frac{1}{2} n^2 \cdot \Delta M_{noise} = \frac{\zeta n^2}{2n\sqrt{\lambda}},$$

где  $\Delta M_{noise}$  – допустимая разность математических ожиданий для областей  $D_1, D_2$ .

Упрощенная оценка по критерию Уэлча  $\Delta M_{noise} = \frac{\zeta}{n\sqrt{\lambda}}$ , где  $\zeta$  – некая константа, не зависящая от объема выборки.

Поэтому

$$\varepsilon_n = \frac{\pi}{2} \cdot \varepsilon'_n = \frac{\pi \zeta}{2n\sqrt{\lambda}}$$

– погрешность для углов, измеряемых в радианах. Правая часть равенства (3) содержит неизвестную константу, определяемую эмпирически и зависящую от характеристик печатающего и сканирующих устройств. В частности, на основе эмпирической оценки количества «шумовых» пикселей для окна конкретного заданного размера было выбрано значение  $\zeta/\sqrt{\lambda} \approx 0,25$ , и получена приблизительная оценка

$$\varepsilon_n \approx \frac{\pi}{8n} \quad (4)$$

Сделанные упрощения не позволяют дать точную оценку абсолютной погрешности (фактически, точная оценка заменена значением  $\zeta$ , выбираемым в зависимости от предполагаемой интенсивности импульсного шума), но позволяют учесть изменение абсолютной погрешности при изменении размера окна. Зависимость абсолютной погрешности от количества черных пикселей в окне не учтена в данной

приближенной оценке, но может учитываться путем вычисления для конкретного окна значения  $N_{black}$ .

#### Учет направления серий при оценке линиатуры

Серии, учитываемые в статистике, могут пересекать объекты на изображении под разными углами в зависимости от их ориентации и от поворота документа при сканировании. Так как изображение обрабатывается дважды – в вертикальном и горизонтальном направлении, то в худшем случае длины серий могут увеличиться в  $\max_{\alpha \in [0; \pi/2]} \sin \alpha = \sqrt{2}$  раз.

На рис. 4 темными линиями показаны серии при двух проходах изображения, а светлой линией – реальная линиатура.



Рис. 4. Линиатура и длины серий

Для получения минимального значения линиатуры необходимо корректировать длины серий, участвующих в статистике, с учетом угла между преобладающим направлением в точках серии и направлением развертки. Преобладающее направление серии определяется как среднее арифметическое направлений на ее концах.

При проведении экспериментов использовались два способа коррекции длин серий:

- для каждой серии,
- для групп серий одной длины.

Способ (а) методологически является более корректным, однако, как показали эксперименты, не является устойчивым к ошибкам оценки направлений. Мотивацией для использования способа (б) является априорная информация о наличии преобладающей ориентации объектов на большинстве изображений исследуемого класса (рис. 5), вследствие чего можно установить приближенное соответствие между длинами серий и их ориентацией. Этот способ ввиду усреднения направлений дает результат, более стойкий к погрешностям их вычисления.



Рис. 5. Примеры изображений с ярко выраженным преобладающим направлением объектов

**Описание алгоритма**

**Шаг 1.** Преобразование цветового пространства RGB–Lab.

**Шаг 2.** Обработка L-компоненты:

**Шаг 1.2.** Адаптивная пороговая обработка.

**Шаг 1.3.** Поэлементная обработка изображения (цикл по всем пикселям):

**Шаг 1.3.1.** Выделение серии черных пикселей в строке.

**Шаг 1.3.2.** Нахождение локальных полей направлений в точках начала и конца серии.

**Шаг 1.3.3.** При условии близости полей направлений добавление длины текущей серии в статистику, а ее направления – в отдельную статистику.

**Шаг 2.3.** Нахождение длины серии с максимальным значением статистики.

**Шаг 2.4.** Транспонирование изображения.

**Шаг 2.5.** Повторение пунктов 2.2-2.3 для транспонированного изображения (новая статистика).

**Шаг 3.** Нахождение минимальной из двух найденных длин серий.

**Шаг 4.** Линиатура равна полученной длине серии, помноженной на косинус соответствующего ей направления.

**Вычислительный эксперимент**

При проведении эксперимента использовались отсканированные изображения документов, напечатанных различными устройствами печати.

