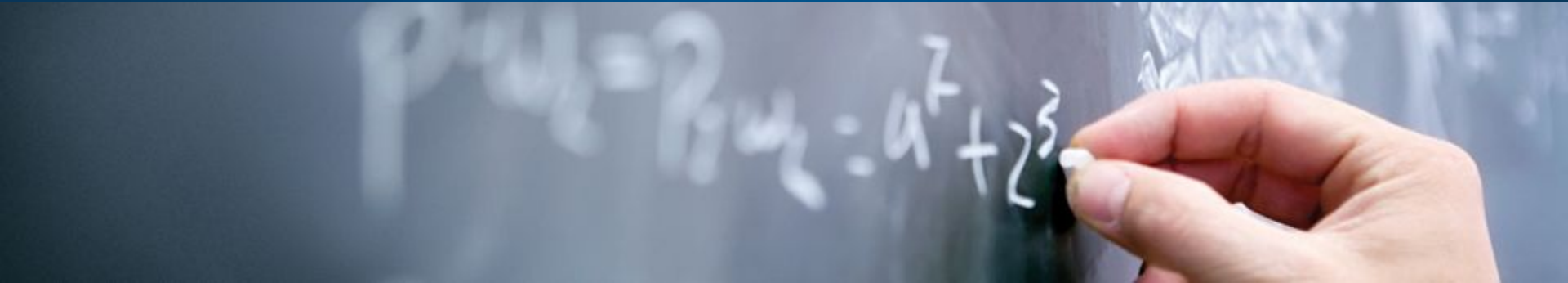


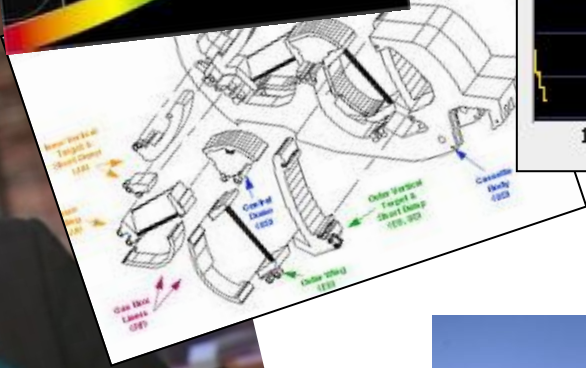
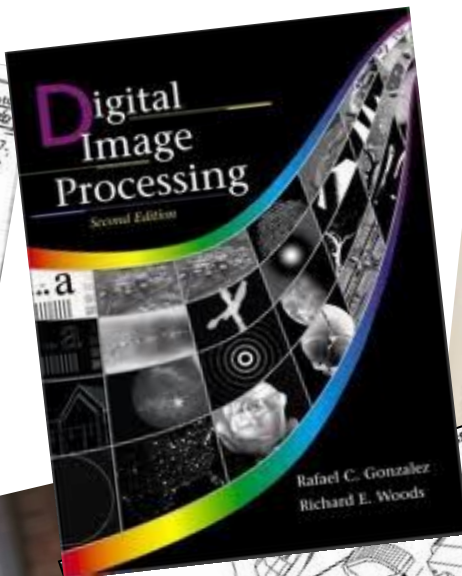
Обнаружение текста на изображениях

Наталья Васильева

nvassilieva@hp.com

HP Labs Russia





Handwritten notes and signatures, including the phrase 'I think we should do this' and 'I would do it but before agreeing to it we must insist upon full access to discs recovered. For instance in the case the Army grabbed it I would not let me bid it for enemy information'.

ANNING JOE
Decision 108
NASTINESS IN POLITICS
AL QAEDA LEADER IN NORTHERN IRAQ

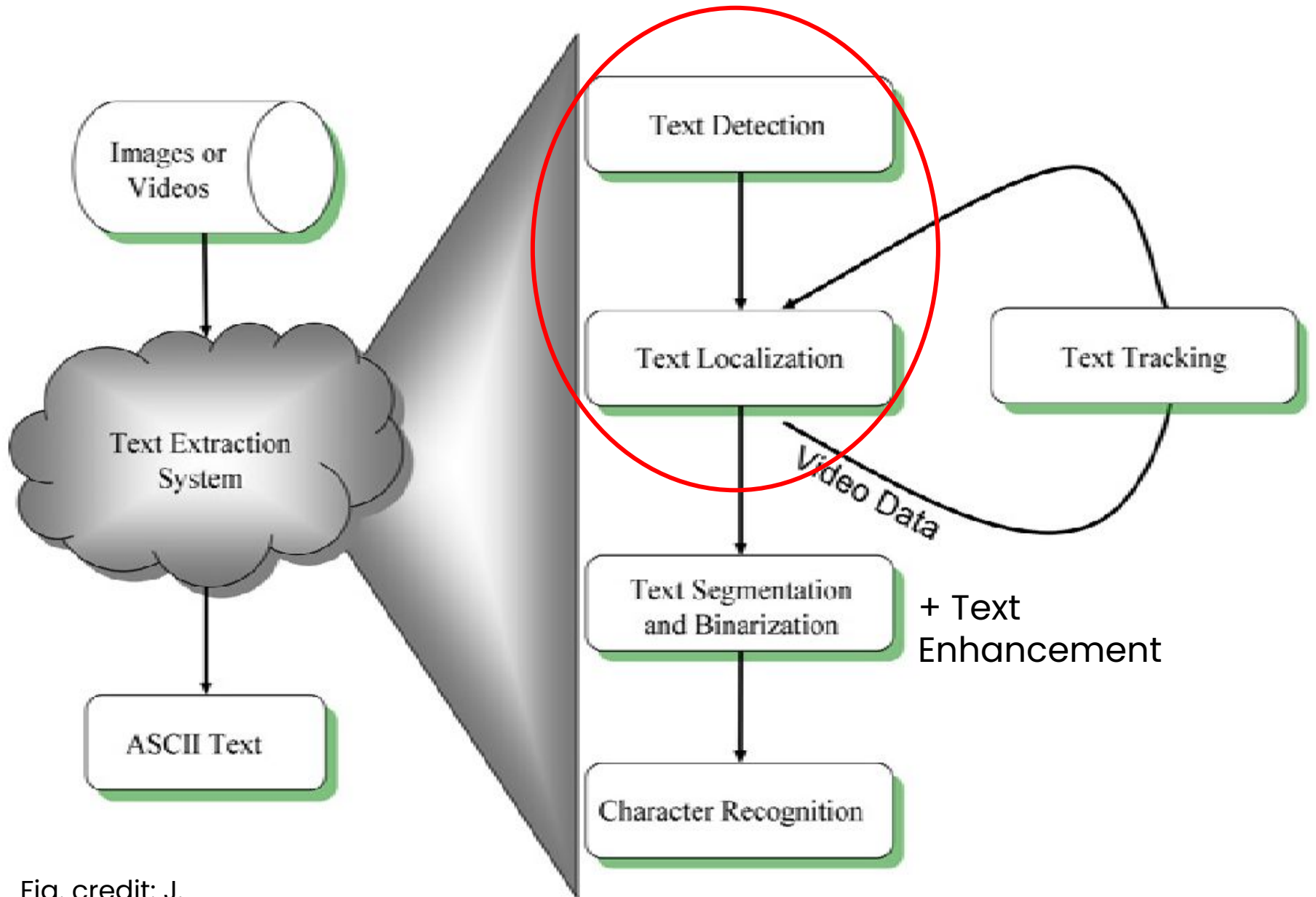


Зачем?

- Необходимо для дальнейшего распознавания текста (OCR)
 - Документы – page layout analysis
 - Фотографии, чертежи, графики – text detection and localization
- Самостоятельные приложения
 - Автоматическое построение коллажей
 - Автоматическое изменение размера изображений



Компоненты системы извлечения текста



Приложения

- Оцифровка документов
- Индексирование и извлечение информации из графиков и чертежей
- Индексирование и поиск изображений, автоматическое построение аннотаций
- Переводчик в кармане
пример: Word Lens (<http://questvisual.com/>)
- Помощь слабовидящим

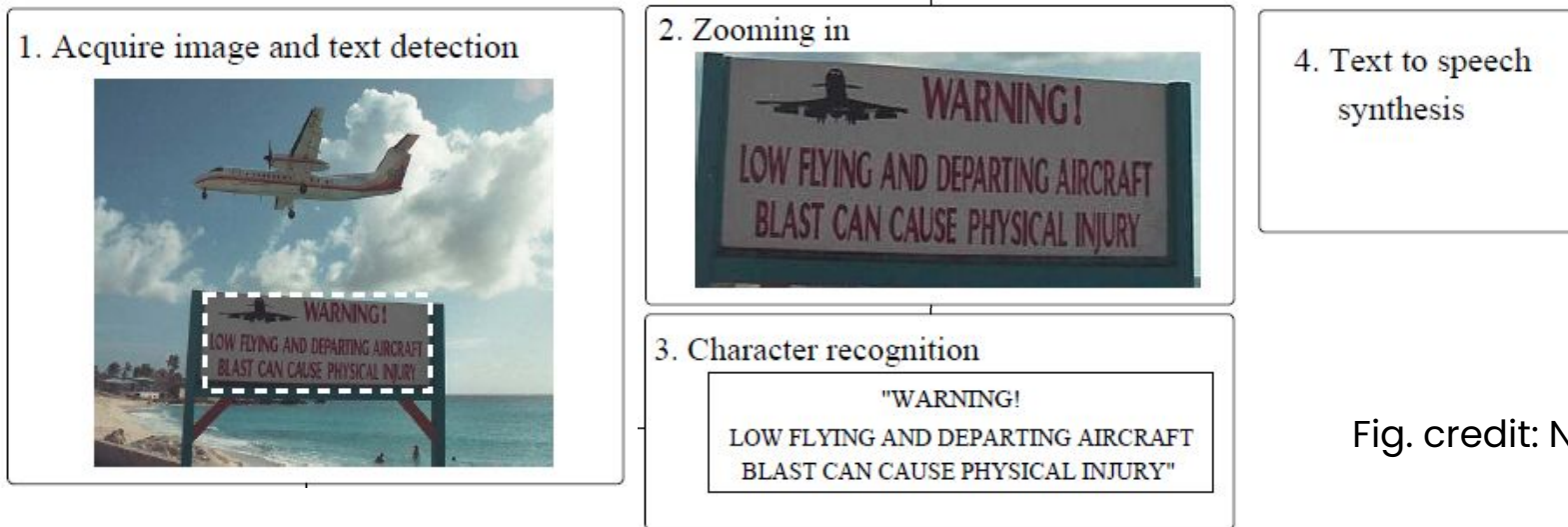


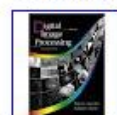
Fig. credit: N. Ezaki et al.

- Навигация роботов в помещениях, в городских условиях

Похожие изображения - Пожаловаться на картинки



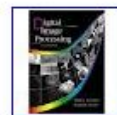
Страницы с подходящими изображениями



300 × 300

[Amazon.com: Digital Image Processing \(2nd Edition ...](http://www.amazon.com/Digital-Image-Processing-2nd-Edition-...)

www.amazon.com/Digital-Image-Processing-2nd-Edition-... - Перевести эту страницу
Digital Image Processing has been the leading textbook in its field for more than 20 years. As was the case with the 1977 and 1987 editions by **Gonzalez** and ...



240 × 240

[Digital Image Processing 2Ed Gonzalez-woods | Free eBooks ...](http://ebookee.org/Digital-Image-Processing-2Ed-...)

ebookee.org/Digital-Image-Processing-2Ed-... - Перевести эту страницу
Jun 6, 2007 – **Digital Image Processing** has been the world-wide leading textbook in its field for almost 30 years. As the 1977 and 1987 editions by **Gonzalez** ...



62 × 80

[Digital Image Processing](http://books.google.com/.../Digital_Image_Processing)

[books.google.com > ... > Image Processing](http://books.google.com/.../Digital_Image_Processing) - Перевести эту страницу

★★★★★ Отзывы: 16
books.google.com - THE leader in the field for more than twenty years, this introduction to basic concepts and methodologies for **digital image processing** ...

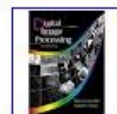


62 × 80

[Digital Image Processing](http://books.google.com/.../Digital_Image_Processing)

books.google.com/.../Digital_Image_Processing - Перевести эту страницу

Digital Image Processing. Front Cover. **Gonzalez** - 1
Reviewhttp://books.google.com/books/about/
/Digital_Image_Processing.html?id=a62xQ2r_f8wC. Pearson ...

