

Document Page Layout Analysis

Bhabatosh Chanda

Electronics and Communication Sciences Unit

Indian Statistical Institute

Kolkata 700108, India

Acknowledgement

- Amit Das, IEST, Sibpur
- Sekhar Mandal, IEST, Sibpur
- Sanjoy Kumar Saha, Jadavpur Univeristy
- Ranjan Mandal, Indian Statistical Institute

Outline

- Introduction
- Projection method
 - Zone content classification
- Morphological operators
 - Skew correction
- Morphology based method
- Deep learning based method
- Performance evaluation
- Database: examples
- Conclusion

Introduction

- Problem description
- Motivation
 - Improve performance of OCR
 - Data compression
 - Graphics recognition
 - Browsing and navigation
- Physical and logical structure

Problem Description




Figure 1: Example of figure inclusion.

This page is used to develop and algorithm for document page image segmentation. Page should contain mathematical equations, tables, section headings, graphics or line drawing and half-tone images alongwith regular text. The page is generated using LaTeX word processing system. Then the page is printed on a HP-5MP laserjet printer, followed by scanned in by HP 4C scanner. Scanned in image may have till.

1 Math equations

Below we give an example of mathematical equation computing mean of a set of data points followed by the formula for computing variances of the same. Equations with or without equation number are given.

$$\bar{x} = \frac{1}{n} \sum_{i=0}^n x_i$$

Following one is an example of equation array with equation number. However, other style of equations are not considered here because of lack

of space. So the following equation is the example of the second type that is equation with number.

$$\begin{aligned} \sigma_x &= \frac{1}{n} \sum_{i=0}^n (x_i - \bar{x})^2 \\ &= \frac{1}{n} \sum_{i=0}^n x_i^2 - \bar{x}^2 \end{aligned} \quad (1)$$

Now we give an example of table that has column and row separators as well as captions and number. Finally we put a line drawing which is basically a normal probability density function. Hence this page illustrates all the basic entities we like to deal with in a document page: image segmentation. Here is the end of this page.

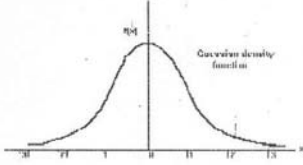


Figure 2: Example of Graphics inclusion.

Thanks for your attention. Resolution of scanning is 100 dpi (i.e., dots per inch).

Roll	Lit.	Sci.	Soc.	Math.
1	38	95	67	85
23	67	83	19	78
43	17	93	13	92
67	19	16	34	56

Objective

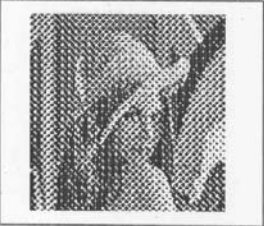


Figure 1: Example of figure inclusion.

Roll	Lit.	Sci.	Soc.	Math.
1	38	95	67	85
23	67	83	19	78
43	17	93	13	92
67	19	16	34	56

This page is used to develop an algorithm for document page image regeneration. Page should contain mathematical equations, tables, section headings, graphics or line drawing and half-tone images alongwith regular text. The page is generated using LaTeX word processing system. Then the page is printed on a HP-5MP laserjet printer, followed by scanned in by HP 4C scanner. Scanned in image may have fill.

of space. So the following equation in the example of the second type that is equation with number.

$$\begin{aligned} \sigma_w &= \frac{1}{n} \sum_{i=0}^n (x_i - \bar{x})^2 \\ &= \frac{1}{n} \sum_{i=0}^n x_i^2 - \bar{x}^2 \end{aligned} \quad (1)$$

Now we give an example of table that has column and row separators as well as captions and number. Finally we put a line drawing which is basically a normal probability density function. Hence this page illustrates all the basic entities we like to deal with in a document page: image regeneration. Here is the end of this page.

1 Math equations

Below we give an example of mathematical equation computing mean of a set of data points followed by the formula for computing variances of the same. Equations with or without equation number are given.

$$\bar{x} = \frac{1}{n} \sum_{i=0}^n x_i$$

Following one is an example of equation array with equation number. However, other style of equations are not considered here because of lack

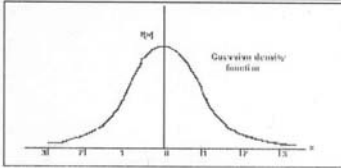
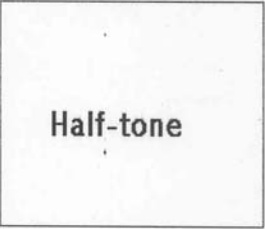


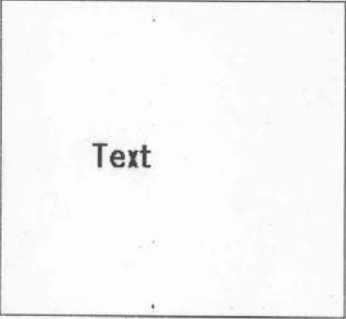
Figure 2: Example of Graphics inclusion.

Thanks for your attention. Resolution of scanning is 100 dpi (i.e., dots per inch).

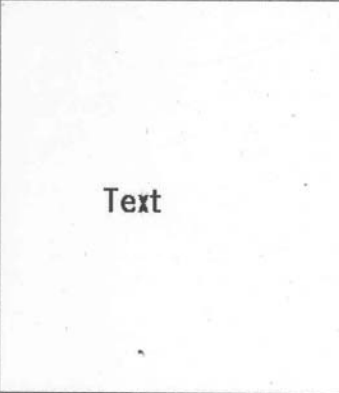


Half-tone

Text
Table

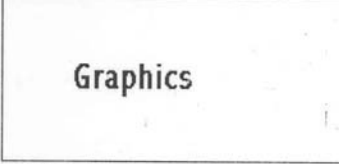


Text

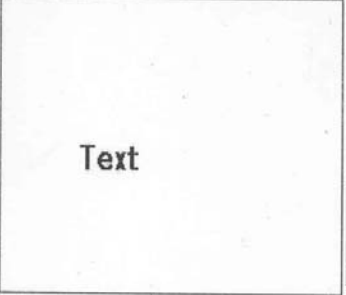


Text

Heading
Text



Graphics



Text

Text
Text

Major Source of Document Pages

1. Books
2. Journals
3. Magazines
4. Newspapers
5. Forms and leaflets
6. Reports

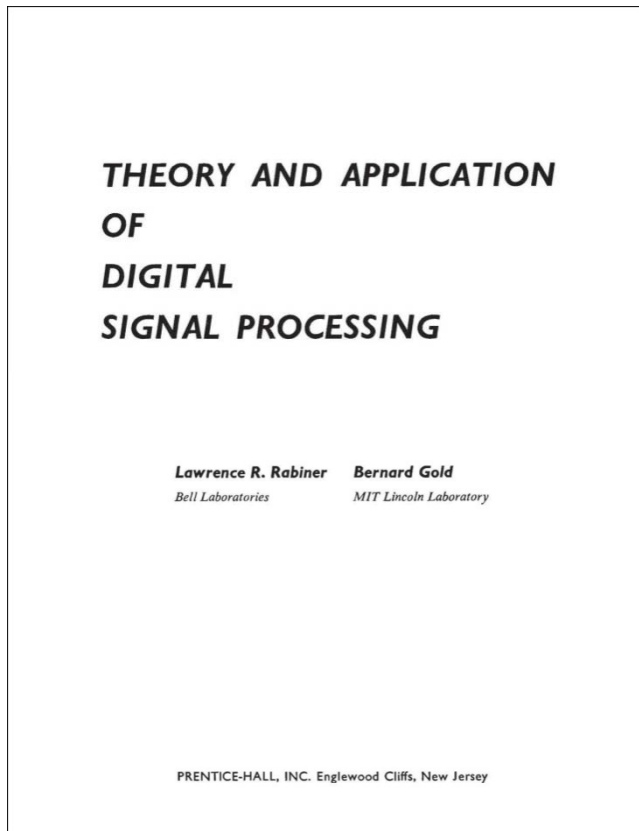
Types of document pages

Consider books and journals

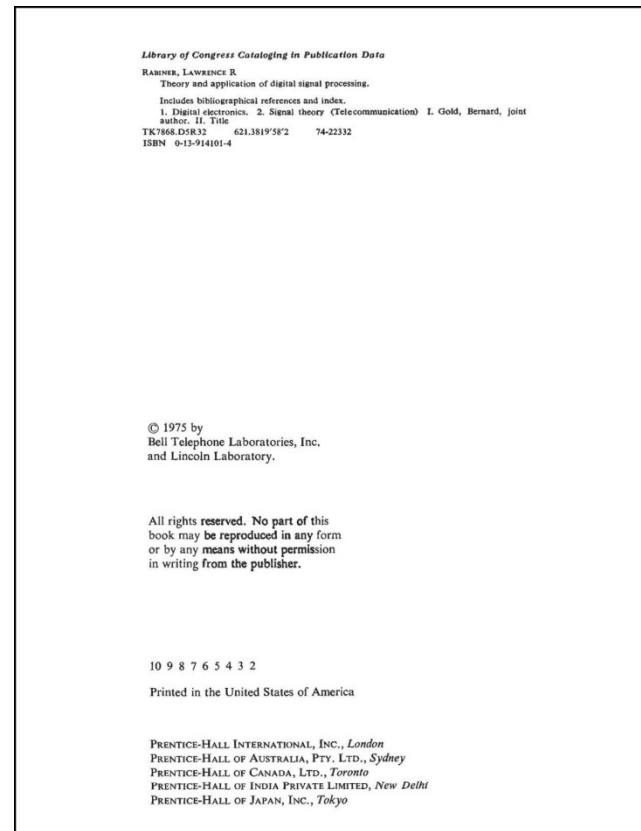
- Title page
- Publisher's page
- Table of Contents
- Text page
- Index page

Different types of pages

Title page



Publisher's page



Different types of pages

Table of Content page

vi	Contents	
2.13	The z Transform	28
2.14	Relation Between the z Transform and the Fourier Transform of a Sequence	30
2.15	The Inverse z Transform	32
2.16	Properties of the z Transform	33
2.17	Solution of Difference Equations Using the One-Sided z Transform	37
2.18	Geometric Evaluation of the Fourier Transform	39
2.19	Digital Filter Realizations (Structures)	40
2.20	Structures For All-Zero Filters	46
2.21	The Discrete Fourier Transform	50
2.22	Properties of the DFT	57
2.23	Convolution of Sequences	59
2.24	Linear Convolution of Finite Duration Sequences	61
2.25	Sectioned Convolutions	63
2.26	The Discrete Hilbert Transform	67
2.27	Hilbert-Transform Relations for Real Signals	70
	References	73
3 THE THEORY AND APPROXIMATION OF FINITE DURATION IMPULSE RESPONSE DIGITAL FILTERS		
		75
3.1	Introduction	75
3.2	Issues in Filter Design	75
3.3	Discussion of FIR Filters	76
3.4	Characteristics of FIR Filters with Linear Phase	77
3.5	Frequency Response of Linear Phase FIR Filters	81
3.6	Positions of the Zeros of Linear Phase FIR Filters	84
3.7	Design Techniques for Linear Phase FIR Filters	88
3.8	Design Technique No. 1—Windowing	88
3.9	Rectangular Window	90
3.10	“Generalized” Hamming Window	91
3.11	Kaiser Window	93
3.12	Examples of a Windowed Lowpass Filter	94
3.13	Issues with Windowing	101
3.14	Some Practical Techniques with Windows	102
3.15	Additional Examples of Window Designed Filters	103
3.16	Summary of Windows	105
3.17	Design Technique No. 2—Frequency Sampling	105
3.18	Solution for Optimization	108
3.19	Linear Programming	110
3.20	Types 1 and 2	112
3.21	Type 1 Designs—Linear Phase Constraints	113
3.22	Type 2 Designs—Linear Phase Constraints	115
3.23	Some General Results on the Design of Frequency Sampling Filters	117
3.24	Summary of Frequency Sampling Design	121
3.25	Design Technique No. 3—Optimal (Minimax Error) Filters	123

