

# Simple Algorithm for Page Layout Analysis

A. O. Shigarov and R. K. Fedorov

*Institute for System Dynamics and Control Theory, Siberian Branch, Russian Academy of Sciences,  
ul. Lermontova 134, Irkutsk, 664033 Russia*

**Abstract**—An algorithm for page layout analysis (segmentation) is suggested in the paper. It allows whitespace between text blocks to be detected on a document page. The algorithm could be used in document analysis and recognition problems. In particular, it can be used for column recognition in multicolumn text and tables. The suggested algorithm is quite simple for implementation.

## INTRODUCTION

The number of documents in the world is growing rapidly; this is discussed in [6] in particular. Document analysis and recognition (DAR) systems are developed for automation of information extraction from document images [7]. Document layout analysis or page segmentation is used for dividing document into specific parts (e.g., columns, figures, tables). This is an important DAR problem. Different methods of document layout analysis are discussed in [3].

Conventionally, there are two approaches to the problem of page or table column segmentation. The first approach is to analyze of text layout (text blocks); this usually requires the use of complex data structures. For example, such an approach with the use of a Voronoi diagram for page segmentation is suggested in [5]. The second approach is to analyze whitespace (free of text) on a page. Gaps separate text and table columns on a page, as is shown in Fig. 1. Algorithms using whitespace analysis are suggested in [1, 2, 8]. Algorithms [1, 8] are briefly considered in [2]. They can be used for detecting whitespace between text blocks. The authors of [2] point out that the algorithms [1, 8] are difficult for implementation. A geometrical layout analysis algorithm, simple for implementation, is presented in [2]; it provides for detecting whitespace gaps in multicolumn text and is described in terms of the largest empty rectangle problem [4]. The algorithm input is a bounding box including obstacles (rectangles). The algorithm [2] finds the largest empty rectangle among obstacles inside the bounding box. To find whitespace gaps on a page,  $n$ -better solutions (empty rectangles) in descending order are found on the page. The algorithm [2] is greedy. The found  $i$ -largest empty rectangle becomes an obstacle when searching for the next  $i + 1$ -largest empty rectangle. However, it is probable that some of  $n$ -better solutions (largest rectangles found) are not gaps between columns, but, e.g., horizontal gaps between paragraphs or

tables on the page. This is the disadvantage of the algorithm [2].

In the paper, we suggest an unconventional simple algorithm for detecting whitespace gaps on a document page. It allows detecting vertical gaps (visually they are stretched up vertically), as well as horizontal gaps if the X and Y axis are exchanged. The suggested algorithm is simple; its Object Pascal implementation consists of about 60 lines of code (expressions).

## 1. PROBLEM FORMULATION

The geometrical objects considered in this work are presented in the Cartesian coordinate system, where the  $x$ -coordinates increase from left to right and  $y$ -coordinates increase from the top down. The following assumptions are used in this work. The rectangle (e.g., an obstacle, bounding box, or gap)  $r = (x_1, y_1, x_2, y_2)$  is defined by the coordinates of its sides (boundaries): left  $x_1 = x_1(r)$ , top  $y_1 = y_1(r)$ , right  $x_2 = x_2(r)$ , and bottom  $y_2 = y_2(r)$ ; in addition, its sides are parallel to the corresponding coordinate axes. The vertical line  $l = (x, y_1, y_2)$  is normal to the X axis and is defined by its coordinates:  $x = x(l)$  by the  $x$ -coordinate,  $y_1 = y_1(l)$  by the minimum (top)  $y$ -coordinate, and  $y_2 = y_2(l)$  by the maximum (bottom)  $y$ -coordinate.

It is assumed that the bounding rectangle  $b$  and the set of obstacles (rectangles)  $R = \{r_1, \dots, r_n\}$ ,  $n \in \mathbb{N}$  are specified. The rectangle  $b$  usually bounds a document page or its part (e.g., a table), while the obstacles are bounding rectangles for text blocks (e.g., words or lines). The obstacles from the set  $R$  are totally inside the bounding rectangle  $b$  and do not overlap each other.