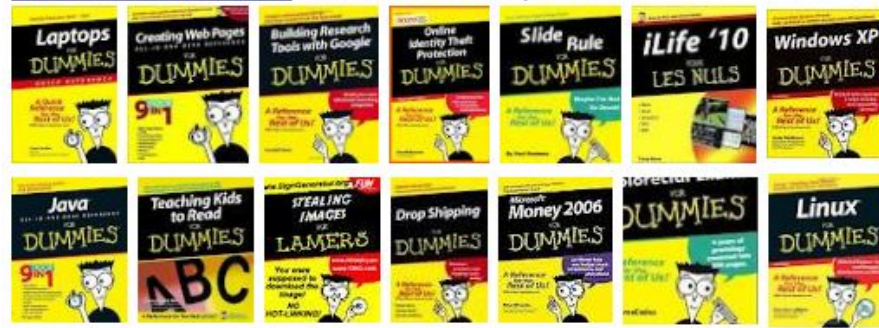


300 × 300

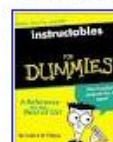
[Digital Image Processing: Amazon.co.uk: Rafael C. Gonzalez ...](http://www.amazon.co.uk/.../Applied-Optics)

[www.amazon.co.uk > ... > Applied Optics](http://www.amazon.co.uk/.../Applied-Optics) - Перевести эту страницу
Product Description. For courses in **Image Processing** and Computer Vision. Completely self-contained—and heavily illustrated—this introduction to basic ...

Похожие изображения - Пожаловаться на картинки



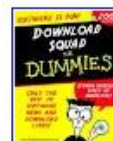
Страницы с подходящими изображениями



400 × 506

[... for dummies book cover](http://www.instructables.com/.../for-dummies-bo-...)

www.instructables.com/.../for-dummies-bo-... - Перевести эту страницу
22 Sep 2009 – This instructables will teach you how to make your very own ... for **dummies book cover**. I did not make this site, nor do I endorse it. I take no cr...



250 × 313

[The 'For Dummies' book cover generator](http://downloadsquad.switched.com/.../the-for-du-...)

downloadsquad.switched.com/.../the-for-du-... - Перевести эту страницу
6 Apr 2006 – Like the magazine cover generator that we linked to previously, the **For Dummies book cover generator** can create a **For Dummies book cover** ...



315 × 400

[For Dummies Book Cover Design Template - Neowin Forums](http://www.neowin.net/.../Jokes-&FunnyStuff)

[www.neowin.net > ... > Jokes & Funny Stuff](http://www.neowin.net/.../Jokes-&FunnyStuff) - Перевести эту страницу
18 Apr 2008 – I found these on another site and thought they were amusing. Make some yourself and post them here. I'd love to see what you can come up...



184 × 233

[Buzz Marketing with Blogs: Buying the Book](http://www.buzzmarketingwithblogs.com/.../buyi-...)

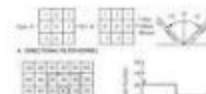
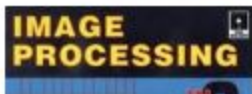
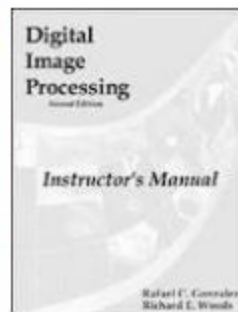
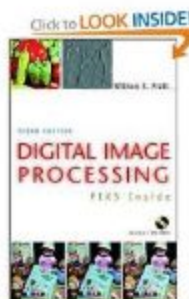
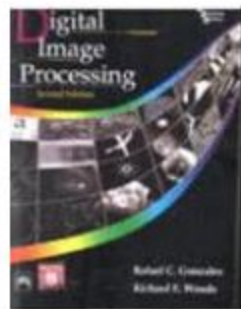
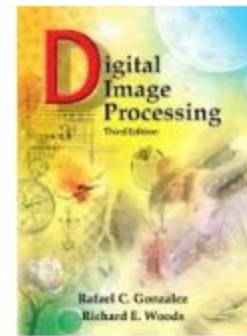
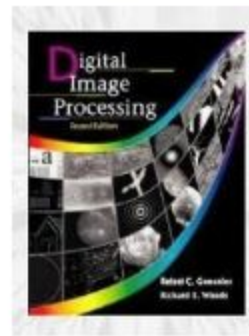
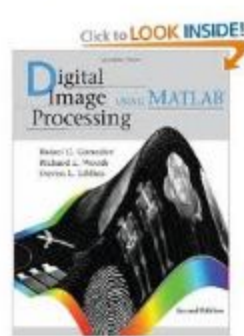
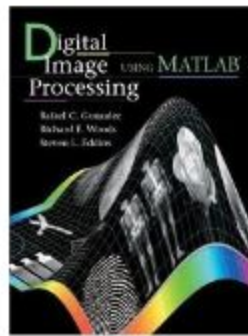
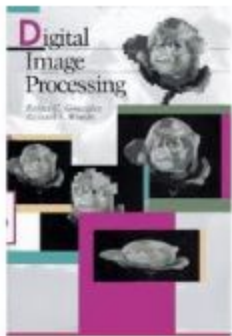
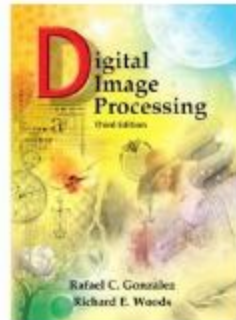
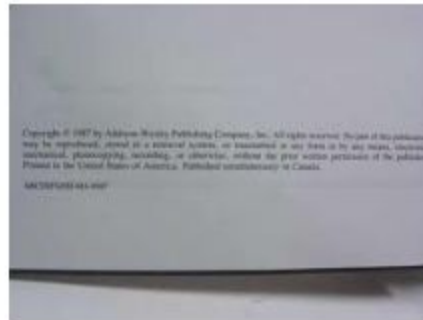
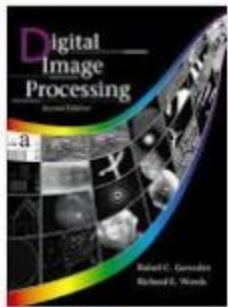
www.buzzmarketingwithblogs.com/.../buyi-... - Перевести эту страницу
Buying the Book. Buzz Marketing with Blogs for **Dummies book cover** image. Buy from Amazon (U.S.) · Buy from Amazon (Canada) · Buy from Amazon (U.K.) ...



[Popular Diets - Chattanooga Public Library - Chattanooga, Tennessee](http://www.chattanooga-public-library.org/.../Popular-Diets)

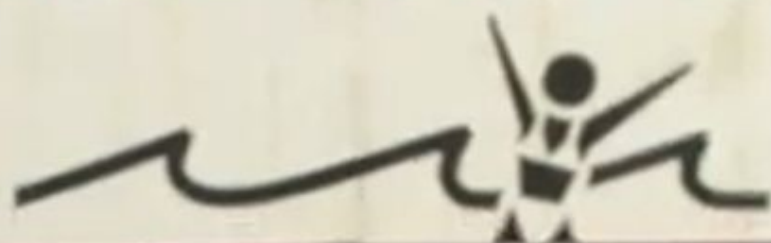


Результатов: примерно 35 100 (0,45 сек.)



PLAYA CERRADA

RECIENTE ATAQUE DE TIBURÓN



Spanish to English



BEACH CLOSED

RECENT ATTACK OF SHARK



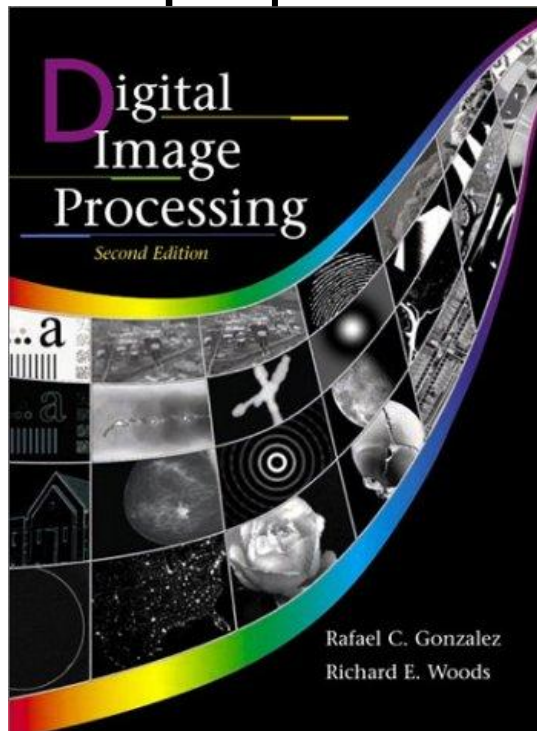
Обнаружение текста – газеты, журналы, книги



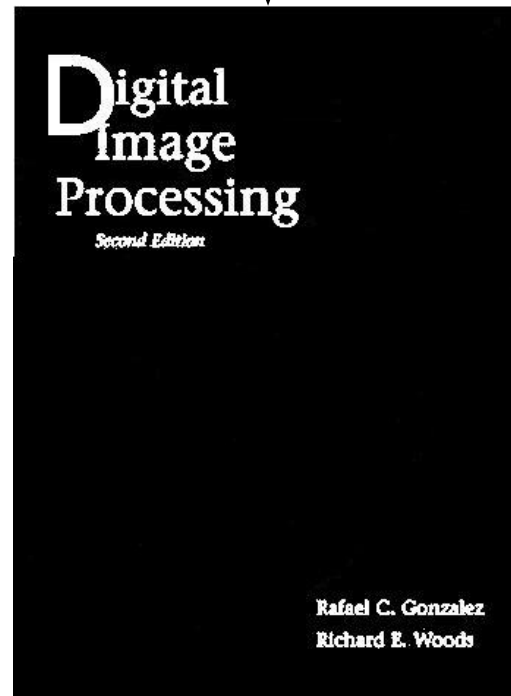
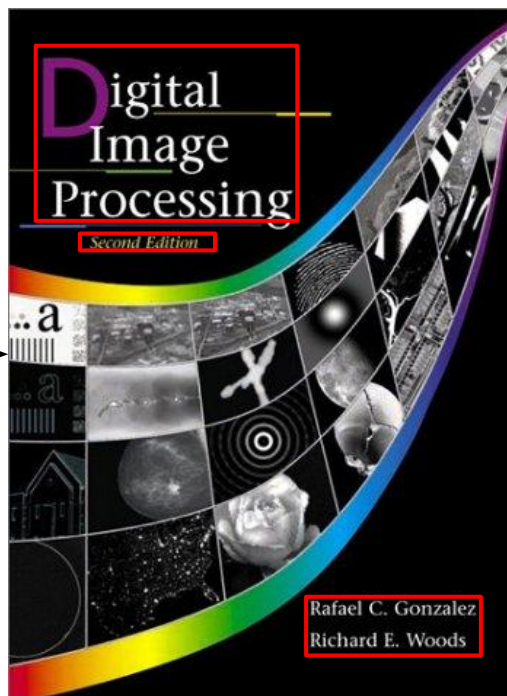
- обнаружение текстовых областей
- определение угла поворота текста (skew detection)
- определение порядка чтения



Обнаружение текста – произвольные изображения

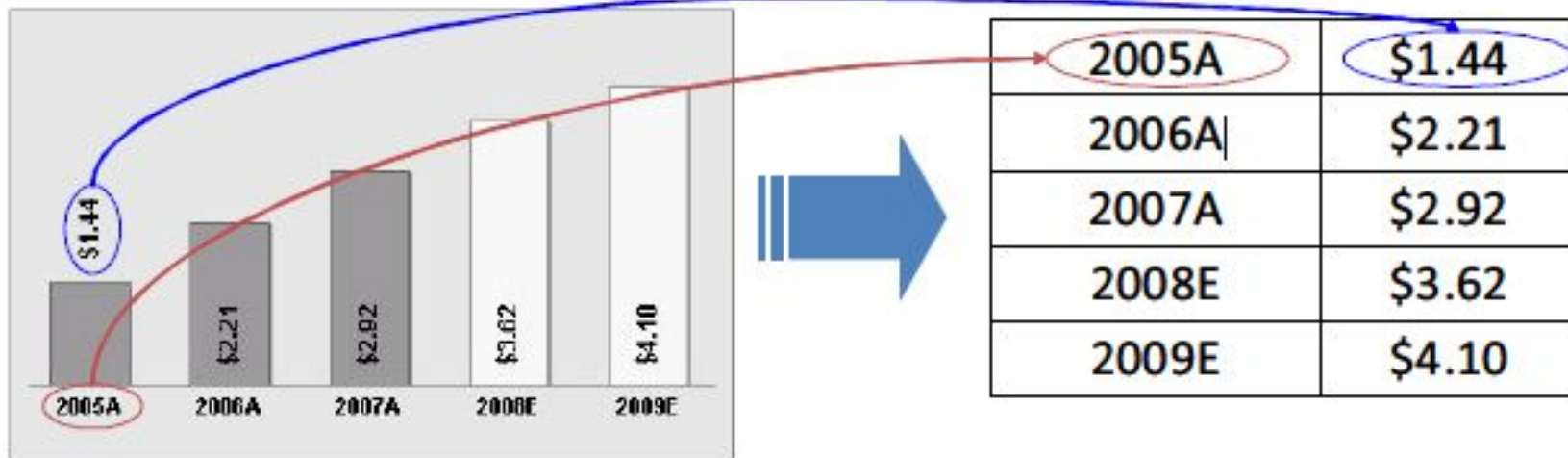


Исходное изображение



Возможные результаты работы алгоритмов обнаружения текста

Обнаружение текста – чертежи и графики



- Обнаружение текстовых областей
- Определение угла поворота текстовых строк

- Короткие фрагменты текста
- Разнообразие шрифтов, текст под разными углами
- Однородный фон
- Высокая контрастность