Table of Content page

	Contents	
	Preface	xxix
<hr/>		
Part I	The Java Language	
1	The History and Evolution of Java	3
	Java's Lineage	3
	The Birth of Modern Programming: C	4
	C++: The Next Step	5
	The Stage Is Set for Java	6
	The Creation of Java	6
	The C# Connection	8
	How Java Changed the Internet	8
	Java Applets	8
	Security	9
	Portability	9
	Java's Magic: The Bytecode	9
	Servlets: Java on the Server Side	10
	The Java Buzzwords	10
	Simple	11
	Object-Oriented	11
	Robust	11
	Multithreaded	12
	Architecture-Neutral	12
	Interpreted and High Performance	12
	Distributed	12
	Dynamic	13
	The Evolution of Java	13
	Java SE 6	14
	A Culture of Innovation	14
2	An Overview of Java	15
	Object-Oriented Programming	15
	Two Paradigms	15
	Abstraction	16
	The Three OOP Principles	16
	A First Simple Program	21
	Entering the Program	21
	Compiling the Program	22
	A Closer Look at the First Sample Program	22
		vii

Different types of pages

Text page-1

20 Theory of Discrete-Time Linear Systems

Fig. 2.8 Realization of a second-order difference equation.

can be realized as shown in Fig. 2.8. Again separate delays are used for both the input and output sequences.

As seen later in this chapter, the specific cases of first- and second-order systems are very important in the realization of higher-order systems because these higher-order systems can be decomposed into a cascade or a parallel combination of first- and second-order systems.

2.7 Frequency Response

Until now we have been concerned with the response of LTI systems to arbitrary input sequences. In this section we restrict ourselves to a special class of inputs, $x(n) = e^{j\omega n}$, in order to investigate the frequency domain representation of LTI systems. As will be shown, this class of inputs is the set of eigenfunctions of discrete-time, LTI systems; i.e., the output sequence is the input sequence multiplied by a complex weighting factor that depends solely on ω .

Consider a class of input sequences of the form

$$x(n) = e^{j\omega n} \quad -\infty < n < \infty \quad (2.29)$$

If this input is applied to an LTI system with impulse response $h(n)$, then from Eq. (2.11a) the output is

$$y(n) = \sum_{m=-\infty}^{\infty} h(m)e^{j\omega(n-m)} \quad (2.30)$$

$$= e^{j\omega n} \sum_{m=-\infty}^{\infty} h(m)e^{-j\omega m} \quad (2.31)$$

$$= x(n)H(e^{j\omega}) \quad (2.32)$$

Thus for this special class of inputs, we see from Eq. (2.32) that the output is identical to the input to within a complex multiplier $H(e^{j\omega})$, which is defined from the impulse response as

$$H(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h(n)e^{-j\omega n} \quad (2.33)$$

Text page-2

LI et al.: EXPANDABLE DATA-DRIVEN GRAPHICAL MODELING OF HUMAN ACTIONS BASED ON SALIENT POSTURES 1807

Fig. 7. Sample silhouette sequences from Lapedis's data set: (a) and (b) noisy; (c) and (d) run.

TABLE II
CROSS-DATA-SET TEST: RECOGNITION ERRORS (OUT OF 68) VERSUS NUMBER OF POSTURES FOR BMLD, BGVD, AND ASVD

# of Postures/Decoding Method	28	30	45	60
BMLD	11	5	1	0
BGVD	11	8	1	0
ASVD	14	7	5	3

different subjects. Fig. 7 shows a few examples of the extracted silhouettes. It can be seen that the silhouettes are noisy and, in Fig. 7(d), the subject wore a trench coat that distorted the silhouette shape. Table II is the number of recognition errors (out of 68) versus number of postures. As seen, the recognition rates are over 95% for BMLD, BGVD, and ASVD when the number of postures is 60. Notice that BMLD and BGVD performed better than ASVD. This is probably because ASVD is less generalized than BMLD and BGVD.

3) *Learning New Actions:* With respect to learning new actions, we first evaluated the significance of *smoothing*. Fig. 8(a) shows the recognition errors for the cases of sharing postures versus not sharing postures and smoothing versus not smoothing when the number of training samples for the new action is one. In sharing, we forced the algorithm not to create any new postures. In the case of not sharing, the algorithm was forced to create three new postures specifically for the new action. In each test, one sample of the action was used as training sample and the rest samples of the same action were used as test samples. The errors showed in the figure were averaged over all actions and all samples in each action. It is apparent that sharing and smoothing significantly reduced the recognition errors and are essential to learning a new action. Notice that, in the case of not sharing, the ASVD scheme is equivalent to the conventional methods where the model for each action is trained independently. It is obvious that our method outperforms the conventional ones.

Fig. 8(b) is the recognition errors of the added new action against the number of training samples. Surprisingly, the BMLD constantly outperformed BGVD and ASVD. On average, we achieved over 85% recognition rate for the new action even though there were only three to four training samples. When the number of training samples reached eight, the recognition rate was improved to over 95%.

We also evaluated the impact on the recognition of previously learned actions when a new action was added. We trained a system by leaving one action out and tested the trained system against the training samples at $M = 30$. In all cases, the training samples were recognized without any error. We then added the left-out action to the system using the proposed method. The

Fig. 8. Learning new actions. (a) Study on the importance of sharing and smoothing. (b) Overall recognition rates of new actions versus number of postures. (c) Impact on the existing system when a new action is added. (d) A typical case that a new posture (S30) was created whenever it was added to the system as a new action.

Authorized licensed use limited to: MICROSOFT. Downloaded on March 25, 2009 at 17:11 from IEEE Xplore. Restrictions apply.

Different types of pages

Text page-3

Image and Vision Computing 28 (2010) 976–990

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

A survey on vision-based human action recognition

Ronald Poppe^{*}

Human Media Interaction Group, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, PO Box 311, 7500 AE Enschede, The Netherlands

ARTICLE INFO

Article history:
Received 20 February 2009
Received in revised form 5 September 2009
Accepted 30 November 2009

ABSTRACT

Vision-based human action recognition is the process of labeling image sequences with action labels. Robust solutions to this problem have applications in domains such as visual surveillance, video retrieval and human-computer interaction. The task is challenging due to variations in motion performance, recording settings and inter-personal differences. In this survey, we explicitly address these challenges. We provide a detailed overview of current advances in the field. Image representations and the subsequent classification process are discussed separately to focus on the novelties of recent research. Moreover, we discuss limitations of the state of the art and outline promising directions of research.

Keywords:
Human action recognition
Motion analysis
Action detection

1. Introduction

We consider the task of labeling videos containing human motion with action classes. The interest in the topic is motivated by the promise of many applications, both offline and online. Automatic annotation of video enables more efficient searching for example finding tackles in soccer matches, handshakes in news footage or typical dance moves in music videos. Online processing allows for automatic surveillance, for example in shopping malls, but also in smart homes for the elderly to support aging in place. Interactive applications, for example in human-computer interaction or games, also benefit from the advances in automatic human action recognition.

In this section, we first discuss related surveys and present the scope of this overview. Also, we outline the main characteristics and challenges of the field at these motivate the various approaches that are reported in literature. Finally, we briefly describe the most common datasets. In its simplest form, vision-based human action recognition can be regarded as a combination of feature extraction, and subsequent classification of these image representations. We discuss these two tasks in Sections 2 and 3, respectively. While many works will be described and analyzed in more detail, we do not intend to give complete coverage of all works in the area. In Section 4, we discuss limitations of the state of the art and outline future directions to address these.

1.1. Scope of this overview

The area of human action recognition is closely related to other lines of research that analyze human motion from images and

video. The recognition of movement can be performed at various levels of abstraction. Different taxonomies have been proposed and here we adopt the hierarchy used by Moeslund et al. [90]: action primitive, action and activity. An action primitive is an atomic movement that can be described at the limb level. An action consists of action primitives and describes a possibly cyclic, whole-body movement. Finally, activities contain a number of subsequent actions, and give an interpretation of the movement that is being performed. For example, "left leg forward" is an action primitive, whereas "running" is an action. "Jumping hurdles" is an activity that contains starting, jumping and running actions.

We focus on actions and do not explicitly consider context such as the environment (e.g. [119]), interactions between persons (e.g. [105,122]) or objects (e.g. [47,91]). Moreover, we consider only full-body movements, which excludes the work on gesture recognition (see [30,88]).

In the field of gait recognition, the focus is on identifying personal styles of walking movement, to be used as a biometric cue. The aim of human action recognition is opposite: to generalize over these variations. This is an arbitrary process as there is often significant intra-class variation. Recently, there have been several approaches that aim at simultaneous recognition of both action, and style (e.g. [22,28,152]). In this overview, we will discuss mainly those approaches that can deal with a variety of actions.

1.2. Surveys and taxonomies

There are several existing surveys within the area of vision-based human motion analysis and recognition. Recent overviews by Forsyth et al. [38] and Poppe [109] focus on the recovery of human poses and motion from image sequences. This can be regarded as a regression problem, whereas human action recognition is a

© 2009 Elsevier B.V. All rights reserved.