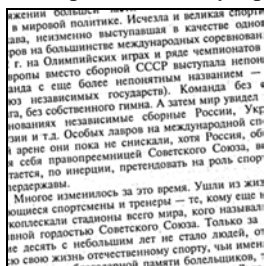


Рис. 6. Изображение «Текст»



Рис. 7. Изображение «Портрет»

В частности, на рис. 8-9 приведены гистограммы длин горизонтальных серий для изображений на рис. 6-7. Следует отметить, что для изображений на рис. 7, которое было напечатано и отсканировано с высоким разрешением, всплеск статистики пришелся

ся не только на линиатуру, но и на длины серий, кратные линиатуре. Однако для большинства проанализированных изображений гистограмма имеет вид, сходный с гистограммой первого изображения.

**Оценка времени работы алгоритма**

Очевидно, время работы  $T$  предлагаемого алгоритма линейно зависит от общего количества пикселей анализируемого изображения, от количества пикселей, обрабатываемых при подсчете локальных полей направлений, и может быть выражено соотношением:

$$T = a \cdot V \cdot H + b \cdot \sum_{k=1}^K \left[ \left( \min \{ L_{black}^k, L_{white}^{k,0} \} \right)^2 + \left( \min \{ L_{black}^k, L_{white}^{k,1} \} \right)^2 \right], \tag{5}$$

где  $V, H$  – ширина и высота изображения,  $k$  – порядковый номер черной серии,  $K$  – общее количество черных серий,  $L_{black}^k$  – длина  $k$ -ой черной серии,  $L_{white}^{k,0}$  – длина белой серии, предшествующей  $k$ -ой черной серии,  $L_{white}^{k,1}$  – длина белой серии, следующей за  $k$ -ой черной серией,  $a$  и  $b$  – постоянные величины.

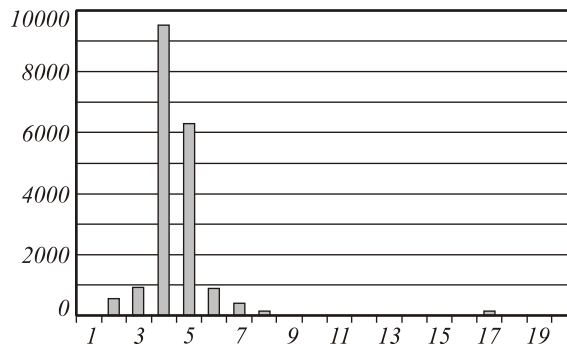


Рис. 8. Гистограмма длин серий для изображения «Текст»

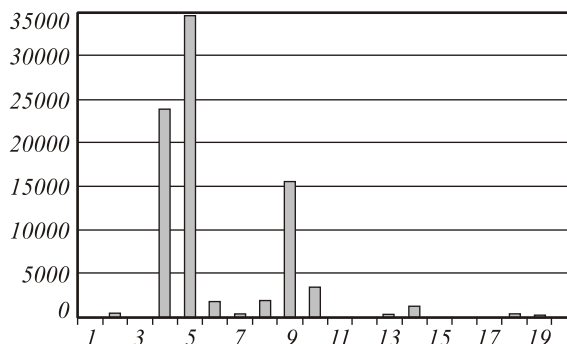


Рис. 9. Гистограмма длин серий для изображения «Портрет»

Очевидно, в силу зависимости размера локального окна при оценке направлений от длин двух серий и невозможности определить общее количество серий и их длины (поскольку в статистику попадают далеко не все), практическая проверка формулы (5) по результатам работы программы невозможна без

дополнительных алгоритмических изменений. Однако приближенно время работы алгоритма может быть оценено по формуле:

$$T \approx \alpha \cdot V \cdot H + \beta \cdot \sum_{n=1}^{n_{\max}} s_n n^2, \quad (6)$$

где  $n$  – текущая длина серии,  $n_{\max}$  – наибольшая анализируемая длина серии,  $s_n$  – количество серий длины  $n$ , попавших в статистику,  $\alpha$  и  $\beta$  – постоянные.

В формуле (6) завышены размеры окна, однако не учитываются не включенные в статистику серии. Как показали сравнения значений времени, полученных по формуле (6), с реальным временем работы программы, эти допущения в некоторой степени «нейтрализуют» друг друга, и оценка получается близкой к экспериментально полученным значениям.