фотографи
и

печатные
документы

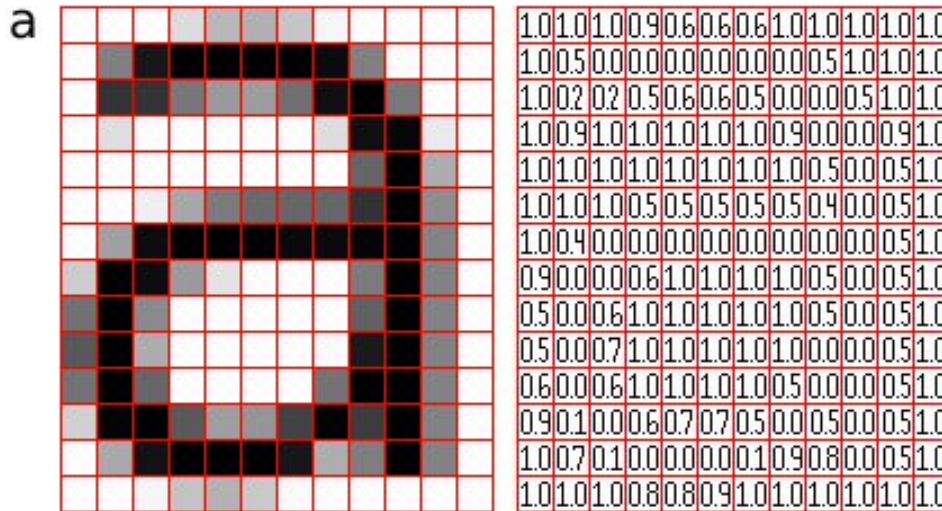
План лекции

- Зачем нужны алгоритмы обнаружения текста?
- Что такое цифровое изображение?
 - Представление цифровых изображений
 - Границы, компоненты связности, бинаризация, преобразование Хафа
- Печатные документы (document images)
- Фотографии (natural scenes)
- Чертежи, графики, обложки



Представление цифровых изображений

Растровое изображение



$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1,N-1) \end{bmatrix}.$$

$$0 \leq f(x, y) \leq L, \quad \text{and typically } L = 255$$

Представление цифровых изображений

Цветное растровое изображение:

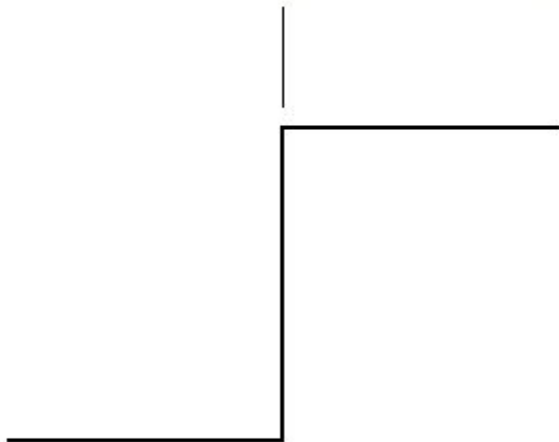
RGB – распространенная модель цвета

Каждый пиксель задается тремя значениями: red, green, blue



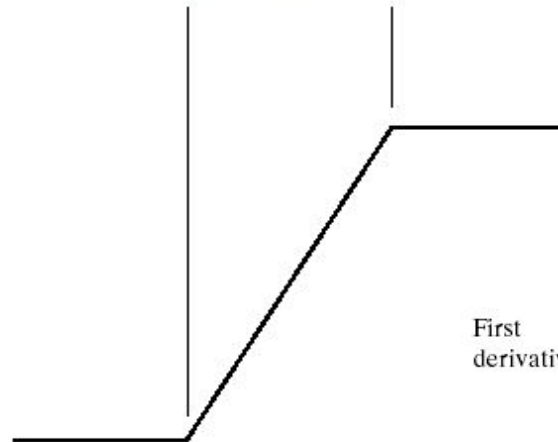
Обнаружение границ

Model of an ideal digital edge

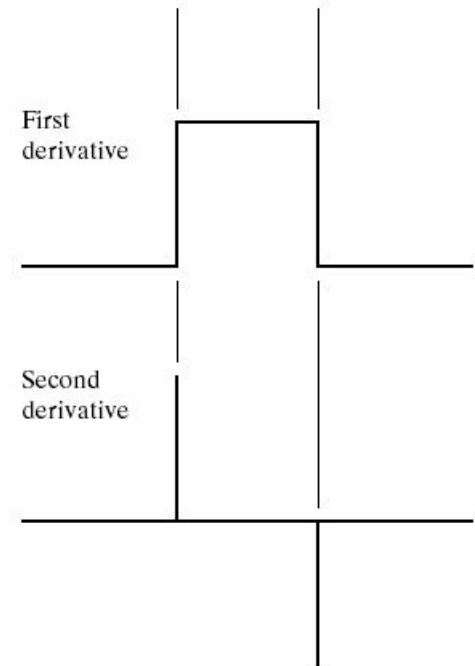


Gray-level profile of a horizontal line through the image

Model of a ramp digital edge



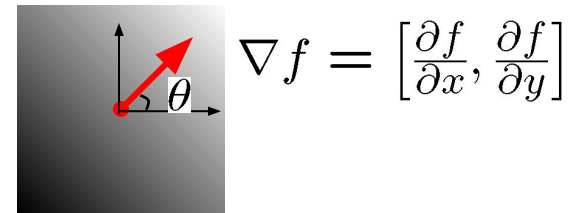
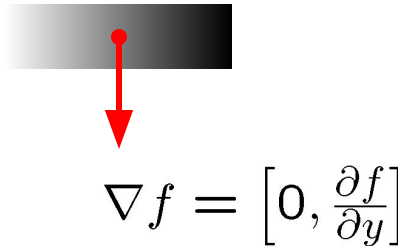
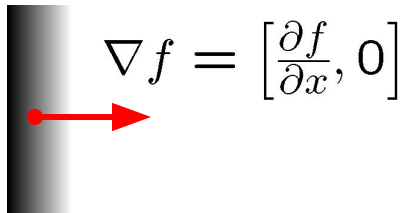
Gray-level profile of a horizontal line through the image



Градиент изображения

$$\nabla f = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right]$$

Градиент направлен в сторону
наибольшего изменения
ИНТЕНСИВНОСТИ



Направление градиента:

$$\theta = \tan^{-1} \left(\frac{\partial f}{\partial y} / \frac{\partial f}{\partial x} \right)$$

Величина градиента:

$$\|\nabla f\| = \sqrt{\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}$$

Вычисление градиента изображения

Дискретный случай:

$$\frac{\partial f}{\partial x}[x, y] \approx f[x + 1, y] - f[x, y]$$

z_1	z_2	z_3
z_4	z_5	z_6
z_7	z_8	z_9

Roberts:

$$G_x = (z_9 - z_5)$$

$$G_y = (z_8 - z_6)$$

-1	0	0	-1
0	1	1	0

Prewitt:

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3)$$

$$G_y = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7)$$

-1	-1	-1	-1	0	1
0	0	0	-1	0	1
1	1	1	-1	0	1

Sobel:

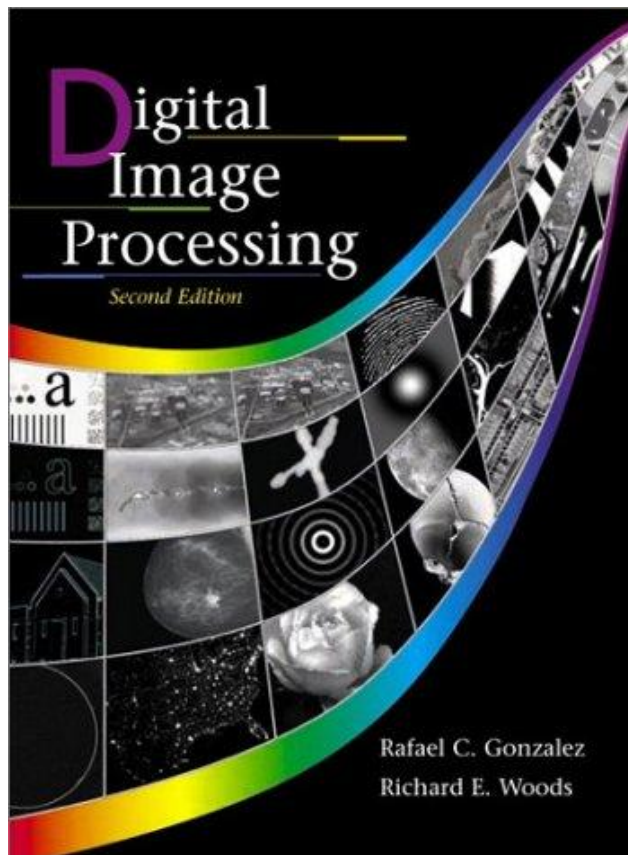
$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$G_y = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

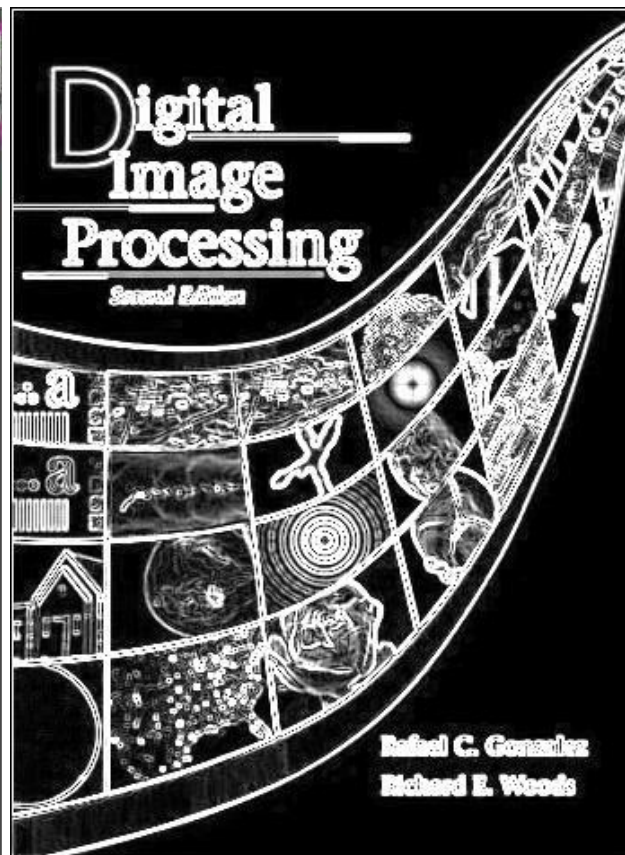
-1	-2	-1	-1	0	1
0	0	0	-2	0	2
1	2	1	-1	0	1