Let us define two sets of obstacles (rectangles)  $R^+ = \{r_1, r_2, \dots, r_m, r_3\}$  and  $R^- = \{r_1, r_2, \dots, r_m, r_3\}$ , where  $r_1 = (x_1, y_1, x_2, y_2)$ ,  $r_2 = (x_m, y_m, x_m, y_2)$ ,  $r_3 = (x_1, y_1, x_m, y_2)$ ,  $r_3 = (x_1, y_2, x_2, y_2)$ ,  $x_1 = x_1(b)$ ,  $y_1 = y_1(b)$ ,  $x_2 = x_2(b)$ , and  $y_2 = y_2(b)$ . Let us also define the gap as a rectangle bounding a certain part of whitespace inside the bounding rectangle  $b$ .

# Simple Algorithm Page Layout Analysis Springerlink

**J Dewey**



**Simple Algorithm Page Layout Analysis Springerlink:**

The Captivating Realm of Kindle Books: A Thorough Guide Revealing the Pros of E-book Books: A Realm of Ease and Versatility E-book books, with their inherent portability and simplicity of availability, have freed readers from the constraints of hardcopy books. Gone are the days of lugging cumbersome novels or carefully searching for particular titles in shops. E-book devices, sleek and lightweight, effortlessly store an extensive library of books, allowing readers to immerse in their favorite reads whenever, everywhere. Whether traveling on a busy train, relaxing on a sun-kissed beach, or just cozying up in bed, E-book books provide an exceptional level of convenience. A Literary Universe Unfolded: Discovering the Vast Array of Kindle Simple Algorithm Page Layout Analysis Springerlink Simple Algorithm Page Layout Analysis Springerlink The E-book Store, a virtual treasure trove of bookish gems, boasts an extensive collection of books spanning diverse genres, catering to every readers preference and preference. From gripping fiction and mind-stimulating non-fiction to classic classics and modern bestsellers, the Kindle Store offers an unparalleled abundance of titles to discover. Whether seeking escape through immersive tales of fantasy and exploration, delving into the depths of past narratives, or broadening ones knowledge with insightful works of scientific and philosophical, the Kindle Shop provides a doorway to a literary universe brimming with limitless possibilities. A Game-changing Factor in the Literary Landscape: The Enduring Influence of Kindle Books Simple Algorithm Page Layout Analysis Springerlink The advent of E-book books has certainly reshaped the bookish landscape, introducing a model shift in the way books are released, distributed, and consumed. Traditional publication houses have embraced the digital revolution, adapting their strategies to accommodate the growing demand for e-books. This has led to a rise in the accessibility of E-book titles, ensuring that readers have entry to a vast array of bookish works at their fingers. Moreover, Kindle books have equalized entry to books, breaking down geographical limits and providing readers worldwide with equal opportunities to engage with the written word. Regardless of their location or socioeconomic background, individuals can now engross themselves in the captivating world of literature, fostering a global community of readers. Conclusion: Embracing the Kindle Experience Simple Algorithm Page Layout Analysis Springerlink E-book books Simple Algorithm Page Layout Analysis Springerlink, with their inherent convenience, flexibility, and wide array of titles, have unquestionably transformed the way we encounter literature. They offer readers the freedom to discover the limitless realm of written expression, anytime, anywhere. As we continue to travel the ever-evolving online landscape, E-book books stand as testament to the persistent power of storytelling, ensuring that the joy of reading remains reachable to all.