Index page

760 Index

Not-in-place algorithms, 575
Number theoretic transforms, 419–33

O

One-bit adder, 493–95
One's-complement, 302
One-sided z transform, 36–37
Optimal (minimax error) design, 123–83
Optimization design of IIR filters, 267–83
Optimization methods:
 frequency sampling design, 107–9
Ordering, 323–25
OR function, 491
Overflow oscillations, 328
Overlap-add method, 63–65
Overlap-save method, 65–67

P

Parseval expansion, 266
Pairing, 323–25
Parallel form, 45
Parseval equations, 36
Partial fraction expansion, 33, 45
Particular solution, 18–19
Peak scaling, 333
Periodic convolution, 59–60
Periodogram, 415
Phase delay, 80
Phase response, 210
Picture processing, 478–82
Pipeline FFTs, 598, 602–14
Pitch detection, 681–87
Pitch period, 663
P-MOS (P-channel MOS), 504
Poles, 29
Pole-zero pairing, 44
Pole-zero plot, 31
Post filter, 301
Power spectral density, 401
Power spectrum analysis, 399–419
Printed circuit construction, 487
Prolate spheroidal wavefunctions, 93

PROM (programmable ROM), 513
Pseudo-random number generators, 565–71

Q

Quantized coefficients:
 optimization methods, 341–44
Quantizing noise, 296

R

Radar applications, 709–53
Radar environments, 711
Radar parameters, 711–15
Radar principles, 709–11
Radar pulse compression, 747–48
Radar signal design, 715–24
Radix 2 FFTs, 357–71, 379–81, 594–98
RAM (random access memory), 513
Random access bipolar memories, 513
Range and range resolution, 713
Rayleigh distributed random numbers, 571
Realizability (see Casuality)
Reconstruction filter, 301
Rectangular window, 90–91
Recursive realization, 40
Remez exchange algorithm, 136–40, 227
Residue theorem, 33
Ripple, 90
ROM (read-only memories), 513
Rotation factors, 363
Rounding, 296–97, 308–9
Roundoff noise:
 cascade form, 323
 cascade form nonrecursive structure, 331–34
 direct form nonrecursive structure, 328–30
 direct form recursive structure, 320–21
 dynamic range constraints, 315–28
 L_p norm, 318–20
 nonrecursive structures:
 fixed point case, 328–34
 parallel form, 321–22

Issues in document page scanning

- Resolution
- Back page impression
- Granular noise
- Blotted text (*specially in old documents*)
- Bending of pages at the binding
- Skew
(due to placement of the page in the scanner)

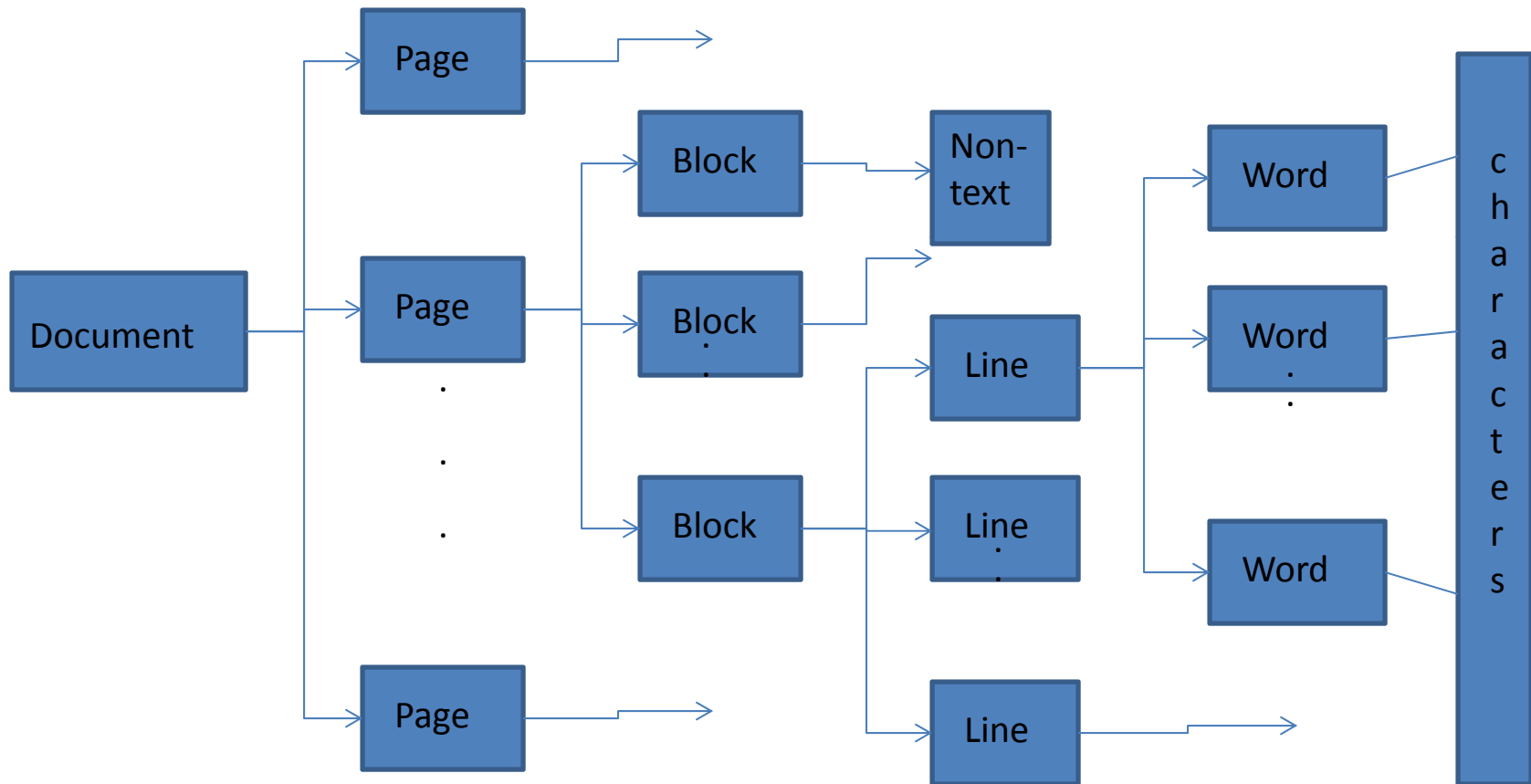
Entities of Document Page

- Text
 - Body text
 - Line → Word → Character
 - Heading
- Non-text
 - Half-tone
 - Table
 - Graphics or line drawing

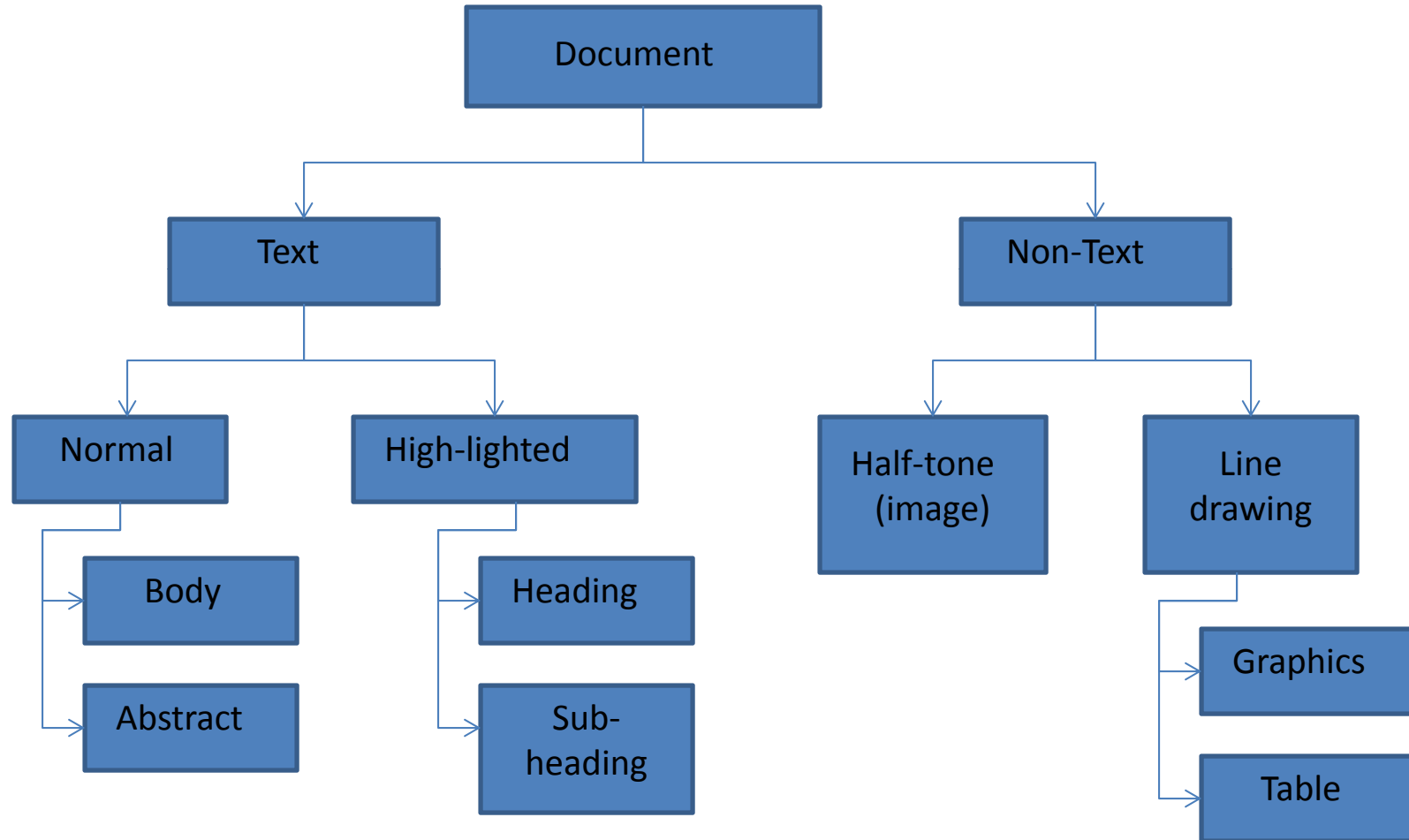
Entities of Document Page

- Each detected zone or block must be homogeneous in terms of content or entity
- Each zone will be input to one of the suitable modules based on entity.
 - OCR system
 - Image compressor
 - Vectorization system
- Output of these modules may be compiled and archived using suitable structure.