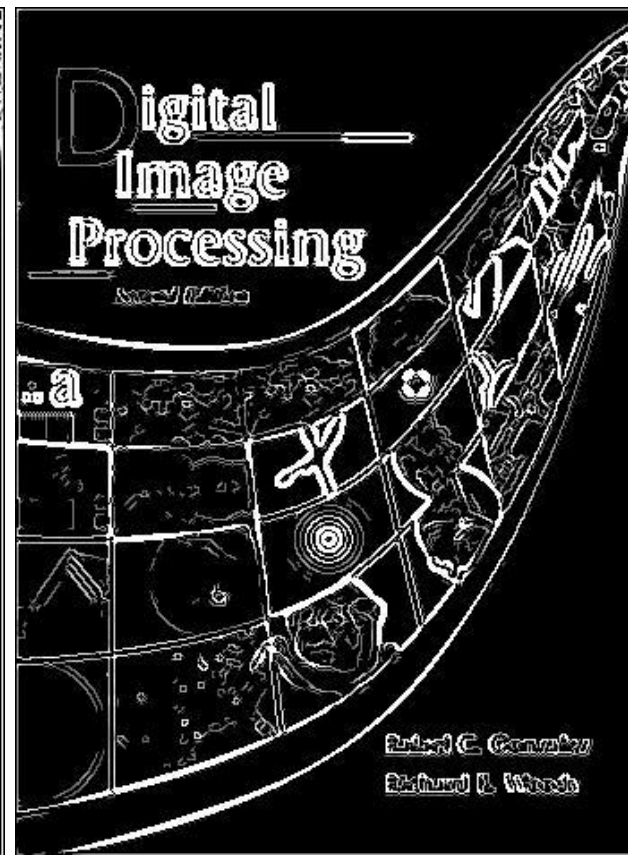
Выделение границ: примеры



Исходное



Sobel



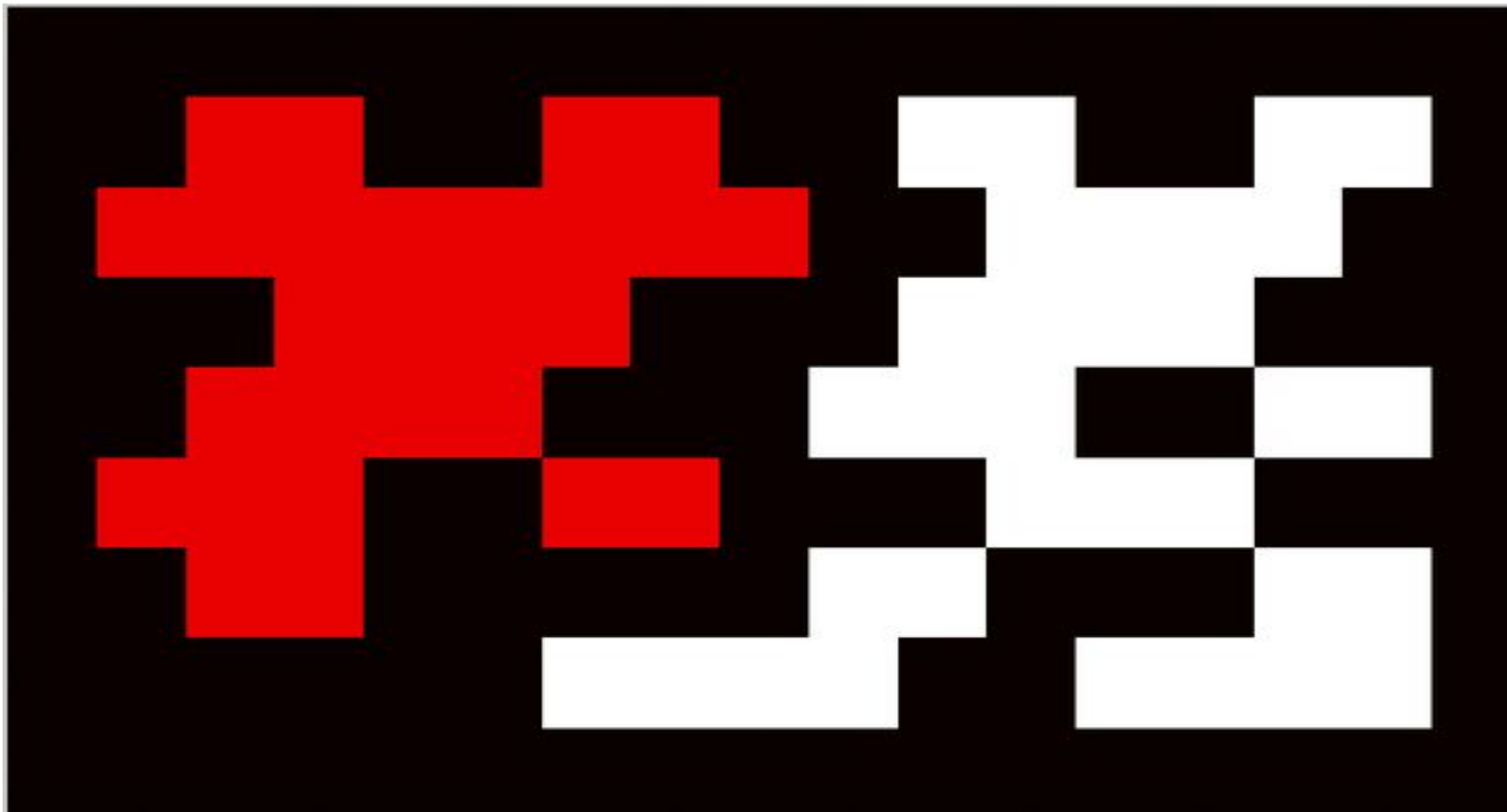
Canny

Компоненты связности

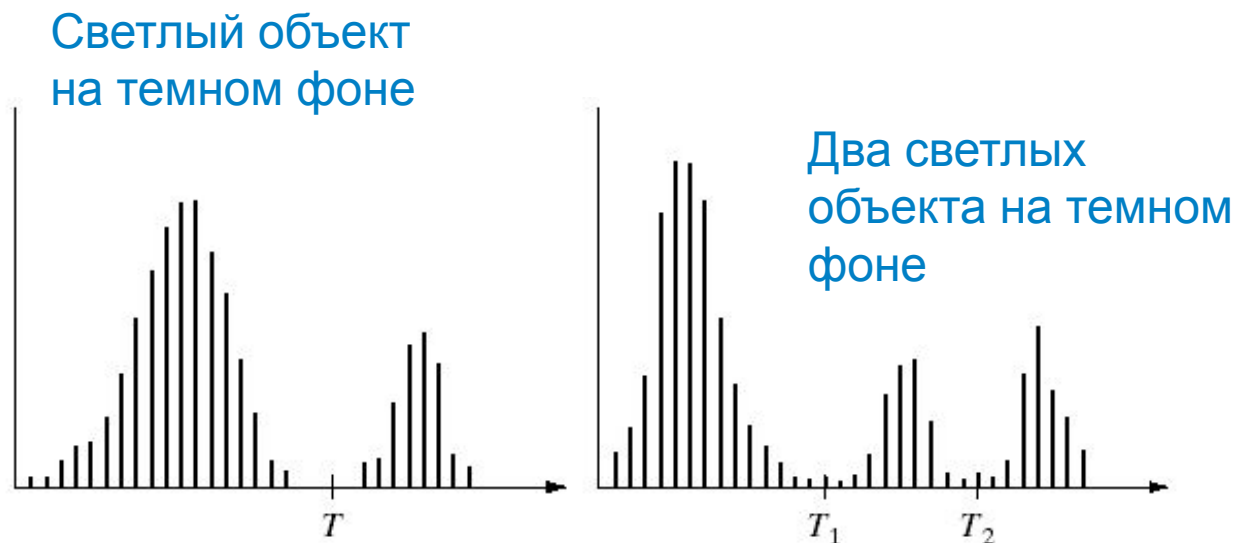
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1	1	0	0	2	2	0	0	3	3	0	0	4	4	0
0	1	1	1	1	1	1	1	1	0	0	3	3	3	3	0	0
0	0	0	1	1	1	1	0	0	0	3	3	3	3	0	0	0
0	0	1	1	1	1	0	0	0	3	3	3	0	0	3	3	0
0	1	1	1	0	0	1	1	0	0	0	3	3	3	0	0	0
0	0	1	1	0	0	0	0	0	5	3	0	0	0	3	3	0
0	0	0	0	0	0	6	6	5	3	0	0	7	3	3	3	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



Компоненты связности



Пороговая бинаризация



a b

(a) Gray-level histograms that can be partitioned by (a) a single threshold, and (b) multiple thresholds.

Глобальная – порог единый для всех точек изображения

Локальная или **Динамическая** – когда порог зависит от координат точки (x,y)

Адаптивная – когда порог зависит от значения яркости в точке $I(x,y)$

Бинаризация

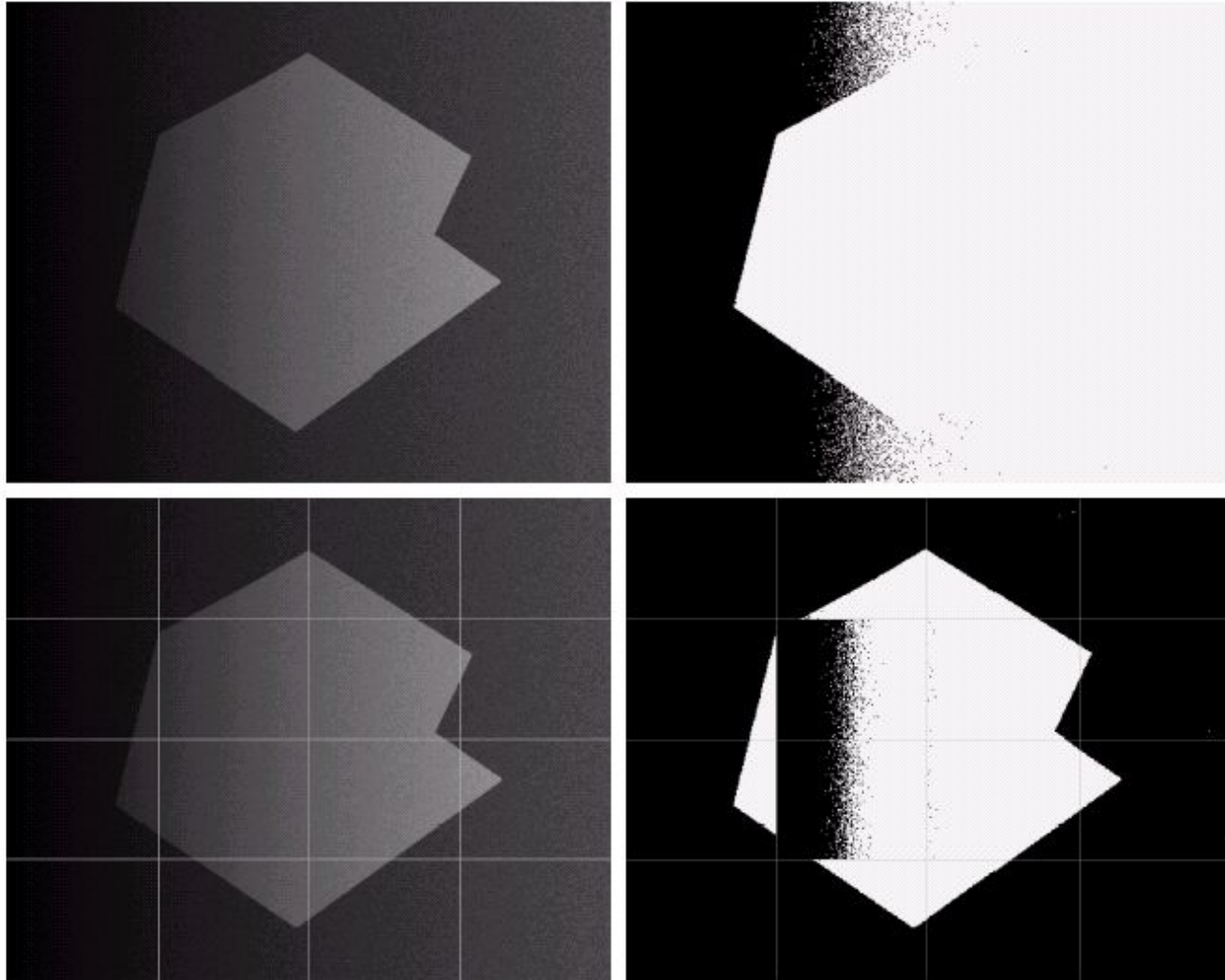
a b
c d

FIGURE 10.30

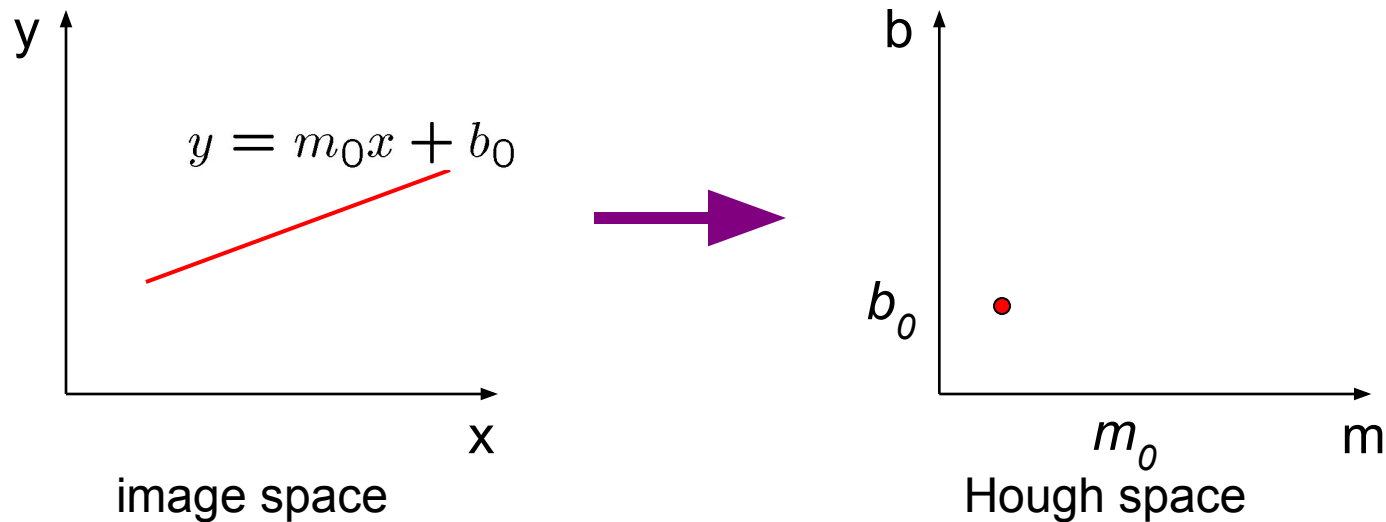
(a) Original image. (b) Result of global thresholding.

(c) Image subdivided into individual subimages.

(d) Result of adaptive thresholding.