[https://matrix.jamesarcher.co/About/publication/HomePages/gibi\\_do\\_chaves.pdf](https://matrix.jamesarcher.co/About/publication/HomePages/gibi_do_chaves.pdf)

## **Table of Contents Simple Algorithm Page Layout Analysis Springerlink**

1. Understanding the eBook Simple Algorithm Page Layout Analysis Springerlink
  - The Rise of Digital Reading Simple Algorithm Page Layout Analysis Springerlink
  - Advantages of eBooks Over Traditional Books
2. Identifying Simple Algorithm Page Layout Analysis Springerlink
  - Exploring Different Genres
  - Considering Fiction vs. Non-Fiction
  - Determining Your Reading Goals
3. Choosing the Right eBook Platform
  - Popular eBook Platforms
  - Features to Look for in an Simple Algorithm Page Layout Analysis Springerlink
  - User-Friendly Interface
4. Exploring eBook Recommendations from Simple Algorithm Page Layout Analysis Springerlink
  - Personalized Recommendations
  - Simple Algorithm Page Layout Analysis Springerlink User Reviews and Ratings
  - Simple Algorithm Page Layout Analysis Springerlink and Bestseller Lists
5. Accessing Simple Algorithm Page Layout Analysis Springerlink Free and Paid eBooks
  - Simple Algorithm Page Layout Analysis Springerlink Public Domain eBooks
  - Simple Algorithm Page Layout Analysis Springerlink eBook Subscription Services
  - Simple Algorithm Page Layout Analysis Springerlink Budget-Friendly Options
6. Navigating Simple Algorithm Page Layout Analysis Springerlink eBook Formats
  - ePub, PDF, MOBI, and More
  - Simple Algorithm Page Layout Analysis Springerlink Compatibility with Devices
  - Simple Algorithm Page Layout Analysis Springerlink Enhanced eBook Features
7. Enhancing Your Reading Experience
  - Adjustable Fonts and Text Sizes of Simple Algorithm Page Layout Analysis Springerlink
  - Highlighting and Note-Taking Simple Algorithm Page Layout Analysis Springerlink
  - Interactive Elements Simple Algorithm Page Layout Analysis Springerlink
8. Staying Engaged with Simple Algorithm Page Layout Analysis Springerlink

- Joining Online Reading Communities
  - Participating in Virtual Book Clubs
  - Following Authors and Publishers [Simple Algorithm Page Layout Analysis Springerlink](#)
9. [Balancing eBooks and Physical Books Simple Algorithm Page Layout Analysis Springerlink](#)
- Benefits of a Digital Library
  - Creating a Diverse Reading Collection [Simple Algorithm Page Layout Analysis Springerlink](#)
10. [Overcoming Reading Challenges](#)
- Dealing with Digital Eye Strain
  - Minimizing Distractions
  - Managing Screen Time
11. [Cultivating a Reading Routine Simple Algorithm Page Layout Analysis Springerlink](#)
- Setting Reading Goals [Simple Algorithm Page Layout Analysis Springerlink](#)
  - Carving Out Dedicated Reading Time
12. [Sourcing Reliable Information of Simple Algorithm Page Layout Analysis Springerlink](#)
- Fact-Checking eBook Content of [Simple Algorithm Page Layout Analysis Springerlink](#)
  - Distinguishing Credible Sources
13. [Promoting Lifelong Learning](#)
- Utilizing eBooks for Skill Development
  - Exploring Educational eBooks
14. [Embracing eBook Trends](#)
- Integration of Multimedia Elements
  - Interactive and Gamified eBooks

### **Simple Algorithm Page Layout Analysis Springerlink Introduction**

Free PDF Books and Manuals for Download: Unlocking Knowledge at Your Fingertips In today's fast-paced digital age, obtaining valuable knowledge has become easier than ever. Thanks to the internet, a vast array of books and manuals are now available for free download in PDF format. Whether you are a student, professional, or simply an avid reader, this treasure trove of downloadable resources offers a wealth of information, conveniently accessible anytime, anywhere. The advent of online libraries and platforms dedicated to sharing knowledge has revolutionized the way we consume information. No longer confined to physical libraries or bookstores, readers can now access an extensive collection of digital books and