В качестве примера в таблице 1 приведены значения  $V \cdot H$  и  $\sum_{n=1}^{n_{\max}} s_n n^2$ , а также практического  $T_{pr}$  и теоретического  $T_{tr}$  (по формуле (6) времени для четырех изображений. По первым двум из них определены коэффициенты  $\alpha$  и  $\beta$  из условия равенства  $T_{pr}$  и  $T_{tr}$ , которые затем подставлены в формулы для других двух изображений. Полученные теоретические значения оказались довольно близкими к практическим.

Таблица 1. Расчет времени работы программы

№ изображения	$V \cdot H$ , пикселей	$\sum_{n=1}^{n_{\max}} s_n n^2$ , пикселей	$T_{pr}$ , сек.	$T_{tr}$ , сек.
1	5 719 728	4 394 467	19	19
2	1 444 851	668 581	4	4
3	3 247 200	2 932 689	10	11,5
4	9 606 790	2 932 704	21	23,8

### Заключение

В данной работе предложен прямой алгоритм оценки линиатуры отсканированного изображения, базирующийся на подсчете длин серий и вычислении локального поля направлений. Время работы алгоритма линейно зависит от размеров изображения, что существенно меньше времени работы известных алгоритмов решения данной задачи.

### Благодарности

Работа выполнена при финансовой поддержке российско-американской программы «Фундаментальные исследования и высшее образование» ("BRHE"), а также грантов РФФИ № 07-07-97610-р\_офи, № 07-07-97603-р\_офи, № 07-01-96612-р.

### Литература

1. Wang Y., Phillips I.T., Haralick R.M. Document Zone Content Classification and Its Performance Evaluation. Pattern Recognition, Jan. 2006. – 39(1). – Pp. 57-73.
2. Longjiang Yu, Xiamu Niu, Shenghe Sun. Print-and-scan model and the watermarking countermeasure. Image and Vision Computing, September 2005. – Vol. 23. – Issue 9. – Pp. 807-814.
3. Sporing J., Colios Ch. Uncommitted Size-Estimation of Image Structure. Technical Report FORTH-ICS, 1999.
4. Методы компьютерной обработки изображений. Под ред. В.А. Сойфера. – М.: Физматлит, 2001.
5. Pes P.J.W. Average Cell Orientation, Eccentricity and Size Estimated from Tissue Images. Waterloo, – Canada, 2005.

**A METHOD FOR ESTIMATING THE SIZE OF BASE ELEMENTS OF STRUCTURES OF SCANNED DIGITAL IMAGES***V.A. Mitekin<sup>1,2</sup>, V.A. Fedoseev<sup>1,2</sup>**<sup>1</sup>Image Processing Systems Institute of the RAS, Samara, Russia,**<sup>2</sup>Samara State Aerospace University (SSAU), Samara, Russia***Abstract:**

This paper considers an algorithm for estimating the size of base elements of structures of scanned digital images based on calculation of run lengths and on evaluation of a direction local field. Experimental results are given and the algorithm's efficiency is analyzed.

**Keywords:** digital images, base elements of structures, direction local field, run lengths

**Citation:** Mitekin VA, Fedoseev VA. A method for estimating the size of base elements of structures of scanned digital images [In Russian]. *Computer Optics* 2007; 31(2): 86-90.

**Acknowledgements:** The work was partly supported by the Russian-American Basic Research and Higher Education Program (BRHE) and the Russian Foundation for Basic Research (RFBR) grants Nos. 07-07-97610-r\_ofi, 07-07-97603-r\_ofi, 07-01-96612-r.

**References:**

- [1] Wang Y, Phillips IT, Haralick RM. Document Zone Content Classification and Its Performance Evaluation. *Pattern Recognition* 2006; 39(1): 57-73.
- [2] Longjiang Yu, Xiamu Niu, Shenghe Sun. Print-and-scan model and the watermarking countermeasure. *Image and Vision Computing* 2005; 23(9): 807-814.
- [3] Sporning J, Colios Ch. Uncommitted Size-Estimation of Image Structure. Technical Report FORTH-ICS, 1999.
- [4] Soifer VA (ed.), Gashnikov MV, Glumov NI, Ilyasova NYu, Myasnikov VV, Popov SB, Sergeev VV, Khramov AG, Chernov AV, Chernov VM, Chicheva MA, Fursov VA. *Methods of computer image processing* [In Russian]. Moscow: "Fizmatlit" Publisher; 2001.
- [5] Iles PJW. Average Cell Orientation, Eccentricity and Size Estimated from Tissue Images. Waterloo, Canada, 2005.