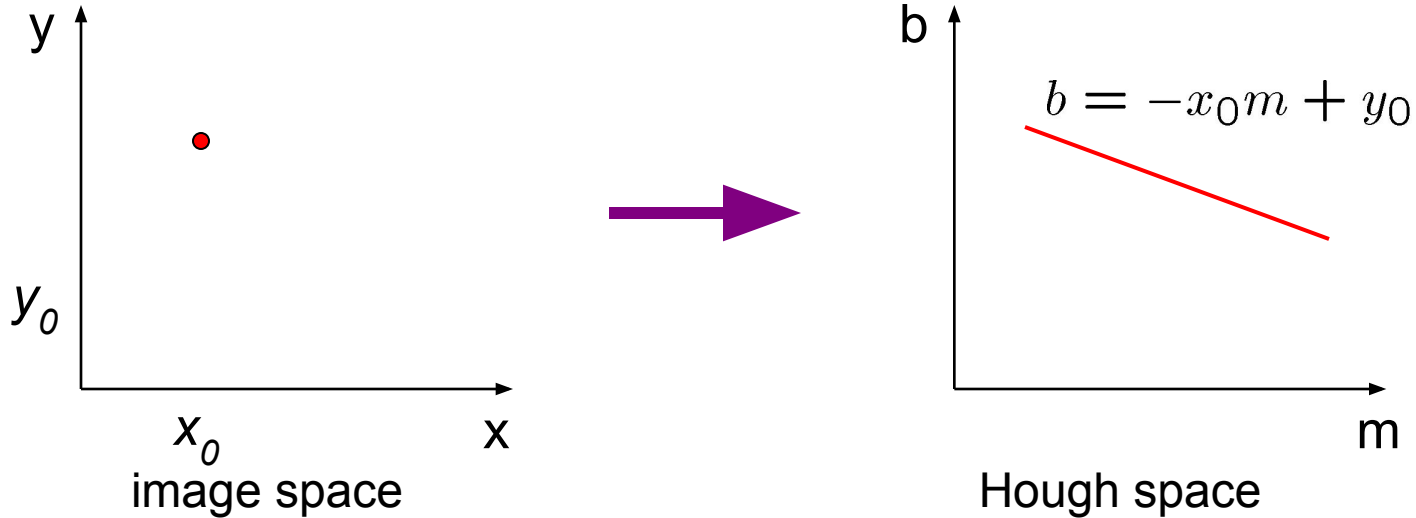


Преобразование Хафа (Hough transform)

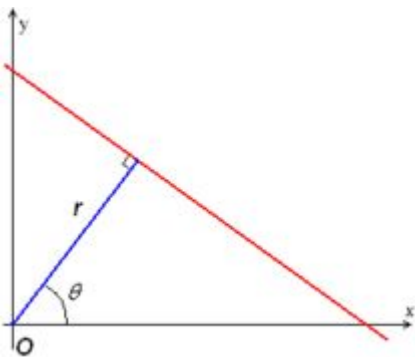


- Для данного набора точек (x, y) найти все точки (m, b) , такие что $y = mx + b$

Преобразование Хафа (Hough transform)



- Для данного набора точек (x, y) найти все точки (m, b) , такие что $y = mx + b$
- Точке (x_0, y_0) соответствует прямая в пространстве Хафа: $b = -x_0 m + y_0$



Обычно, используют полярные

координаты:

$$y = \left(-\frac{\cos \theta}{\sin \theta} \right) x + \left(\frac{r}{\sin \theta} \right), \quad r = x \cos \theta + y \sin \theta.$$

План лекции

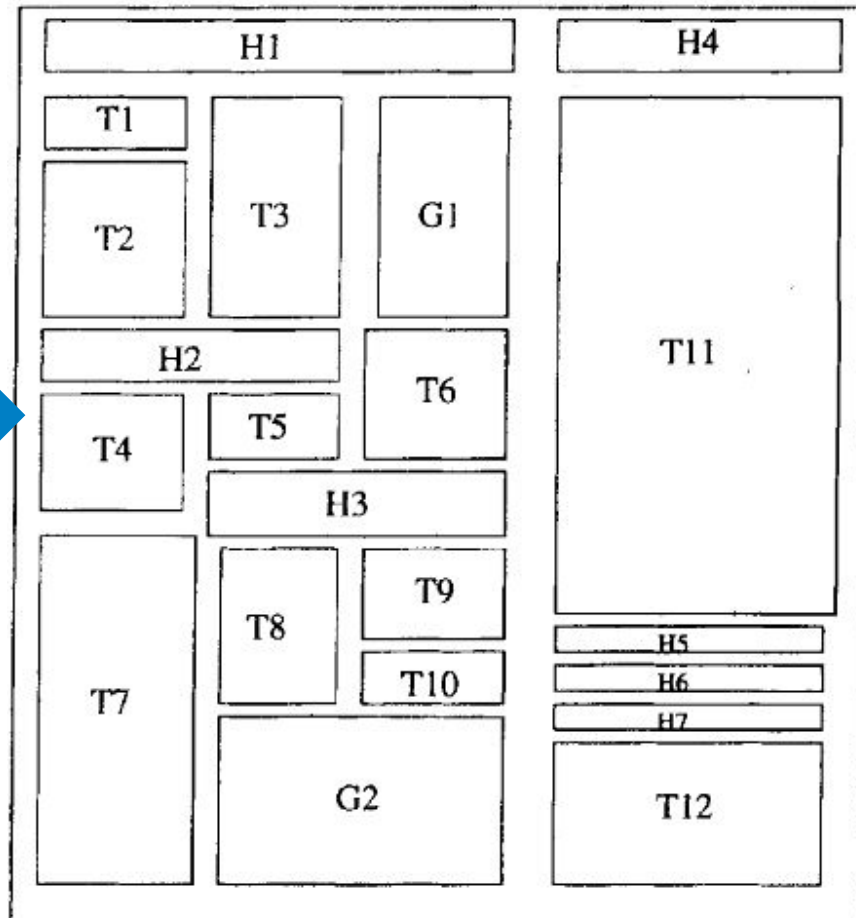
- Зачем нужны алгоритмы обнаружения текста?
- Что такое цифровое изображение?
 - Представление цифровых изображений
 - Границы, компоненты связности, бинаризация, преобразование Хафа
- Печатные документы (document images)
- Фотографии (natural scenes)
- Графики, диаграммы, обложки



Печатные документы

Основные задачи

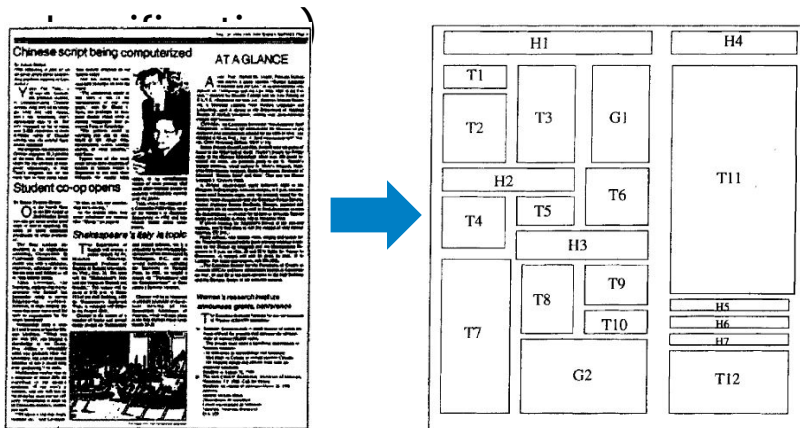
- Анализ структуры страницы
(layout analysis, geometric structure analysis, page segmentation, region



Печатные документы

Основные задачи

- Анализ структуры страницы
(layout analysis, geometric structure analysis, page segmentation, region



- Методы
 - «Сверху-вниз» (top-down)
 - XY-cuts, whitespace segmentation
 - «Снизу-вверх» (bottom-up)
 - группировка ближайших соседей, диаграммы Вороного

Projection profiles and XY-cuts



Precipitation of Lead Films with Copper, Gold, and Silver

Minimum lead film thicknesses were determined by means of electron spectroscopy and electron microscopy. Minimum thicknesses indicated that films obtained with the lowest thicknesses were thicker than films prepared with Ag or Au. Deposits of the different metals showed Ag to be the only metal of the three to form an intermetallic compound with lead, because of the associated formation with its substrate.

Introduction

The process of metal nucleation of silver on lead surfaces has been extensively studied in the past and is well understood. It is known that the nucleation of silver on lead surfaces is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface.

Because the binding energy of a film tends to be lower on lead than on silver, nucleation of silver on lead surfaces is expected to be more favorable than on silver surfaces. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface.

Experimental techniques: The lead, copper, gold, and silver were all obtained by electrolysis from the same solution. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface.



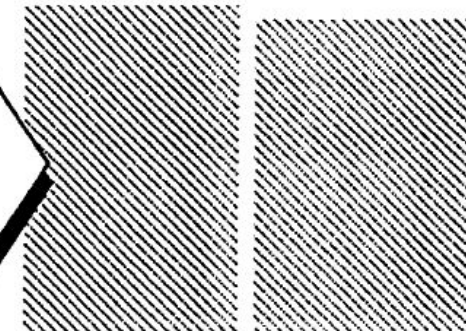
Вертикальна
я проекция

Introduction

The process of metal nucleation of silver on lead surfaces has been extensively studied in the past and is well understood. It is known that the nucleation of silver on lead surfaces is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface.

Because the binding energy of a film tends to be lower on lead than on silver, nucleation of silver on lead surfaces is expected to be more favorable than on silver surfaces. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface.

Experimental techniques: The lead, copper, gold, and silver were all obtained by electrolysis from the same solution. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface. The surface energy of the lead surface is a function of the surface energy of the silver surface and the surface energy of the lead surface. The surface energy of the silver surface is a function of the surface energy of the lead surface and the surface energy of the silver surface.



Горизонтальна
я проекция



Методы «снизу-вверх»

Результат алгоритма Docstrum

J. G. Steele,¹ E. Treasure,² N. B. Pitts,³ J. Morris,⁴ and G. Bradnock,⁵

The 1998 Adult Dental Health Survey, published this year, showed that the number of people without teeth should fall over the next three decades, to only 4% of the UK population. Patterns of tooth loss and retention are also changing. This article, the first of a series on the interpretation of the Adult Dental Health Survey, discusses the implications of these trends for dentistry.

At the time of the first national survey of adult dental health, which was held in 1968 and covered only England and Wales, over one third of the population (37%) had no natural teeth. Even amongst people aged 35–44 at that time, an edentulous mouth was a common finding (22%).¹ Times have changed. This paper will use data from the most recent United Kingdom Adult Dental Health Survey,² to describe the oral health of the nation in 1998. The data were also used to predict what is likely to happen over the next 20 or 30 years and these projections and their

implications may be necessary if an accurate indication of oral health is to be obtained from all patient groups. Data relating to these are also reported here in order to illustrate and discuss some of the important implications for dental practice from the findings of the survey.

The national surveys of Adult Dental Health have given a 10-yearly summary of the clinical condition of adults in the United Kingdom (England and Wales only in 1968; Scotland and Northern Ireland were surveyed later) on three previous occasions.^{1–3} The fourth report in the series was published

by the Office of National Statistics together with the Universities of Birmingham, Dundee, Newcastle-upon-Tyne and Wales

Who had no natural teeth at all in 1998?

The irreversible nature of the two main destructive dental diseases (caries and periodontal disease) dictate that age is always likely to be a principal factor associated with total tooth loss. Figure 1 shows the proportion who do and do not have teeth, plotted against age. Although 87% of all adults had some natural teeth, up to the age of 45 the figure was almost 100%, while over the age of 54 being edentate was still a relatively common occurrence. Amongst people aged 75 and over, those without natural teeth were still in the majority (58%). Nevertheless, the retention of some natural teeth is now sufficiently common that, amongst the 'younger-old' population nearly two thirds (64%) of the 65–74 year age group and more than half of all of the people of 'pen-



Методы «снизу-вверх»

Использование диаграмм Вороного

Myocardial Tumor Necrosis Factor- α Expression Does Not Correlate With Clinical Indices of Heart Failure in Patients on Left Ventricular Assist Device Support

Peter Razeghi, MD, Madhuri Mukhopadhyay, BS, Timothy J. Myers, BS, Janelle N. Williams, BS, Christine S. Moravec, PhD, O. Howard Frazier, MD, and Heinrich Taegtmeier, MD, DPhI

Division of Cardiology, The University of Texas-Houston Medical School, Houston, Texas, St. Luke's Episcopal Hospital and Texas Heart Institute, Houston, Texas, and Cleveland Clinic Foundation, Cleveland, Ohio

Background. Mechanical unloading with a left ventricular assist device (LVAD) can improve clinical indices of heart failure and alter myocardial tumor necrosis factor- α (TNF α) expression, but a correlation between clinical and molecular indices has not been established.

Methods. We enrolled 14 patients with end-stage heart failure treated with drugs and mechanical unloading in a protocol including the collection of myocardial tissue samples at LVAD implantation and explantation. Ten nonfailing donor hearts served as controls. TNF α expression was measured by quantitative reverse transcription polymerase chain reaction. Clinical indices of heart failure were retrospectively analyzed and correlated with myocardial TNF α expression.