manuals with just a few clicks. These resources, available in PDF, Microsoft Word, and PowerPoint formats, cater to a wide range of interests, including literature, technology, science, history, and much more. One notable platform where you can explore and download free Simple Algorithm Page Layout Analysis Springerlink PDF books and manuals is the internet's largest free library. Hosted online, this catalog compiles a vast assortment of documents, making it a veritable goldmine of knowledge. With its easy-to-use website interface and customizable PDF generator, this platform offers a user-friendly experience, allowing individuals to effortlessly navigate and access the information they seek. The availability of free PDF books and manuals on this platform demonstrates its commitment to democratizing education and empowering individuals with the tools needed to succeed in their chosen fields. It allows anyone, regardless of their background or financial limitations, to expand their horizons and gain insights from experts in various disciplines. One of the most significant advantages of downloading PDF books and manuals lies in their portability. Unlike physical copies, digital books can be stored and carried on a single device, such as a tablet or smartphone, saving valuable space and weight. This convenience makes it possible for readers to have their entire library at their fingertips, whether they are commuting, traveling, or simply enjoying a lazy afternoon at home. Additionally, digital files are easily searchable, enabling readers to locate specific information within seconds. With a few keystrokes, users can search for keywords, topics, or phrases, making research and finding relevant information a breeze. This efficiency saves time and effort, streamlining the learning process and allowing individuals to focus on extracting the information they need. Furthermore, the availability of free PDF books and manuals fosters a culture of continuous learning. By removing financial barriers, more people can access educational resources and pursue lifelong learning, contributing to personal growth and professional development. This democratization of knowledge promotes intellectual curiosity and empowers individuals to become lifelong learners, promoting progress and innovation in various fields. It is worth noting that while accessing free Simple Algorithm Page Layout Analysis Springerlink PDF books and manuals is convenient and cost-effective, it is vital to respect copyright laws and intellectual property rights. Platforms offering free downloads often operate within legal boundaries, ensuring that the materials they provide are either in the public domain or authorized for distribution. By adhering to copyright laws, users can enjoy the benefits of free access to knowledge while supporting the authors and publishers who make these resources available. In conclusion, the availability of Simple Algorithm Page Layout Analysis Springerlink free PDF books and manuals for download has revolutionized the way we access and consume knowledge. With just a few clicks, individuals can explore a vast collection of resources across different disciplines, all free of charge. This accessibility empowers individuals to become lifelong learners, contributing to personal growth, professional development, and the advancement of society as a whole. So why not unlock a world of knowledge today? Start exploring the vast sea of free PDF books and manuals waiting to be discovered right at your fingertips.

## FAQs About Simple Algorithm Page Layout Analysis Springerlink Books

1. Where can I buy Simple Algorithm Page Layout Analysis Springerlink books? Bookstores: Physical bookstores like Barnes & Noble, Waterstones, and independent local stores. Online Retailers: Amazon, Book Depository, and various online bookstores offer a wide range of books in physical and digital formats.
2. What are the different book formats available? Hardcover: Sturdy and durable, usually more expensive. Paperback: Cheaper, lighter, and more portable than hardcovers. E-books: Digital books available for e-readers like Kindle or software like Apple Books, Kindle, and Google Play Books.
3. How do I choose a Simple Algorithm Page Layout Analysis Springerlink book to read? Genres: Consider the genre you enjoy (fiction, non-fiction, mystery, sci-fi, etc.). Recommendations: Ask friends, join book clubs, or explore online reviews and recommendations. Author: If you like a particular author, you might enjoy more of their work.
4. How do I take care of Simple Algorithm Page Layout Analysis Springerlink books? Storage: Keep them away from direct sunlight and in a dry environment. Handling: Avoid folding pages, use bookmarks, and handle them with clean hands. Cleaning: Gently dust the covers and pages occasionally.
5. Can I borrow books without buying them? Public Libraries: Local libraries offer a wide range of books for borrowing. Book Swaps: Community book exchanges or online platforms where people exchange books.
6. How can I track my reading progress or manage my book collection? Book Tracking Apps: Goodreads, LibraryThing, and Book Catalogue are popular apps for tracking your reading progress and managing book collections. Spreadsheets: You can create your own spreadsheet to track books read, ratings, and other details.
7. What are Simple Algorithm Page Layout Analysis Springerlink audiobooks, and where can I find them? Audiobooks: Audio recordings of books, perfect for listening while commuting or multitasking. Platforms: Audible, LibriVox, and Google Play Books offer a wide selection of audiobooks.
8. How do I support authors or the book industry? Buy Books: Purchase books from authors or independent bookstores. Reviews: Leave reviews on platforms like Goodreads or Amazon. Promotion: Share your favorite books on social media or recommend them to friends.
9. Are there book clubs or reading communities I can join? Local Clubs: Check for local book clubs in libraries or community centers. Online Communities: Platforms like Goodreads have virtual book clubs and discussion groups.
10. Can I read Simple Algorithm Page Layout Analysis Springerlink books for free? Public Domain Books: Many classic books are available for free as they're in the public domain. Free E-books: Some websites offer free e-books legally, like Project Gutenberg or Open Library.