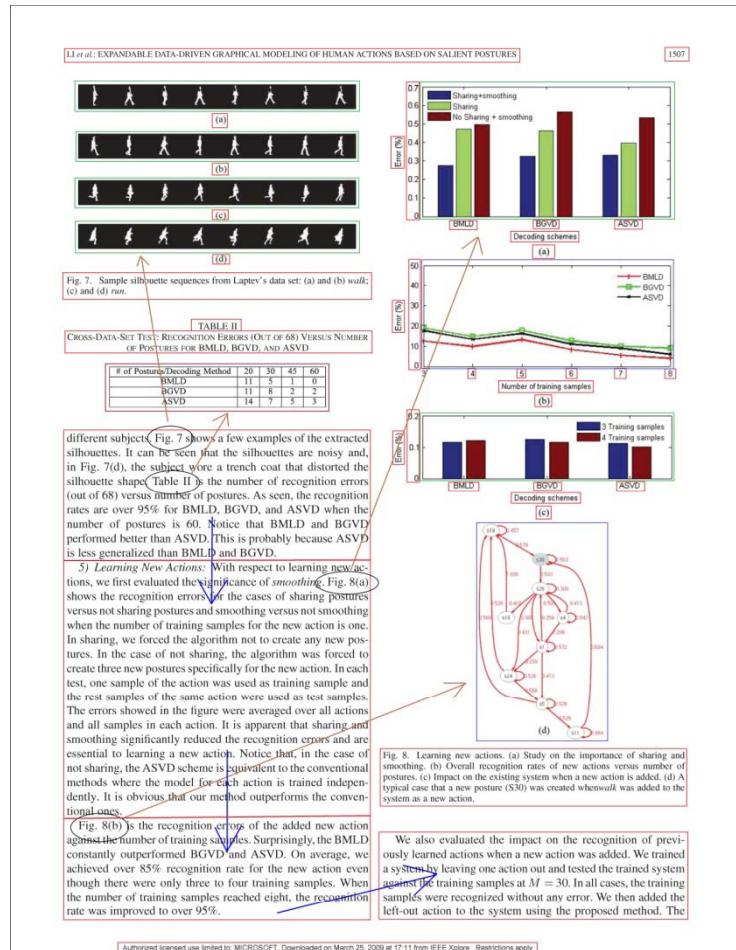
Geometrical / Physical structure



Logical structure



Logical structure



- Different entities:
 - Text (red box)
 - Halftone (green box)
 - Table (magenta box)
 - Line drawing (blue box)
- Reading direction (dark blue arrow)
- Link between entities (brown arrow)

Zone / block detection

- One of the simple way is ***Projection method***.
- Algorithm
 - Take horizontal (or vertical) projection of foreground pixels. (may be implemented as pixel count)
 - If there exists a characteristic change in projection profile, put a horizontal (resp. vertical) separator.
 - Take horizontal and vertical direction alternately.
 - Continue, until above condition is satisfied.
- Works well for ***structured document***, usually the pages of technical journals, books, etc.

Projection Method: An Example

তিসেখর থেকে এখিল। সুতির পোশাক ফর্মের। সঙ্গে ছাড়া নিয়ে গেলে ছাড়া হয়। নাতলমুখে প্রথম দুখপ্রসঙ্গে মনেমেটেই বুঝি হয়েছে।

কলকাতা-পোর্ট প্রেমার জাহাজভাড়া

এম ডি	এম ডি	এম ডি
জাহাজ	বেইলি	মিলেজ
বিশাল প্রেমার	৩৪৪১	৩৪৪১
মুর্শী প্রেমার	৩৪২১	৩৪২১
সেকেন্ড প্রেমার (এ)	২৪৪৪	২৪৪৪
সেকেন্ড প্রেমার (বি)	১৭২২	১৭২২
মুর্শী প্রেমার	২৪৪১	২৪৪১
মুর্শী প্রেমার	৩৪২১	৩৪২১

এক প্রেমার হাজার কলম্বা কলম্বা কলম্বা কলম্বা কলম্বা ১০০ টাকার প্রিমিয়াম। কলম্বা কলম্বা কলম্বা কলম্বা ১০ টাকার প্রিমিয়াম।

কোথায় থাকবেন
 নানা মানের নানা মানের পেনসনরি কলম্বা কলম্বা আছে পোর্ট প্রেমার ও তার আশেপাশে। কলম্বা কলম্বা আছে কলম্বা কলম্বা প্রেমার হোটেল বে আইসিআর (৩ ৩৩৩১১) ও পিটার্সনস রিসর্ট (৩ ৩৩৩২২), সিন্ডিকেশন বে ডিউ (পোর্ট প্রেমার, ৩ ৩৩৩৩৩) ও হোটেল শাশনি (মিলেজ পোর্ট, ৩ ৩৩৩৪৪) এর ঘরভাড়া ৩০০-১৪০০ টাকার মধ্যে। হোটেল অফিসের (গোবর্ধন, ৩ ৩৩৩৫৫), হোটেল এন কে ইন্টারন্যাশনাল (সি পোর প্রেমার, ৩ ৩৩৩৬৬), হোটেল শালিমার (গোবর্ধন, ৩ ৩৩৩৭৭) ইত্যাদি হোটেল ঘর মেসে ১০০-৪০০ টাকায়। আশেপাশী ঘরভাড়া আছে বেশ কিছু হোটেল।

হোটেল কলম্বা কলম্বা কলম্বা কলম্বা কলম্বা ৩০০ টাকার। হোটেল কলম্বা কলম্বা কলম্বা কলম্বা কলম্বা ২০-২০ টাকার ঘর মেসে। এছাড়া আছে চলিত রিসর্ট (৩ ৩৩৩১১), হোটেল শাশনি (৩ ৩৩৩২২), জাহাজ পোর্ট হাউস (৩ ৩৩৩৩৩) ইত্যাদি।

আনন্দামান পর্যটন বিভাগের হোটেল

আনন্দামান টোল হাউস	বিলম্ব	এম ডি
লেগনিসিউ	৩৪৪১	৩৪৪১
হোটেল কেট	৩৪২১	৩৪২১
কলম্বা কলম্বা	৩৪২১	৩৪২১
হাওয়াইল কেট	৩৪২১	৩৪২১
মিল হাউস	৩৪২১	৩৪২১
কলম্বা কলম্বা	৩৪২১	৩৪২১
হাওয়াইল কেট	৩৪২১	৩৪২১
কলম্বা কলম্বা	৩৪২১	৩৪২১

Director of Tourism (A & N)
 Port Blair-744 101
 ৩ (৩৩১৯২) ৩৩৩৩৩

তিসেখর থেকে এখিল। সুতির পোশাক ফর্মের। সঙ্গে ছাড়া নিয়ে গেলে ছাড়া হয়। নাতলমুখে প্রথম দুখপ্রসঙ্গে মনেমেটেই বুঝি হয়েছে।

কলকাতা-পোর্ট প্রেমার জাহাজভাড়া

এম ডি	এম ডি	এম ডি
জাহাজ	বেইলি	মিলেজ
বিশাল প্রেমার	৩৪৪১	৩৪৪১
মুর্শী প্রেমার	৩৪২১	৩৪২১
সেকেন্ড প্রেমার (এ)	২৪৪৪	২৪৪৪
সেকেন্ড প্রেমার (বি)	১৭২২	১৭২২
মুর্শী প্রেমার	২৪৪১	২৪৪১
মুর্শী প্রেমার	৩৪২১	৩৪২১

এক প্রেমার হাজার কলম্বা কলম্বা কলম্বা কলম্বা কলম্বা ১০০ টাকার প্রিমিয়াম। কলম্বা কলম্বা কলম্বা কলম্বা ১০ টাকার প্রিমিয়াম।

কোথায় থাকবেন
 নানা মানের নানা মানের পেনসনরি কলম্বা কলম্বা আছে পোর্ট প্রেমার ও তার আশেপাশে। কলম্বা কলম্বা আছে কলম্বা কলম্বা প্রেমার হোটেল বে আইসিআর (৩ ৩৩৩১১) ও পিটার্সনস রিসর্ট (৩ ৩৩৩২২), সিন্ডিকেশন বে ডিউ (পোর্ট প্রেমার, ৩ ৩৩৩৩৩) ও হোটেল শাশনি (মিলেজ পোর্ট, ৩ ৩৩৩৪৪) এর ঘরভাড়া ৩০০-১৪০০ টাকার মধ্যে। হোটেল অফিসের (গোবর্ধন, ৩ ৩৩৩৫৫), হোটেল এন কে ইন্টারন্যাশনাল (সি পোর প্রেমার, ৩ ৩৩৩৬৬), হোটেল শালিমার (গোবর্ধন, ৩ ৩৩৩৭৭) ইত্যাদি হোটেল ঘর মেসে ১০০-৪০০ টাকায়। আশেপাশী ঘরভাড়া আছে বেশ কিছু হোটেল।

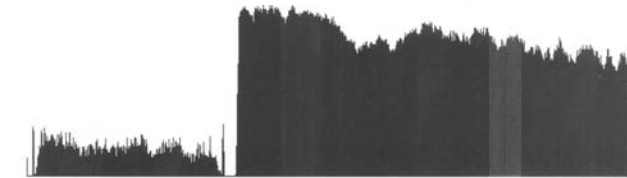
হোটেল কলম্বা কলম্বা কলম্বা কলম্বা কলম্বা ৩০০ টাকার। হোটেল কলম্বা কলম্বা কলম্বা কলম্বা কলম্বা ২০-২০ টাকার ঘর মেসে। এছাড়া আছে চলিত রিসর্ট (৩ ৩৩৩১১), হোটেল শাশনি (৩ ৩৩৩২২), জাহাজ পোর্ট হাউস (৩ ৩৩৩৩৩) ইত্যাদি।

আনন্দামান পর্যটন বিভাগের হোটেল

আনন্দামান টোল হাউস	বিলম্ব	এম ডি
লেগনিসিউ	৩৪৪১	৩৪৪১
হোটেল কেট	৩৪২১	৩৪২১
কলম্বা কলম্বা	৩৪২১	৩৪২১
হাওয়াইল কেট	৩৪২১	৩৪২১
মিল হাউস	৩৪২১	৩৪২১
কলম্বা কলম্বা	৩৪২১	৩৪২১
হাওয়াইল কেট	৩৪২১	৩৪২১
কলম্বা কলম্বা	৩৪২১	৩৪২১

Director of Tourism (A & N)
 Port Blair-744 101
 ৩ (৩৩১৯২) ৩৩৩৩৩

Example (contd.)



বিদেশের থেকে এটি... (Text describing the image)

কনকাতো-পোর্ট প্রোফের জাহাজসমূহ

বিভাগ	জাহাজ	বিশেষ
কনকাতো	১৪৫১	১৪৫২
পোর্ট প্রোফ	১৪৫৩	১৪৫৪
কনকাতো-পোর্ট প্রোফ	১৪৫৫	১৪৫৬
কনকাতো-পোর্ট প্রোফ	১৪৫৭	১৪৫৮
কনকাতো	১৪৫৯	১৪৬০
পোর্ট প্রোফ	১৪৬১	১৪৬২

এই প্রোগ্রামের উদ্দেশ্য... (Text describing the program)

সিদ্ধান্ত নেওয়া

১৯৬০-৬১ সালের মধ্যে... (Text describing the data)

আনামান পোর্ট বিজ্ঞানের হোটেল

বিভাগ	জাহাজ	বিশেষ
আনামান	১৪৬৩	১৪৬৪
পোর্ট প্রোফ	১৪৬৫	১৪৬৬
আনামান-পোর্ট প্রোফ	১৪৬৭	১৪৬৮
আনামান-পোর্ট প্রোফ	১৪৬৯	১৪৭০
আনামান	১৪৭১	১৪৭২
পোর্ট প্রোফ	১৪৭৩	১৪৭৪

সিদ্ধান্ত নেওয়া

বিদেশের থেকে এটি... (Text describing the image)

কনকাতো-পোর্ট প্রোফের জাহাজসমূহ

বিভাগ	জাহাজ	বিশেষ
কনকাতো	১৪৫১	১৪৫২
পোর্ট প্রোফ	১৪৫৩	১৪৫৪
কনকাতো-পোর্ট প্রোফ	১৪৫৫	১৪৫৬
কনকাতো-পোর্ট প্রোফ	১৪৫৭	১৪৫৮
কনকাতো	১৪৫৯	১৪৬০
পোর্ট প্রোফ	১৪৬১	১৪৬২

এই প্রোগ্রামের উদ্দেশ্য... (Text describing the program)

সিদ্ধান্ত নেওয়া

১৯৬০-৬১ সালের মধ্যে... (Text describing the data)

আনামান পোর্ট বিজ্ঞানের হোটেল

বিভাগ	জাহাজ	বিশেষ
আনামান	১৪৬৩	১৪৬৪
পোর্ট প্রোফ	১৪৬৫	১৪৬৬
আনামান-পোর্ট প্রোফ	১৪৬৭	১৪৬৮
আনামান-পোর্ট প্রোফ	১৪৬৯	১৪৭০
আনামান	১৪৭১	১৪৭২
পোর্ট প্রোফ	১৪৭৩	১৪৭৪

সিদ্ধান্ত নেওয়া

Problems of Projection method

- Cannot say what each block contains until further analysis.
 - Extract *features* from a zone
 - Recognize the zone content using a *classifier*
- Results are highly dependent even on small skew in the scanned page.