Results. Left ventricular end-diastolic dimension decreased ($p < 0.01$) and cardiac index ($p < 0.001$) increased with unloading. Abnormal values of serum sodium,

creatinine, blood urea nitrogen, glutamic-oxaloacetic transaminase, glutamic-pyruvic transaminase, and albumin showed a trend toward normalization with mechanical unloading. TNF α expression was increased in 5 of 14 patients and decreased with mechanical unloading in 4 of them. Surprisingly, there was no correlation between mRNA levels of TNF α and any of the clinical indices studied.

Conclusions. Although clinical indices of heart failure improve and elevated levels of myocardial TNF α expression decrease with mechanical unloading, there is no correlation between the two. Thus, clinical and molecular indices of heart failure in LVAD-supported patients do not always correlate.

Ann Thorac Surg 2001;72:2044-50

© 2001 by The Society of Thoracic Surgeons



Печатные документы

Основные задачи

- Определение поворота текста (page rotation, skew detection)



- Projection profiles (для исходного изображения или компонент связности)
- Использование преобразования Хафа
- Определение угла наклона текстовых строк

- Обнаружение текстовых строк (text line finding, baseline finding)

~~Say you want to select a quiet, centrally located Manhattan hotel. Google returns an overwhelming seven million results in response to the query "new york city hotels." Or, say you are trying to assemble a program committee for an annual conference composed of researchers who have published at the conference in previous years, and to balance it~~

(θ, r)
d **Soggy**



План лекции

- Зачем нужны алгоритмы обнаружения текста?
- Что такое цифровое изображение?
 - Представление цифровых изображений
 - Границы, компоненты связности, бинаризация, преобразование Хафа
- Печатные документы (document images)
- Фотографии (natural scenes)
- Графики, диаграммы, обложки



Фотографии

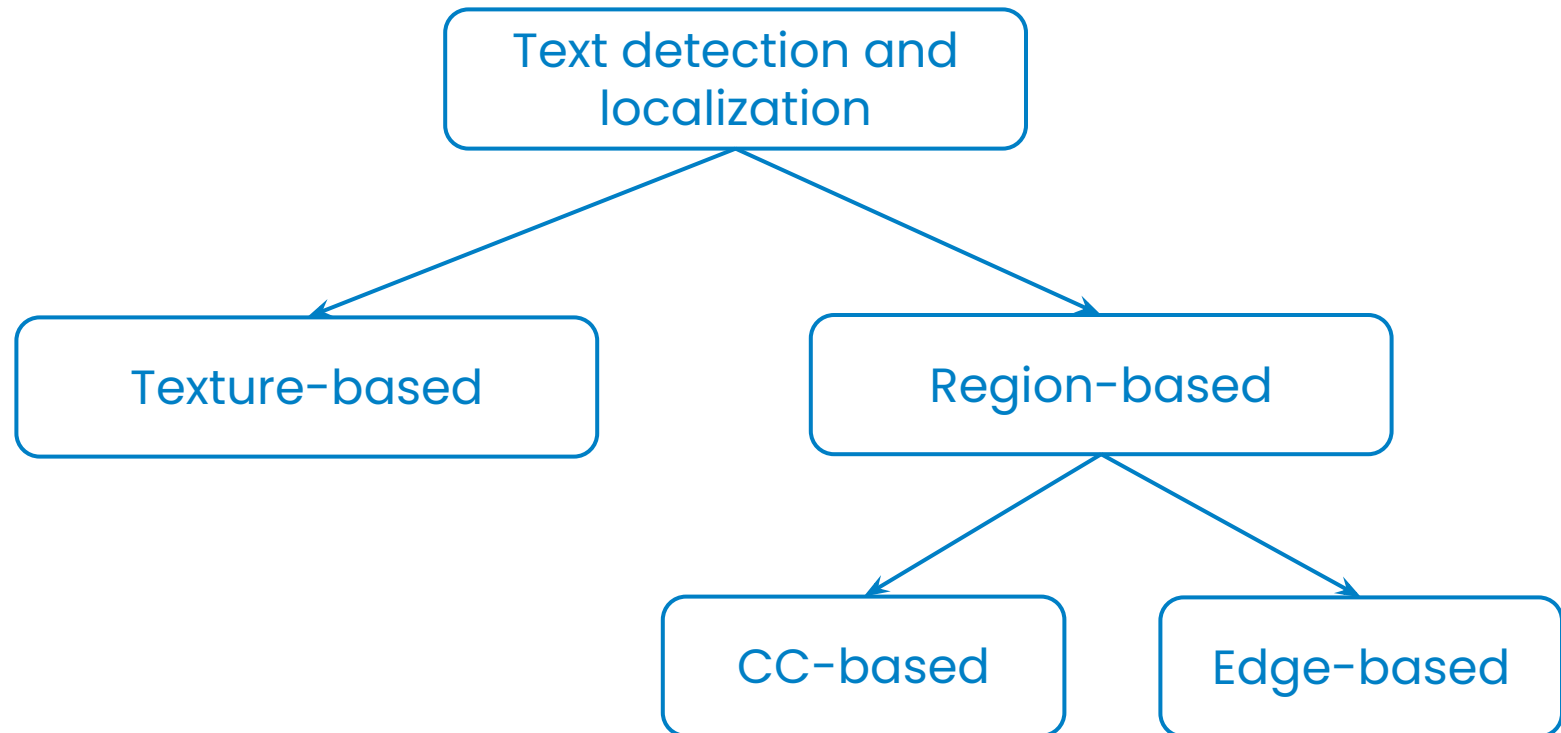
Почему не работают традиционные методы?

- Большое разнообразие шрифтов
- Разнообразии расположений и направлений текстовых строк
- Короткие текстовые строки
- Разнообразии условий съемки (освещение, фокусное расстояние)
- Сложный фон
- Нет определенной структуры страницы
- Наложение объектов (occlusions)



Фотографии

Классификация подходов

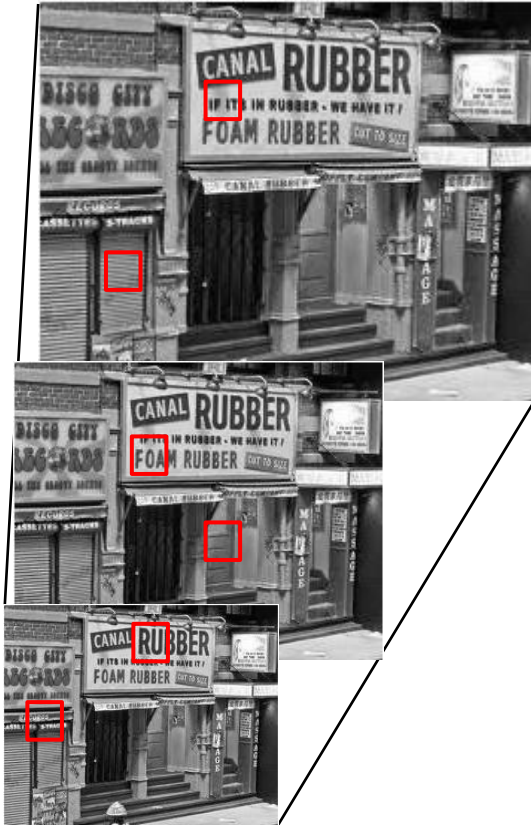


K. Jung et
al.

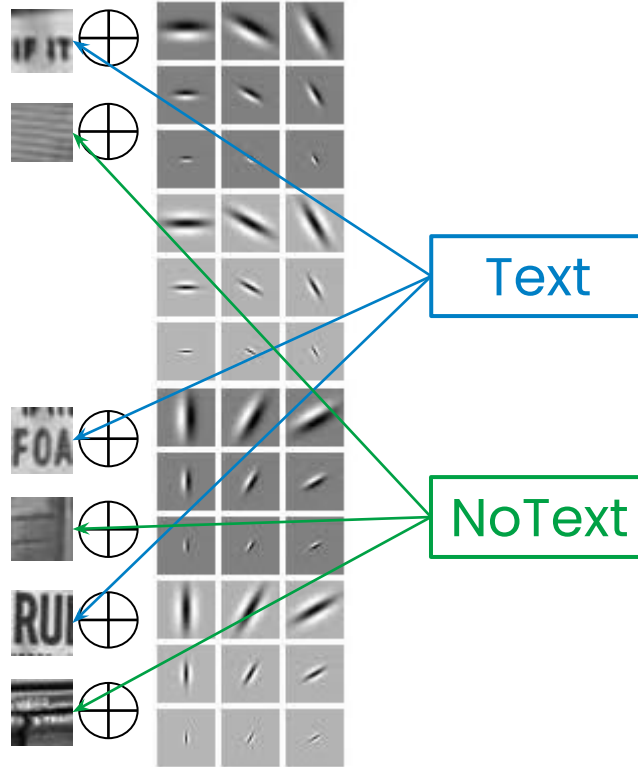


Фотографии

Методы, основанные на анализе текстуры



Построение пирамиды изображений



Извлечение текстурных признаков (Gabor, Wevelets, DCT)

Классификация регионов (SVM)



Объединение результатов

Фотографии

Методы, основанные на анализе текстуры



- Сложный фон



- Вычислительно сложные (обработка нескольких масштабов, операции свертки)
- Произвольная направленность текста (негоризонтальный текст)
- Произвольный размер шрифта



Фотографии

Region-based methods (bottom-up)

- Выделение компонент связности на основе локальных признаков (близкий цвет или принадлежность границе)
- Объединение выделенных компонент связности в группы по признакам близкого расположения и схожих локальных признаков (размер, цвет)



Фотографии

Region-based methods



- Произвольный размер шрифта
- Произвольная направленность текста
- Просты в реализации



- Сложный фон
- Шум и нерезкость изображения
- Используют большое количество эвристик



Stroke Width Transform (SWT)



(a)

Исходное
изображение



(b)

Результат SWT



(c)

После фильтрации
по признаку
постоянства
ширины штриха

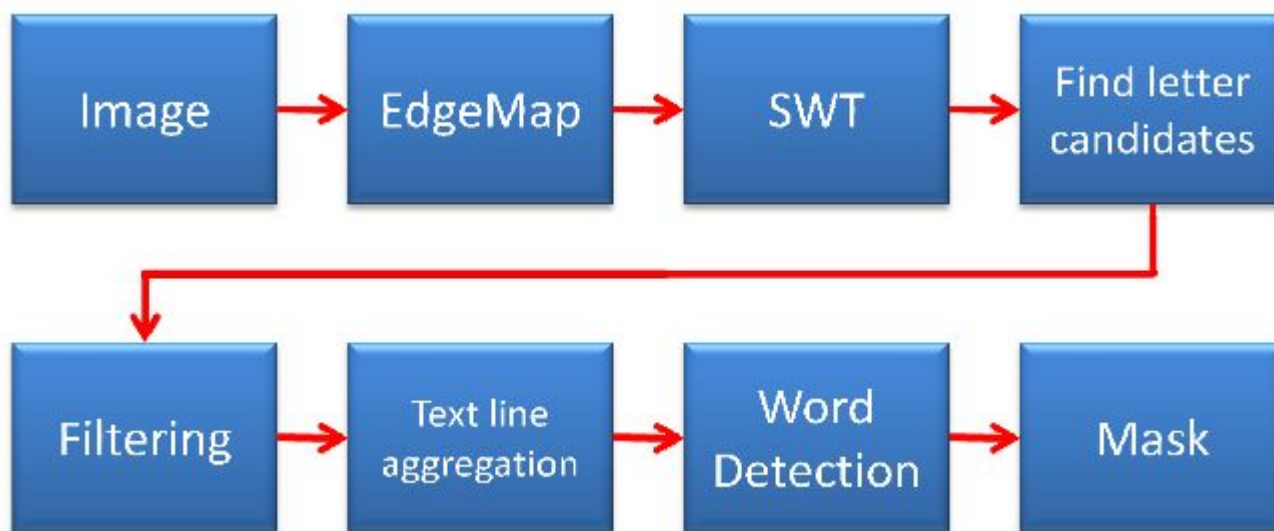


(d)