**Find Simple Algorithm Page Layout Analysis Springerlink :**

[gibi do chaves](#)

[gothic novels of the twentieth century](#)

[grade 3 handwriting workbook mcgraw hill](#)

[glencoe algebra 2 chapter 3 resource masters](#)

[globalization east and west by bryan s turner](#)

[grade 12 life sciences learner notes](#)

**genetics from genes to genomes hartwell genetics**

[gli zii di sicilia](#)

[geography of jammu and kashmir by majid husain](#)

**geography grade 12 past exam papers spadós de**

[generator set synchronising with and running in parallel](#)

**geometry and discrete mathematics solutions**

[gestalt coaching right here right now](#)

[global oncology trends 2017 ims health](#)

**getting started in zbrush an introduction to digital sculpting and illustration**

**Simple Algorithm Page Layout Analysis Springerlink :**

Paraphrase on Dizzy Gillespie's "Manteca" : for two pianos, ... Paraphrase on Dizzy Gillespie's "Manteca" : for two pianos, op. 129. Authors: Nikolai Kapustin, Masahiro Kawakami (Editor), Dizzy Gillespie. Paraphrase on Dizzy Gillespie Manteca for two pianos, op. ... Paraphrase on Dizzy Gillespie Manteca for two pianos, op.129 - Kapustin, Nikolai - listen online, download, sheet music. PARAPHRASE ON DIZZY GILLESPIE'S MANTECA OP.129 ... MUST KAPUSTIN N. - PARAPHRASE ON DIZZY GILLESPIE'S MANTECA OP.129 - TWO PIANOS Classical sheets Piano. German edition. 4.4 4.4 out of 5 stars 2 reviews. MUST ... MUST KAPUSTIN N. - PARAPHRASE ON DIZZY ... MUST KAPUSTIN N. - PARAPHRASE ON DIZZY GILLESPIE'S MANTECA OP.129 - TWO PIANOS Classical sheets Piano - ISBN 10: 4904231562 - ISBN 13: 9784904231562 - MUST. PARAPHRASE ON DIZZY GILLESPIE'S MANTECA OP.129 ... MUST KAPUSTIN N. - PARAPHRASE ON DIZZY GILLESPIE'S MANTECA OP.129 - TWO PIANOS Classical sheets Piano. German edition. 4.4 4.4 out of 5 stars 2 Reviews. MUST ... Paraphrase On Dizzy Gillespie's Manteca Sheet Music - £37.95 - Nikolaj Girshevich Kapustin - Paraphrase On Dizzy Gillespie's Manteca. ... Piano, Keyboard & Organ - Piano Solo. Publisher: MusT Music ... Classical and Jazz Influences in the