Zone content recognition

Features:

- Black pixel ratio (no. of black pixel / zone area)
- Horizontal transition (black to white) count
- Vertical transition (black to white) count
- Normalized mean length of horizontal black pixel run
- Normalized mean length of vertical black pixel run
- Connected component ratio

Classifier:

- Two-class (text and non-text)
SVM with RBF kernel (accuracy 94.89%)

Duong, Emptoz, Côté: Features for Printed Document Image Analysis, ICPR 2002.

Zone content recognition

- Functional classification of text blocks
 - Title / Heading, Sub-heading, Body text ...
- Features:
 - complexity (measured by entropy)
 - visibility values (or relative boldness)
 - directional compactness (horizontal and vertical)
 - geometric characteristics (block height, width, etc.)
- Classifier:
 - K-means clustering followed by min. distance classifier

Bres, Eglin, and Gafneux, *Unsupervised Clustering of Text Entities in Heterogeneous Grey Level Documents*, ICPR, 2002.

Problems of Projection method

- Cannot say what each block contains until further analysis
 - Extract *features* from a zone
 - Recognize the zone content using a *classifier*
- Results are highly dependent even on small skew in the scanned page
 - Detecting base line of each text line of the document
 - Determining orientation (slope) angle of base line
 - Estimation overall skew of the document page

Processing Tool

- Spatial domain operator that can handle shape information directly
- Mathematically well defined
- Neighborhood operator such that hardware implementation should be simple

Mathematical Morphology

- Mathematical morphological operators are good choice.

Objects

- All characters, figures, drawing, i.e., black components against white background

Structuring element

- Regular geometric figures:
 - mostly *line segment, square, circle, etc.*

Morphological Operations

Set theoretic operations (including union, intersection, etc.):

- 1. Dilation**
- 2. Erosion**
- 3. Opening**
- 4. Closing**

Morphological operator: Dilation

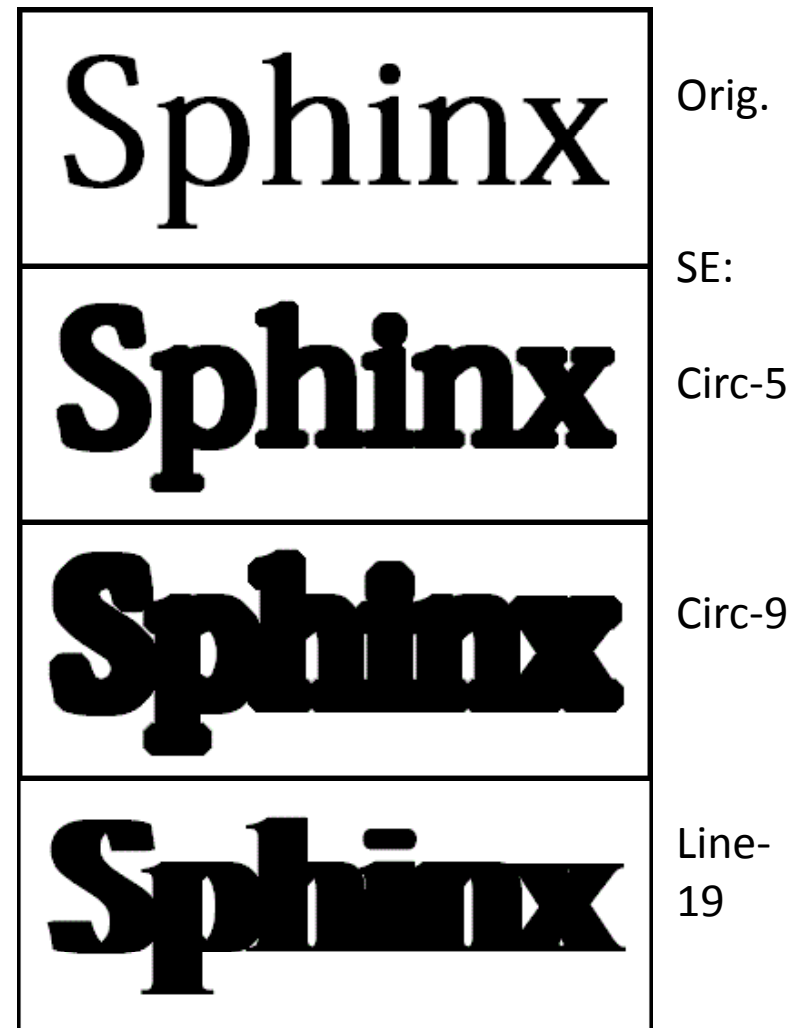
- Expands the objects.

$$A \oplus B = \{a + b \mid a \in A, b \in B\}$$

where A is an object and

B is SE.

- **Properties:**
Commutative,
associative,
distributive (over union),
increasing



Morphological operator: Erosion

- Shrinks the objects.

$$A \ominus B = \{p \mid B + p \in A\}$$

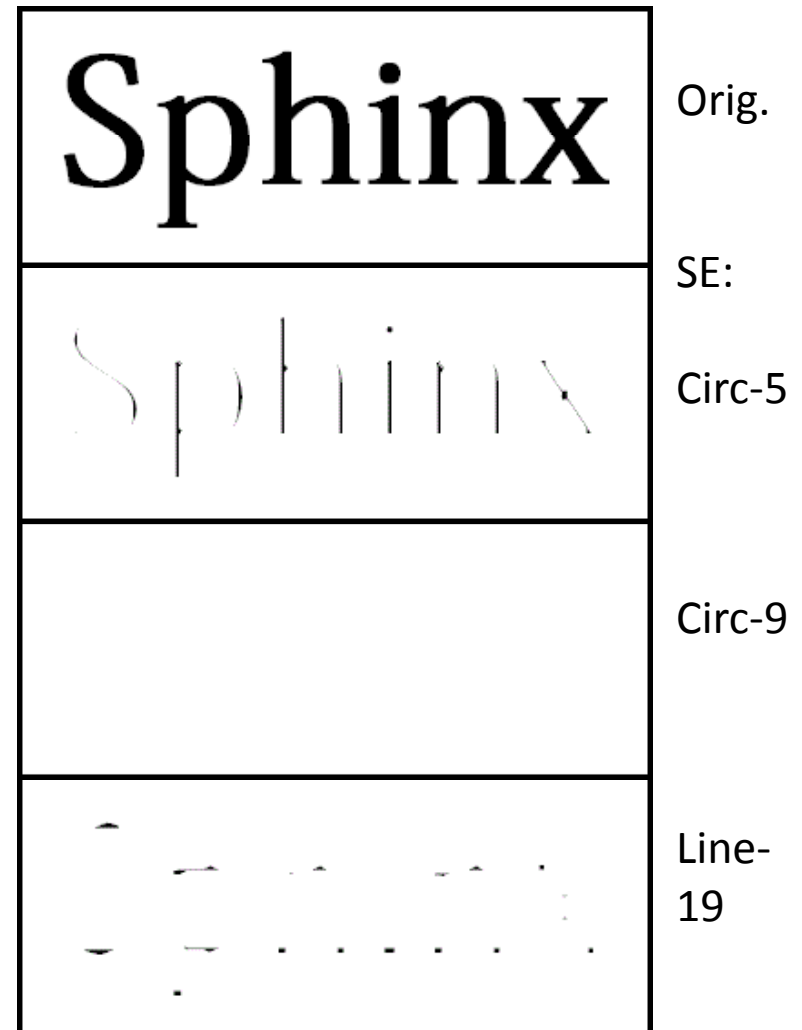
where A is an object and

B is SE.

- **Properties:**

Distributive (over intersection),
increasing.

- *Dilation and erosion are dual.*



Morphological operator: Opening

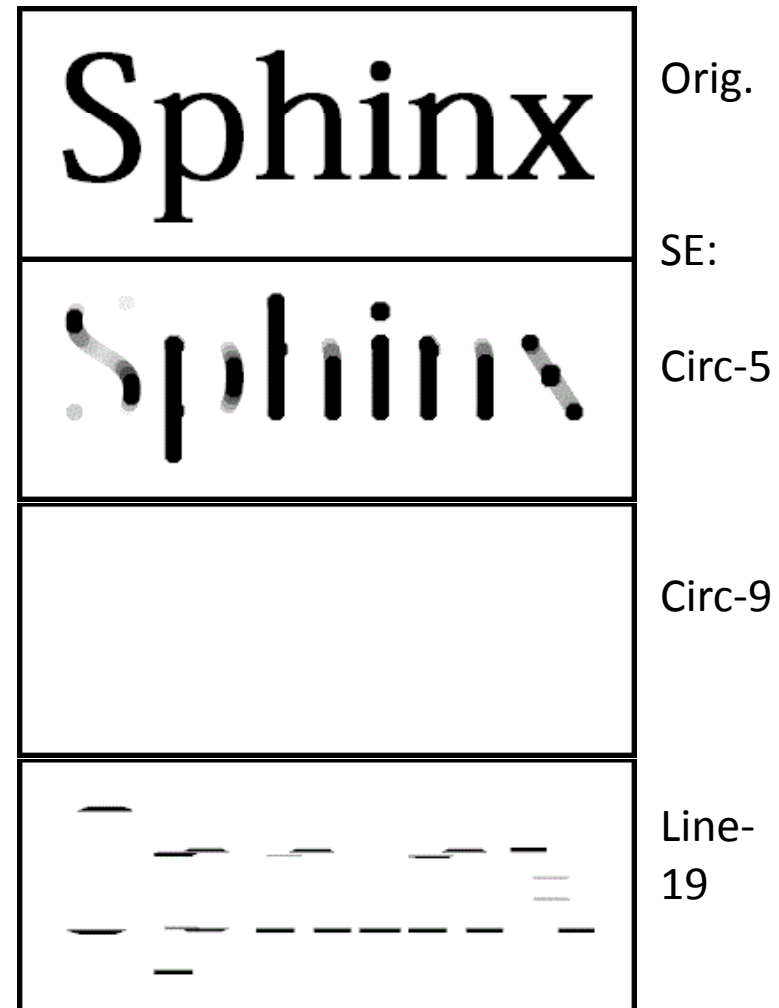
- Removes objects or parts of it that cannot fit in SE.

$$A \circ B = (A \ominus B) \oplus B$$

where A is an object and

B is SE.