Найденный
текст

B. Epshtein et
al.

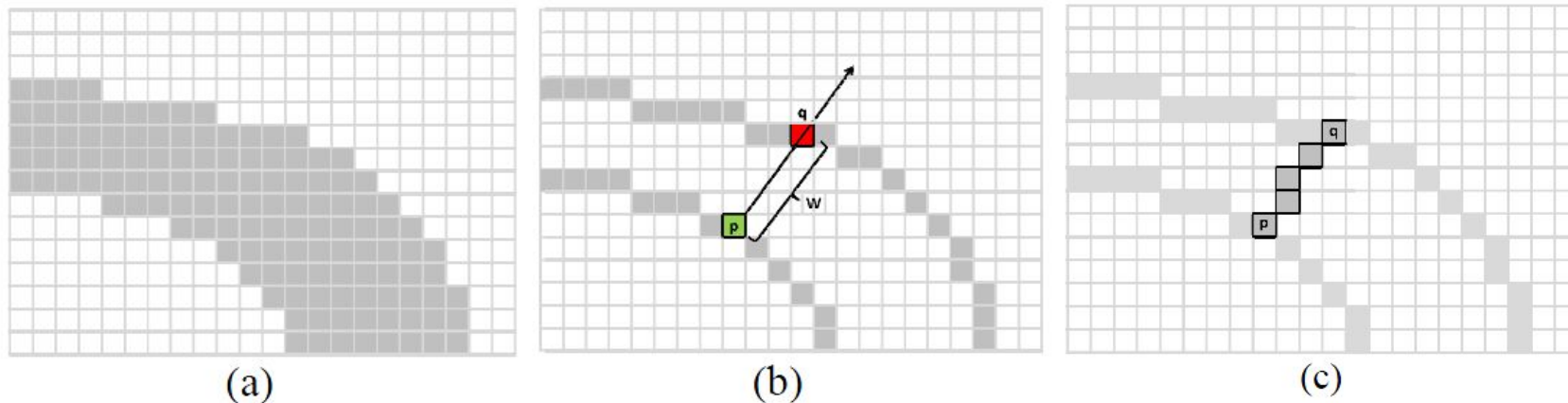
Обнаружение текста при помощи SWT



B. Epshtein et
al.



Вычисление SWT



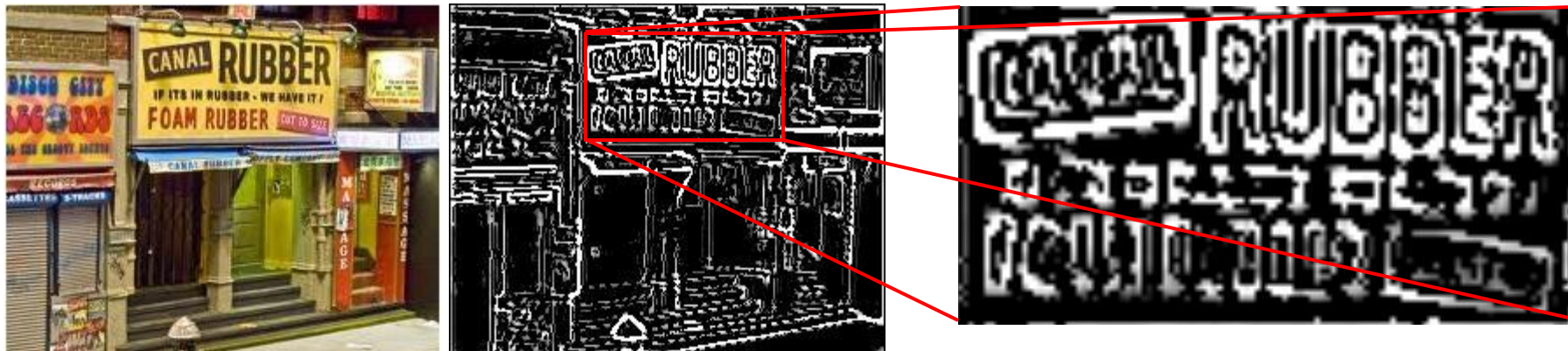
(a) Фрагмент штриха

(b) p – пиксель на границе штриха,
 q – пиксель на противоположной стороне штриха
(градиенты в p и q направлены друг на друга)

(c) Всем пикселям вдоль луча pq присваивается значение ширины штриха

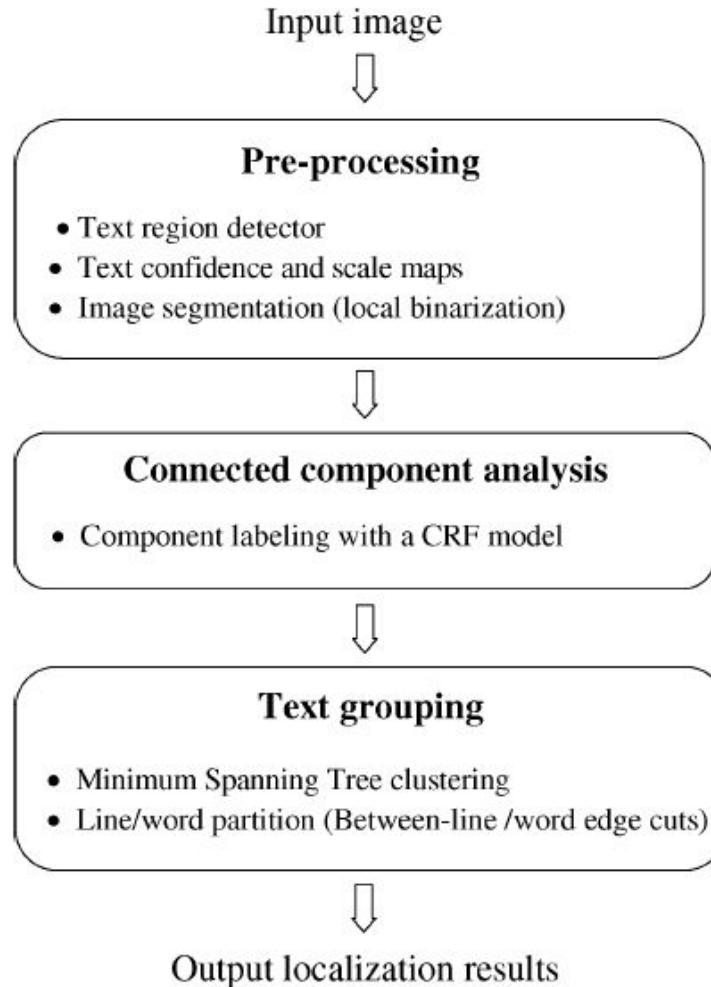
В. Epshtein et al.

Обнаружение текста с помощью SWT



- Границы для нерезких изображений, низкого разрешения – 😞
- Погрешность SWT на стыках штрихов – 😞
- Эвристики для фильтрации компонент – 😞
- Двойной проход и интеграция результатов – 😞

Комбинированные методы



Y.-F. Pan et
al.



Комбинированные методы

Шаг 1 – анализ текстуры

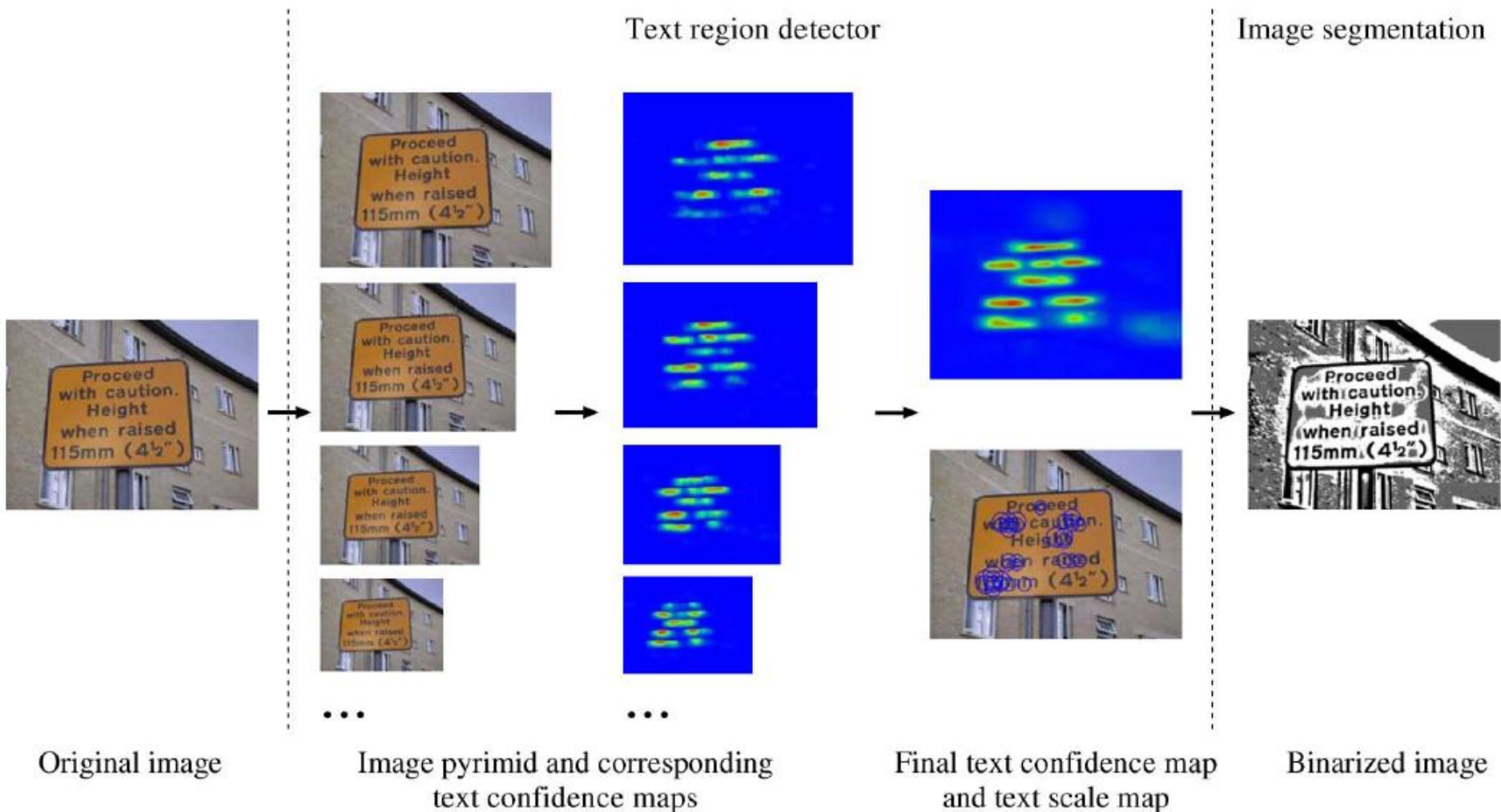


Fig. 5. Example of the pre-processing stage.

Комбинированные методы

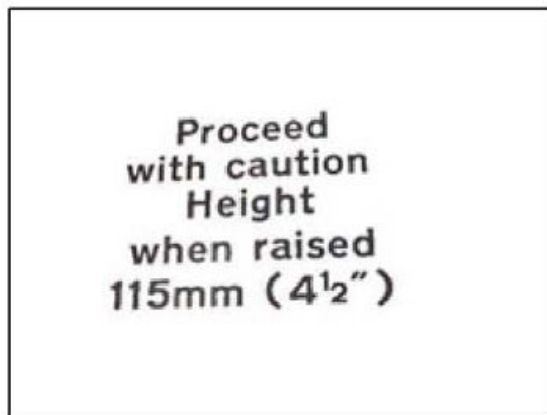
Шаг 2 – анализ компонент связности



(a)



(b)



(c)

Fig. 6. Example of the CCA stage. (a) Components passing through unary classification thresholds. (b) Component neighborhood graph. (c) Text components after CRF labeling.

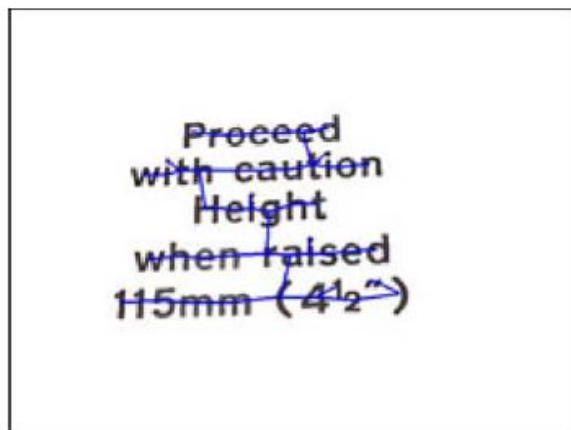
otice.

Y.-F. Pan et al.

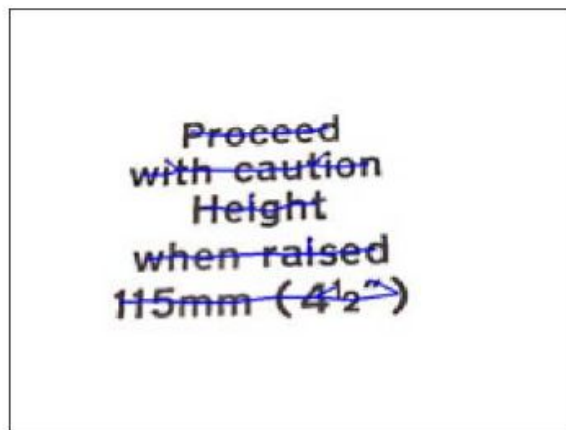


Комбинированные методы

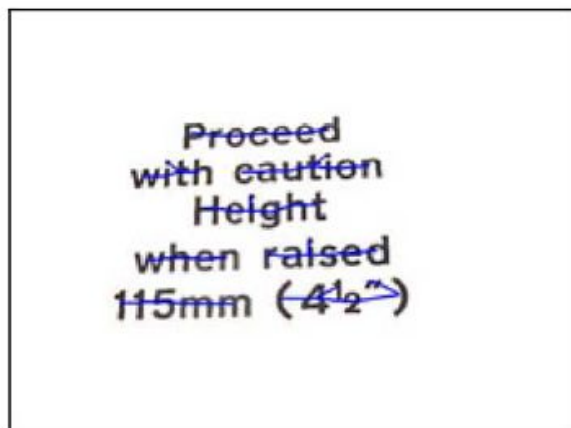
Шаг 3 – выделение текстовых строк и слов



(a)



(b)



(c)



(d)

- построение минимального остовного дерева
- решение оптимизационной задачи

Fig. 7. Example of the text grouping stage. (a) Building component tree with MST. (b) Text line partition. (c) Text word partition. (d) Text localization results.