Music of Nikolai Kapustin by Y Tyulkova · 2015 · Cited by 8 — The topic of this research is the contemporary Russian composer and pianist Nikolai. Kapustin. This paper will focus on the influences from both Classical and ... Kinetic and Potential Energy Worksheet KEY  $g=9.8$  Calculate it. 21. Determine the kinetic energy of a 1000-kg roller coaster car that is moving with a speed of 20.0 m/s. 22. KINETIC AND POTENTIAL ENERGY WORKSHEET Answer the following: a. What is the kinetic energy of a 1-kilogram ball is thrown into the air with an initial velocity of 30 m/sec?  $KE = \frac{1}{2} m v^2$   $\frac{1}{2} (1 \text{ kg}) \dots$  Kinetic Energy (KE) =  $\frac{1}{2}$  mass times velocity squared Potential and Kinetic Energy Worksheet. Kinetic Energy (KE) =  $\frac{1}{2}$  mass times velocity squared.  $KE = \frac{1}{2} m v^2$ . Potential Energy (PE) = mass times the acceleration ... Kinetic and potential energy worksheet answer key k o myaiu kinetic and potential energy worksheet classify the following as type of potential energy or kinetic energy (use the letters or bicyclist pedaling up ... Kinetic and Potential Energy Worksheet Walkthrough - YouTube kinetic and potential energy worksheet Flashcards A. How much kinetic energy does the ball have? B. How much potential energy does the ball have when it reaches the top of the ascent? KINETIC AND POTENTIAL ENERGY WORKSHEET Answer the following: a. What is the kinetic energy of a 1-kilogram ball is thrown into the air with an initial velocity of 30 m/sec? Kinetic vs Potential Energy Practice KEY Page 1. Scanned by CamScanner. Page 2. Scanned by CamScanner. Potential and kinetic energy worksheet and answer key This easy to read, one page passage about potential energy :explains potential energy as stored energy gives examples such as a car ... The Chips Are Down (screenplay) The Chips Are Down is a screenplay written by Jean-Paul Sartre in 1943 and published in 1947. The original title translates literally as "the plays are ... The Chips Are Down (Les Jeux Sont Faits) Amazon.com: The Chips Are Down (Les Jeux Sont Faits): Jean-Paul Sartre, Louise Varese: Movies & TV. ... The Chips Are Down (Les Jeux Sont Faits). 4.7 4.7 out of 5 ... The Chips are Down by Sartre The Chips Are Down (Les Jeux Sont Faits) by Jean-Paul Sartre and a great selection of related books, art and collectibles available now at AbeBooks.com. The chips are down =: Les jeux sont faits: Sartre, Jean Paul The chips are down =: Les jeux sont faits [Sartre, Jean Paul] on Amazon ... Jean-Paul Sartre. 4.5 out of 5 stars 80. Paperback. 48 offers from \$2.04. Explore ... The Chips are Down - Jean-Paul Sartre The story is set in Paris, in a setting vaguely suggestive of German-occupied northern France (or perhaps Vichy France) during World War II. The plot concerns ... The Chips are Down | Jean-Paul SARTRE Hardcover. A novel by Sartre translated from the French by Louise Varese. The basis for a French movie with Micheline prsle and Marcel Pagliero. A clean very ... The chips are down Screenplay written by Jean-Paul Sartre in 1943 and published in 1947. The original title translates literally as "The Plays are Made", an idiomatic French ... Jean-Paul Sartre First Edition The Chips Are Down First US edition of the tragicomic screenplay "The Chips Are Down" by French philosopher Jean-Paul Sartre, adapted from "Les Jeux Sont Faits". Les jeux sont faits by Jean-Paul Sartre The Chips Are Down is a screenplay written by Jean-Paul Sartre in 1943 and published in 1947. Ève and Pierre have never met each other in their respective lives ... The Chips Are Down "The Chips Are Down" is a French idiom used in cards, roughly meaning 'the plays are made'. It may also refer to: The

Chips Are Down (screenplay) (Les jeux ...