- Properties:**
Increasing,
idempotent,
anti-extensive.
- It is a filter.*



Morphological operator: Closing

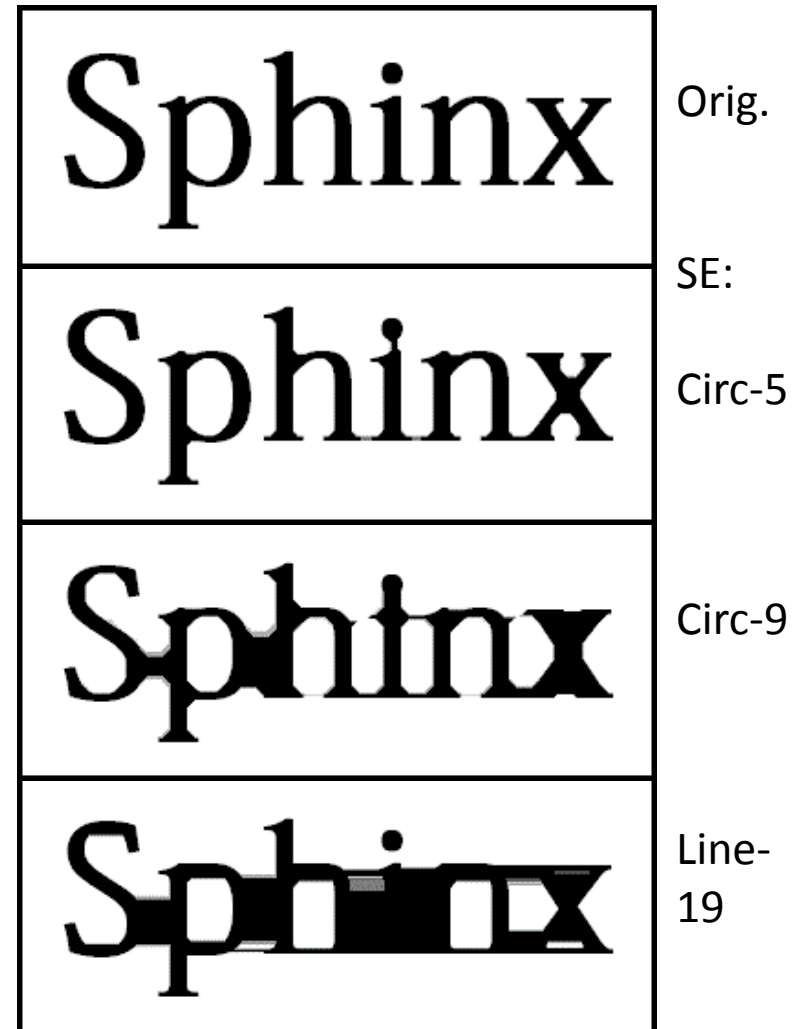
- Appends to objects parts of background if SE does not fit.

$$A \bullet B = (A \oplus B) \ominus B$$

where A is an object and

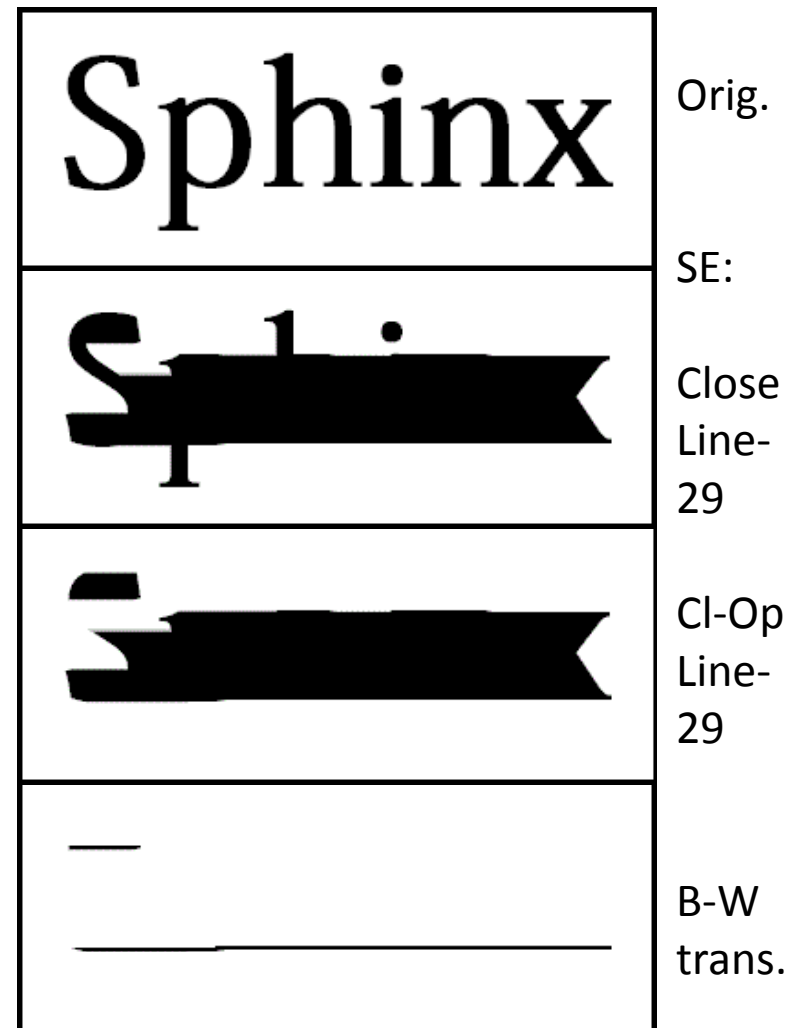
B is SE.

- **Properties:**
Increasing, idempotent, and extensive.
- It is a filter.
- *Opening & closing are dual.*



Detecting base line

- Close the original image with line SE of suitable length.
- Open the close image with same line SE.
- Detect black to white transition in vertical scan.



Font

- Traditionally, in metal typesetting, a **font** is a particular *size*, *weight* and *style* of a typeface.
- The weight of a particular font is the thickness of the character outlines relative to their height.
- Font size is measured in *point* unit.

1 point in is equal to ...

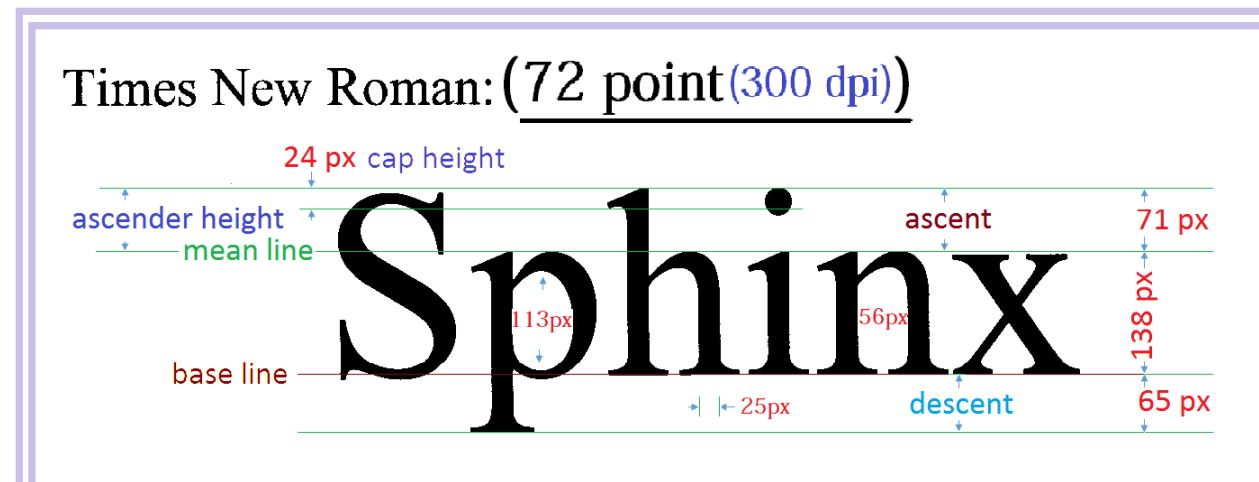
typographic units 1/12 picas

imperial/US units 1/72 inch

metric (SI) units 0.3528 mm

Size related parameters

- X-height or *corpus* height
- Ascender
- Descender



- Scan resolution (in dpi)
- Font style: bold, italics, ornamental

Skew correction: An example

Using a modem is easy in a corporate office, with a staff of computer gurus on call whenever you have a question. It's tougher if you're home alone, or in a small company with no on-site support. We designed Sportster for people like you, who want their modems to work the first time, every time. People who have no time to learn complicated procedures. But we built in all the features a rocket scientist could ask for.

Sportster offers all the speed and convenience big business demands, at mom-and-pop prices. Even the smallest office has room for a Sportster — it's about half the size of this page, but it comes with everything experienced users demand:

[Redacted text block]

[Redacted text block]

[Redacted text block with pruning points]

• Possible pruning points

Pages with complex layout

HERBERT MATTER
BIBLIOGRAPHY & WORK



Herbert Matter
1889-1977

Herbert Matter was a Swiss-born American typographer and designer. He is best known for his work on the design of the *Knopf Catalogue*, which was a landmark in the history of book design. Matter's work was characterized by its clean, modern lines and its use of bold, sans-serif typefaces. He was a pioneer in the use of the "modern" style of typography, which emphasized clarity and simplicity over the ornate, decorative styles of the past.

Herbert Matter was a Swiss-born American typographer and designer. He is best known for his work on the design of the *Knopf Catalogue*, which was a landmark in the history of book design. Matter's work was characterized by its clean, modern lines and its use of bold, sans-serif typefaces. He was a pioneer in the use of the "modern" style of typography, which emphasized clarity and simplicity over the ornate, decorative styles of the past.

Herbert Matter was a Swiss-born American typographer and designer. He is best known for his work on the design of the *Knopf Catalogue*, which was a landmark in the history of book design. Matter's work was characterized by its clean, modern lines and its use of bold, sans-serif typefaces. He was a pioneer in the use of the "modern" style of typography, which emphasized clarity and simplicity over the ornate, decorative styles of the past.

Cooking the Books
The Accounting [scandal] in Copart America



Cooking the Books is a satirical piece about the accounting industry. It uses a mix of bold, serif, and sans-serif fonts to create a visually busy and somewhat chaotic layout. The text is arranged in columns and paragraphs, with some words in larger, bolder fonts to draw attention. The overall design is a blend of traditional and modern typographic styles, reflecting the complexity and often convoluted nature of the subject matter.