ICDAR (2003, 2005, 2009, 2011)

Robust Reading Competitions

- Распознавание символов
- Распознавание слов
- Локализация текста
- Распознавание текста

Algorithm	Precision	Recall	f	Time (sec.)
SWT	0.73	0.60	0.66	0.94
Hinnerk Becker*	0.62	0.67	0.62	14.4
Alex Chen	0.60	0.60	0.58	0.35
Qiang Zhu	0.33	0.40	0.33	1.6
Jisoo Kim	0.22	0.28	0.22	2.2
Nobuo Ezaki	0.18	0.36	0.22	2.8
Ashida	0.55	0.46	0.50	8.7
HWDavid	0.44	0.46	0.45	0.3
Wolf	0.30	0.44	0.35	17.0
Todoran	0.19	0.18	0.18	0.3
Full	0.1	0.06	0.08	0.2
Hybrid Approach	67.4	69.7	68.5	2.43



otice.

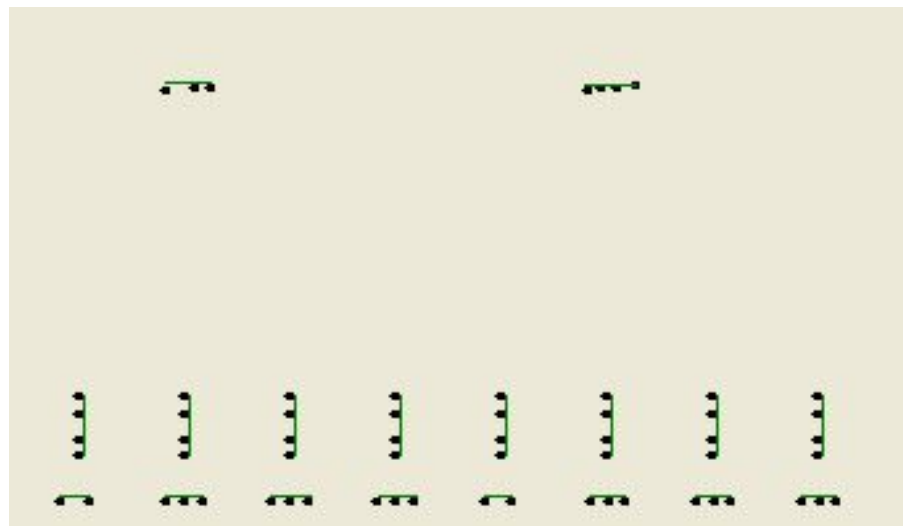
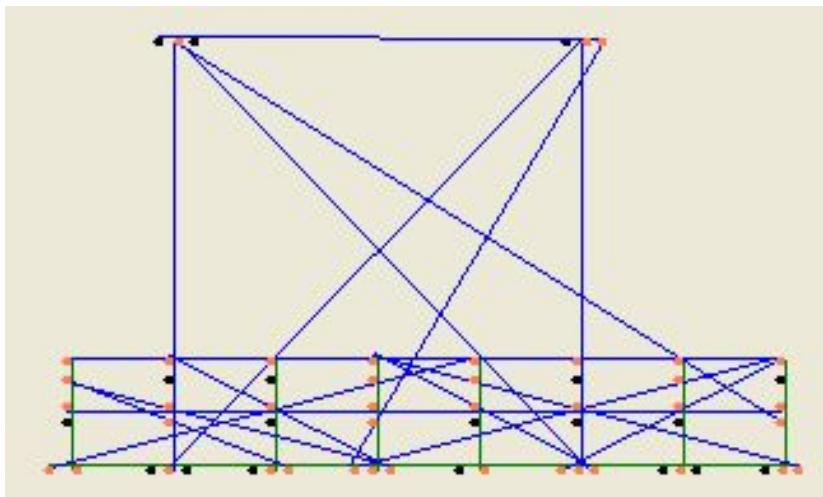
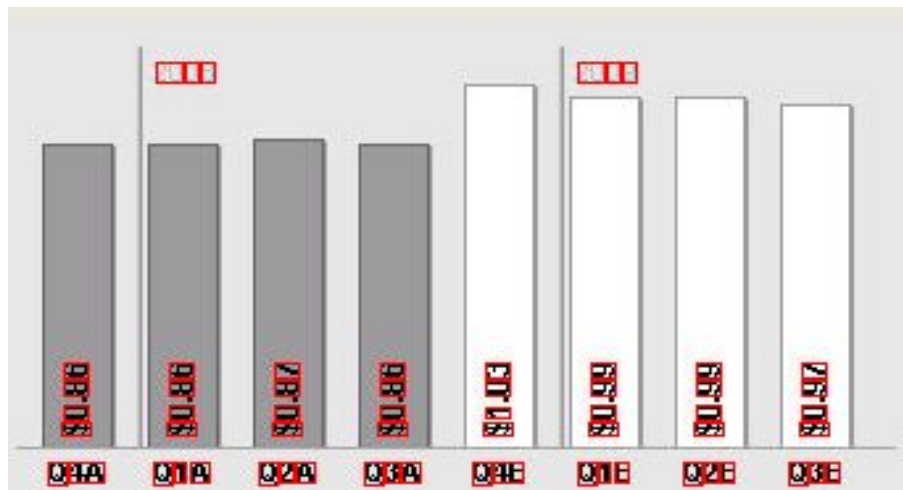
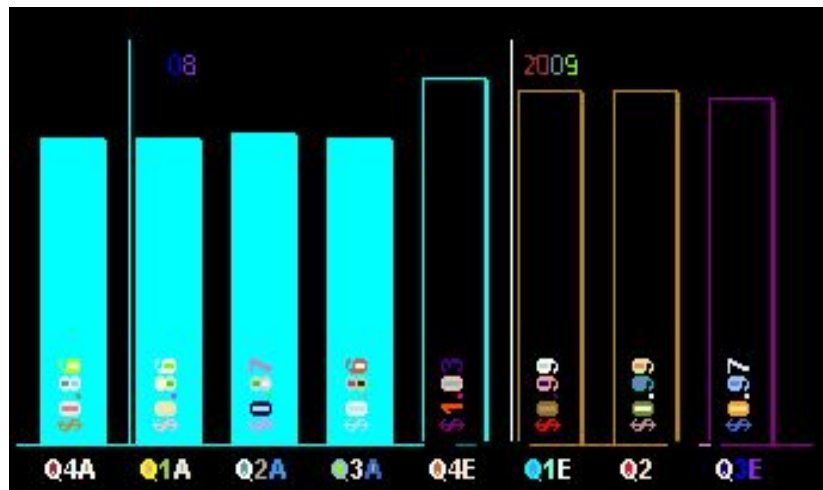
План лекции

- Зачем нужны алгоритмы обнаружения текста?
- Что такое цифровое изображение?
 - Представление цифровых изображений
 - Границы, компоненты связности, бинаризация, преобразование Хафа
- Печатные документы (document images)
- Фотографии (natural scenes)
- Графики, диаграммы, обложки

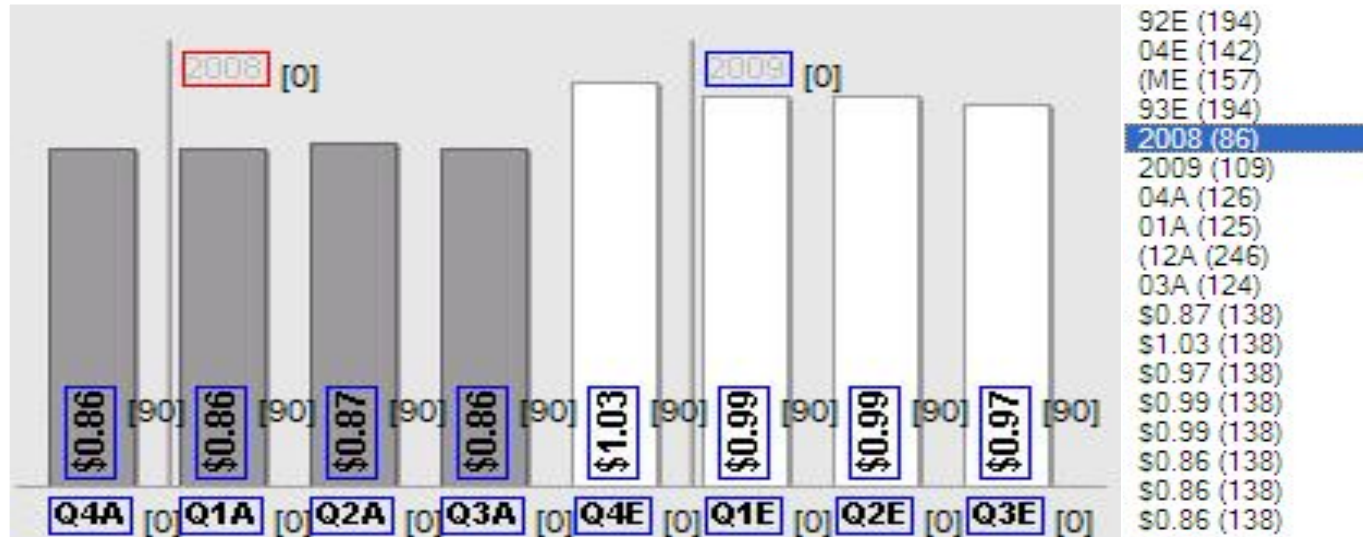


Графики и диаграммы

Анализ компонент связности по цвету



Экспериментальная оценка



Тестовое множество: 1000 диаграмм, сгенерированных SWF/XML tool

	LRR	LPR	TRR	TPR
Preprocess	79.0%	88.7%	44.7%	44.6%
NoPreprocess	33.7%	84.2%	2.5%	2.9%

Location Recognition Rate =

$\frac{N_{Loc}}{N_G}$ Location Precision Rate =

$\frac{N_{Text}}{N_F}$ Text Recognition Rate =

$\frac{N_{Text}}{N_F}$ Text Precision Rate =

$\frac{N_{Text}}{N_F}$

N_G – the total number of text

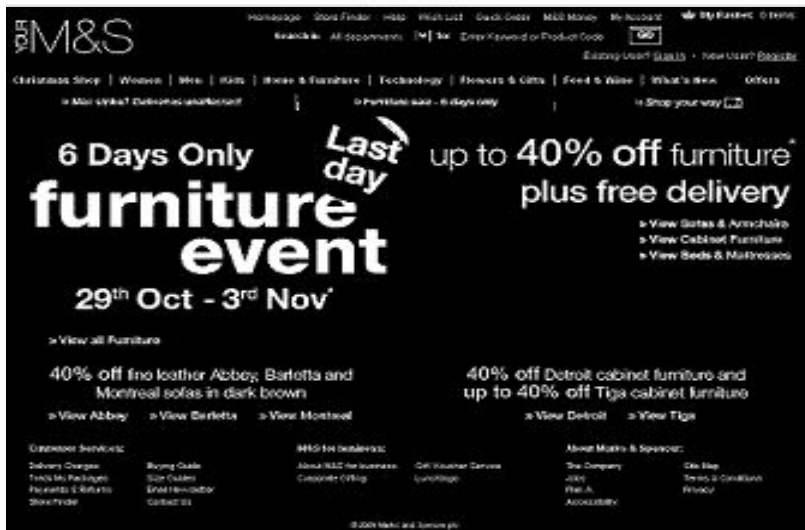
N_F – the total number of detected text blocks

N_{Loc} – the number of correctly localized text

N_{Text} – the number of correctly recognized text blocks



Screenshots



- Выделение границ и пороговая бинаризация
- Удаление длинных горизонтальных и вертикальных границ
- Выделение компонент связности
- Классификация компонент связности и адаптивная бинаризация
- Быстрая бинаризация
 - Время обработки изображения 1600x1008
 - Tesseract OCR: ~6.56 секунд
 - Данный алгоритм: ~0.45 seconds
- (-) Требуется фильтрация компонент
 - Naïve Bayes



Заключение

- Зачем нужны алгоритмы обнаружения текста?
 - Распознавание текста: оцифровка, индексирование, извлечение информации, автоматический перевод, text to speech, навигация
- Печатные документы (document images)
 - Анализ структуры документа, определение поворота, выделение текстовых строк
- Фотографии (natural scenes)
 - Texture-based & region-based
 - Stroke Width Transform, Hybrid approach
- Графики, диаграммы, обложки