Cooking the Books is a satirical piece about the accounting industry. It uses a mix of bold, serif, and sans-serif fonts to create a visually busy and somewhat chaotic layout. The text is arranged in columns and paragraphs, with some words in larger, bolder fonts to draw attention. The overall design is a blend of traditional and modern typographic styles, reflecting the complexity and often convoluted nature of the subject matter.

Little Devilant Tales 77



Little Devilant Tales is a collection of short stories or essays. The layout is highly decorative and artistic, featuring a large, stylized illustration of a person's head with a dark, abstract shape on top. The text is integrated into the design, with some words in large, bold, serif fonts. The overall aesthetic is dark and moody, with a focus on visual storytelling and typography.

Little Devilant Tales is a collection of short stories or essays. The layout is highly decorative and artistic, featuring a large, stylized illustration of a person's head with a dark, abstract shape on top. The text is integrated into the design, with some words in large, bold, serif fonts. The overall aesthetic is dark and moody, with a focus on visual storytelling and typography.

The Quizzers



The Quizzers is a collection of quizzes or tests. The layout is clean and modern, with a large, bold, serif font for the title. The text is arranged in columns and paragraphs, with some words in larger, bolder fonts. The overall design is simple and functional, focusing on readability and ease of use.

The Quizzers is a collection of quizzes or tests. The layout is clean and modern, with a large, bold, serif font for the title. The text is arranged in columns and paragraphs, with some words in larger, bolder fonts. The overall design is simple and functional, focusing on readability and ease of use.

'n Moderne chimera Omstredenheid heers rontom mens-dier-kruisklone.



'n Moderne chimera Omstredenheid heers rontom mens-dier-kruisklone. is a satirical piece about genetic engineering and cloning. The layout is highly decorative and artistic, featuring a large, stylized illustration of a person's head with a dark, abstract shape on top. The text is integrated into the design, with some words in large, bold, serif fonts. The overall aesthetic is dark and moody, with a focus on visual storytelling and typography.

'n Moderne chimera Omstredenheid heers rontom mens-dier-kruisklone. is a satirical piece about genetic engineering and cloning. The layout is highly decorative and artistic, featuring a large, stylized illustration of a person's head with a dark, abstract shape on top. The text is integrated into the design, with some words in large, bold, serif fonts. The overall aesthetic is dark and moody, with a focus on visual storytelling and typography.

I CAN'T SAY THAT I'M SCARED OF ANYTHING.



I CAN'T SAY THAT I'M SCARED OF ANYTHING. is a collection of essays or articles. The layout is clean and modern, with a large, bold, serif font for the title. The text is arranged in columns and paragraphs, with some words in larger, bolder fonts. The overall design is simple and functional, focusing on readability and ease of use.

I CAN'T SAY THAT I'M SCARED OF ANYTHING. is a collection of essays or articles. The layout is clean and modern, with a large, bold, serif font for the title. The text is arranged in columns and paragraphs, with some words in larger, bolder fonts. The overall design is simple and functional, focusing on readability and ease of use.

Morphological algorithm

- Text region is composed of small objects (characters) placed in regular interval.
- Opening the image with small SE removes the thin object parts (strokes of character), but has insignificant effect on large objects in half-tone etc.
- Closing the image with small SE fills in white holes in small objects (space within and between character), but has insignificant effect on large white space or half-tone.
- Thus difference between closed and opened image highlights the text region.
- Difference image is thresholded to detect text region.

Morphological approach: An example



I would like to thank the outgoing Executive Council members and welcome the new ones. Late Aggarwal replaces Michael Duff in the capacity of Past President. Gunilla Bergfors takes over from Saburo Tsuji as 1st Vice-President. They will both provide valuable counsel in the operation of IAPR. Gunilla's fiancée, Dr. Bagg, has demonstrated her effectiveness and efficiency in her chair of the Education Committee, replaces Gun Bergfors as Secretary. Josef Bagan is succeeding S. Tamimoto as Treasurer, taking the responsibility for financial well-being of the Association across the globe for the first time in IAPR's history. Prof. Taniaka is becoming Second Vice-President. The most urgent of a new administration is the appointment of 12 Committees. Vice-President Gunilla Bergfors continues to chair the Membership Committee to cope with the contacts she has established in past roles - potential new member countries.

For the time being, I shall continue to discharge my full responsibility for IAPR Technical Committee, now as of the duties of the President. This task should not naturally come as all IAPR Technical Council Chairmen have already been appointed. David Paul joining TC1 (Statistical Pattern Recognition) Teaching as a co-chairman jointly with Doro Geiger, Marco Per will chair TC2 (Special-Purpose Architectures). F. Leberl assumes the responsibility for TC3 (Application Remote Sensing). Prof. Nobida takes over (Applications in Industry). Jan Gerbrands will chair (Biomedical Pattern Recognition) and Gan Lo succeeds Re-join Platenodon in running TC4 (Applications in Text Processing). Horac Ip and Ar. Seneiders agreed to vacate the newly established TC5 (Multimedia and Communication Systems) and Ce. Maccorone will chair TC6 (Pattern Recognition in Astronomy and Astrophysics). The other Technical Committees will continue under their past chairmen.

Following a well established tradition, the Honorary Committee will be chaired by Past President Aggarwal. Other members of the Nominating Committee are Dr. M. Egan, Prof. E. S. Gebruera, Prof. H. Niemann

me always been happy and willing to serve the IAPR in whatever capacity was called for as borne witness by my numerous jobs with which I had been entrusted in the past such as Chairman of TC1, man of Publications and Publicity Committee, J. Chairman of IAPR Technical Committee, Inter Editor, 2nd Vice President and my membership on IAPR Committees at some time or other. In that I was not all that apprehensive to leave technical work with some new responsibility with which I could be called to the Association. Nevertheless, ending the presidency came to me as a surprise, inducing a mixed feeling of delight from the honour bestowed by the Governing Board and at the same time, conscious of the fact of the huge responsibility this job entails. I would like to thank the Governing Board for electing me to this position and say that I will do my best to meet its challenges.

myself as President follows that of Professor Late Aggarwal. Other members of the Nominating Committee are Dr. M. Egan, Prof. E. S. Gebruera, Prof. H. Niemann



me always been happy and willing to serve the IAPR in whatever capacity was called for as borne witness by my numerous jobs with which I had been entrusted in the past such as Chairman of TC1, man of Publications and Publicity Committee, J. Chairman of IAPR Technical Committee, Inter Editor, 2nd Vice President and my membership on IAPR Committees at some time or other. In that I was not all that apprehensive to leave technical work with some new responsibility with which I could be called to the Association. Nevertheless, ending the presidency came to me as a surprise, inducing a mixed feeling of delight from the honour bestowed by the Governing Board and at the same time, conscious of the fact of the huge responsibility this job entails. I would like to thank the Governing Board for electing me to this position and say that I will do my best to meet its challenges.

me always been happy and willing to serve the IAPR in whatever capacity was called for as borne witness by my numerous jobs with which I had been entrusted in the past such as Chairman of TC1, man of Publications and Publicity Committee, J. Chairman of IAPR Technical Committee, Inter Editor, 2nd Vice President and my membership on IAPR Committees at some time or other. In that I was not all that apprehensive to leave technical work with some new responsibility with which I could be called to the Association. Nevertheless, ending the presidency came to me as a surprise, inducing a mixed feeling of delight from the honour bestowed by the Governing Board and at the same time, conscious of the fact of the huge responsibility this job entails. I would like to thank the Governing Board for electing me to this position and say that I will do my best to meet its challenges.

I would like to thank the outgoing Executive Council members and welcome the new ones. Late Aggarwal replaces Michael Duff in the capacity of Past President. Gunilla Bergfors takes over from Saburo Tsuji as 1st Vice-President. They will both provide valuable counsel in the operation of IAPR. Gunilla's fiancée, Dr. Bagg, has demonstrated her effectiveness and efficiency in her chair of the Education Committee, replaces Gun Bergfors as Secretary. Josef Bagan is succeeding S. Tamimoto as Treasurer, taking the responsibility for financial well-being of the Association across the globe for the first time in IAPR's history. Prof. Taniaka is becoming Second Vice-President. The most urgent of a new administration is the appointment of 12 Committees. Vice-President Gunilla Bergfors continues to chair the Membership Committee to cope with the contacts she has established in past roles - potential new member countries.

For the time being, I shall continue to discharge my full responsibility for IAPR Technical Committee, now as of the duties of the President. This task should not naturally come as all IAPR Technical Council Chairmen have already been appointed. David Paul joining TC1 (Statistical Pattern Recognition) Teaching as a co-chairman jointly with Doro Geiger, Marco Per will chair TC2 (Special-Purpose Architectures). F. Leberl assumes the responsibility for TC3 (Application Remote Sensing). Prof. Nobida takes over (Applications in Industry). Jan Gerbrands will chair (Biomedical Pattern Recognition) and Gan Lo succeeds Re-join Platenodon in running TC4 (Applications in Text Processing). Horac Ip and Ar. Seneiders agreed to vacate the newly established TC5 (Multimedia and Communication Systems) and Ce. Maccorone will chair TC6 (Pattern Recognition in Astronomy and Astrophysics). The other Technical Committees will continue under their past chairmen.

Following a well established tradition, the Honorary Committee will be chaired by Past President Aggarwal. Other members of the Nominating Committee are Dr. M. Egan, Prof. E. S. Gebruera, Prof. H. Niemann



me always been happy and willing to serve the IAPR in whatever capacity was called for as borne witness by my numerous jobs with which I had been entrusted in the past such as Chairman of TC1, man of Publications and Publicity Committee, J. Chairman of IAPR Technical Committee, Inter Editor, 2nd Vice President and my membership on IAPR Committees at some time or other. In that I was not all that apprehensive to leave technical work with some new responsibility with which I could be called to the Association. Nevertheless, ending the presidency came to me as a surprise, inducing a mixed feeling of delight from the honour bestowed by the Governing Board and at the same time, conscious of the fact of the huge responsibility this job entails. I would like to thank the Governing Board for electing me to this position and say that I will do my best to meet its challenges.

(a) Original image

(b) Closed image

(c) Opened image

QUESTION: Size of structuring element?

Results

Input test image

Resultant (labeled) image

HERBERT MATTER
BIBLIOGRAPHY & WORK



1. Like the other ...
2. America calling 1923
3. Container Corporation of America 1926

Analysis of famous works
Swiss National Tourist Office 1934

Bibliography
Take a closer look at the main element of the poster: the white background in white and dark white. The white background and dark white seem to have created a perfect balance which has led to the bottom right corner of the poster. Of the entire poster the face of the poster is dominant, it has the big size, and the dark white which creates a strong movement in the poster. Consequently, its shape is greatly emphasized by this contrast. The main part used is Swiss modern design and is very legible for design. Moreover, the differences between the sizes of font greatly enhance the depth in the poster. The existence of the Helvetica font successfully creates an order which attracts more attention with viewers' eyes.

With none unnecessary elements printed for the poster of a basic principle, the combination of many color or the Swiss design. The main part used is Swiss modern design and is very legible for design. Moreover, the differences between the sizes of font greatly enhance the depth in the poster. The existence of the Helvetica font successfully creates an order which attracts more attention with viewers' eyes.

HERBERT MATTER
BIBLIOGRAPHY & WORK



1. Like the other ...
2. America calling 1923
3. Container Corporation of America 1926

Analysis of famous works
Swiss National Tourist Office 1934

Bibliography
Take a closer look at the main element of the poster: the white background in white and dark white. The white background and dark white seem to have created a perfect balance which has led to the bottom right corner of the poster. Of the entire poster the face of the poster is dominant, it has the big size, and the dark white which creates a strong movement in the poster. Consequently, its shape is greatly emphasized by this contrast. The main part used is Swiss modern design and is very legible for design. Moreover, the differences between the sizes of font greatly enhance the depth in the poster. The existence of the Helvetica font successfully creates an order which attracts more attention with viewers' eyes.

With none unnecessary elements printed for the poster of a basic principle, the combination of many color or the Swiss design. The main part used is Swiss modern design and is very legible for design. Moreover, the differences between the sizes of font greatly enhance the depth in the poster. The existence of the Helvetica font successfully creates an order which attracts more attention with viewers' eyes.

"Herbert had a strong feeling for minute details, and this was exemplified by the distinguished typography he did for the Knoll catalogues"

"Herbert had a strong feeling for minute details, and this was exemplified by the distinguished typography he did for the Knoll catalogues"

Results

Input test image

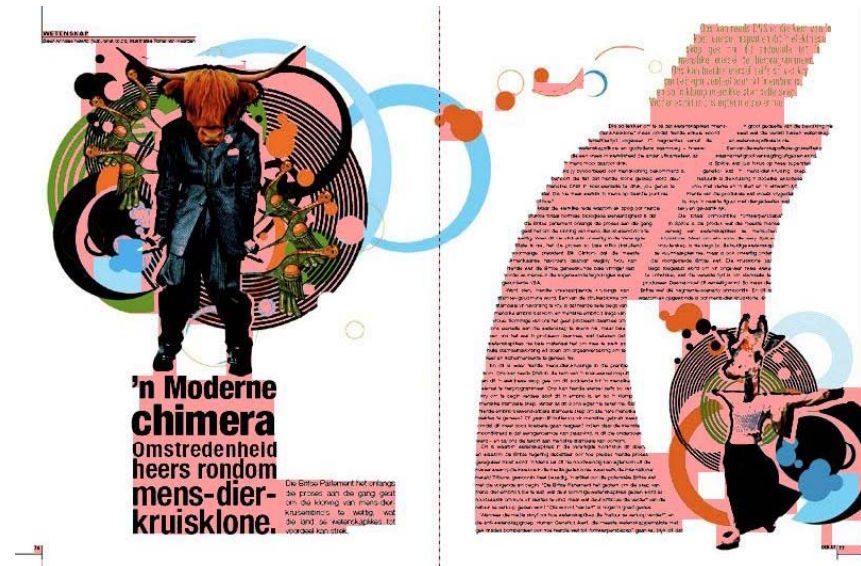
Resultant (labeled) image



Results

Input test image

Resultant (labeled) image



Deep learning

- Popular technique for unsupervised feature extraction for supervised applications
 - Ex. object recognition.
- Utilizes HUGE number of instances to train relatively simpler system to perform more complicated task.
- Training samples may be outcome of controlled or uncontrolled data acquisition.
- Requires very high computational resources for implementing a reasonably meaningful system.

Detect text area using CNN

Input: A document image

Output: Text / Non-text area

5. Experimental results

The method produces very good results for many images within a reasonable time. For example, the results of images in Figure-1 are given in Figure-4, where damages are successfully repaired.



Figure-4: Inpainted results of Figure-1. Note that scratches are removed.

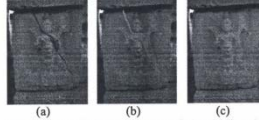


Figure-5: Inpainting of image of damaged artifacts. (a) Image of artifact with crack. (b) Crack region manually marked and declared missing region. (c) Inpainted image.

Image inpainting technique can be employed in various applications including visual restoration of damaged objects (see Figure-5). This may be considered as digital restoration of archeological artifacts where image of the damaged object is inpainted to view the original look of the object without crack. Note that here the problem not with the image, but with the scene or object in the scene.

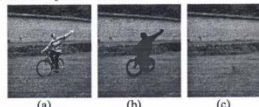


Figure-6: Removal of undesired object by image inpainting. (a) Image with undesired object (bicycle rider). (b) Bicycle rider manually marked and declared missing region. (c) Inpainted image.

Another interesting and important application is removal of undesired object(s) from the scene. Here again the problem is not with the image, rather it is with the scene itself, which contains that undesired

object. However, we treat the image as before. That means we manually cover the undesired object in the image and mark that portion as the missing region or target region. Inpainting algorithm fills up this region from the environment or source region. Thus image of the scene is completed without that undesired object (see Figure-6).

So far we have considered the cases where source region is in the same image that contains the missing region. However, there are some cases where source region is so small that it may not contain sufficient information. In those cases we manually choose relevant images from our database and use them as source of patches to perform inpainting. An example is shown in Figure-7.

5.1 Image expansion by texture synthesis

In the aforementioned experiments the missing (target) regions are always enclosed by the known (source) region, and filling process starts at the boundary and gradually progresses inward from all sides. Question is what would happen if the arrangement is reversed, that means, source region is enclosed by target region. Note that boundary detection is done by local operator (usually logical operation over a small neighbourhood). So this process does not discriminate between interior region and exterior region. Second, the filling process never considers whether it moves inward or outward, it simply moves from known to unknown. These observations suggest that inpainting can be done in regions adjacent but exterior to the source region. In fact, we exploit this strategy to develop the algorithm for image expansion.

Note that both image magnification and image expansion increases the size of the image. However, in the former case the size of objects and background increases by the same factor, while in the latter their scale remains same and they are repeated to expand the extent of the image (see Figure-8).

An interesting application of this method is generating large motif from only a small part of it as shown in Figure-9.

6. Conclusion

In this article we have presented an automated image correction method by inpainting which can successfully fill in missing information resulted from physical damage or due to presence of undesired object in the scene. The method can also expand the image keeping its components in the same scale. We treat the problem in a unified way by marking the region manually and label it as target region to be filled in.



Solution strategy

Transforming the problem into a classification Problem.

- Divide the Input image into $M \times M$ patches.
- **Input:** Image patch of size $M \times M$
- **Output:** Text, Non text, and Ambiguous
 - Text: if $>80\%$ of the patch has text
 - Non-text: if $<20\%$ of the patch has text area
 - Ambiguous: otherwise

Training data

P450 3A and morphine synthesis

Cytochrome P450 3A Enzymes Catalyze the *O*⁶-Demethylation of Thebaine, a Key Step in Endogenous Mammalian Morphine Biosynthesis

Valerie M. Kramlinger, Mónica Alvarado Rojas, Tatsuyuki Kanamori, and F. Peter Guengerich

From the Department of Biochemistry, School of Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146

Running title: *P450 3A and morphine synthesis*

To whom correspondence should be addressed: Prof. F. Peter Guengerich, Department of Biochemistry, Vanderbilt University School of Medicine, 638B Robinson Research Building, 2200 Pierce Avenue, Nashville, Tennessee 37232-0146, Telephone: (615) 322-2261; FAX: (615) 343-0704; E-mail: fguengerich@vanderbilt.edu

Keywords: Cytochrome P450, biosynthesis, natural product, oxygenase, morphine, alkaloids

ABSTRACT

Morphine, first characterized in opium from the poppy *Papaver somniferum*, is one of the strongest known analgesics. Endogenous morphine has been identified in several mammalian cells and tissues. The synthetic pathway of morphine in the opium poppy has been elucidated, and the presence of common intermediates in plants and mammals suggests that biosynthesis occurs through similar pathways (beginning with the amino acid L-tyrosine), and the pathway has been completely delineated in plants. Some of the enzymes in the mammalian pathway have been identified and characterized. Two of the latter steps in the morphine biosynthesis pathway are demethylation of thebaine at the *O*⁵- and the *O*⁶-positions, the latter of which has been difficult to demonstrate. The plant enzymes responsible for both the *O*⁵- and the *O*⁶-demethylation are members of the Fe^{II}/α-ketoglutarate-dependent dioxygenase family. Previous studies showed that human cytochrome P450 (P450) 2D6 can catalyze thebaine *O*⁵-demethylation. We report that demethylation of thebaine at the *O*⁶-position is selectively catalyzed by human P450s 3A4 and 3A5, with the latter being more efficient, and rat P450 3A2. Our results do not support *O*⁶-demethylation of thebaine by an Fe^{II}/α-ketoglutarate-dependent dioxygenase. In rat brain microsomes, *O*⁶-demethylation was inhibited by ketoconazole, but not sulfaphenazole, suggesting that P450 3A enzymes are responsible for this activity in the brain. An alternate pathway to

morphine, oripavine *O*⁶-demethylation, was not detected. The major enzymatic steps in mammalian morphine synthesis have now been identified.

Morphine is the principal benzylisoquinoline alkaloid produced in the opium poppy, *Papaver somniferum*. Its mechanism of action is as an agonist for the μ-opioid receptors that are distributed throughout the brain. Activation of these receptors is associated with analgesia, sedation, euphoria, physical dependence, and respiratory depression. The μ-binding sites are discretely distributed in the human brain, with high densities in the posterior amygdala, hypothalamus, thalamus, nucleus caudatus, putamen, and certain cortical areas (1).

A considerable body of evidence exists that morphine is present in the tissues of various animals that have not been medicated with morphine or other related opioids. In mammals, morphine has been detected in skin, lung, spinal cord, and, most notably, in the brain (2-4). The presence of morphine in brain is of particular interest due to the presence of the μ-opioid receptors. In the rat, morphine levels in structures of the brain have been quantified and range from 26 fmol/g tissue (found globally in the brain) to 7.2 pmol/g tissue measured in the hypothalamus (3,5-7). One study quantified morphine in the temporal lobe from one human brain tissue at 3.4 ng/g tissue (8). The presence of morphine in brain tissue in many mammals and the

1



1

Prepare training data

INPUT: document images with manually labeled text area.

- From each image, overlapping patches of size 100x100 are taken (stride along x, y is 20) and resized to 50x50
- From each image, overlapping patches of size 50x50 are taken (stride along x, y is 10)
- Each 50x50 patch is divided into 4 patches of size 25x25 and are resized back to 50x50.
- We get total number of **825670** patches of size 50x50 as training data from 8 images.

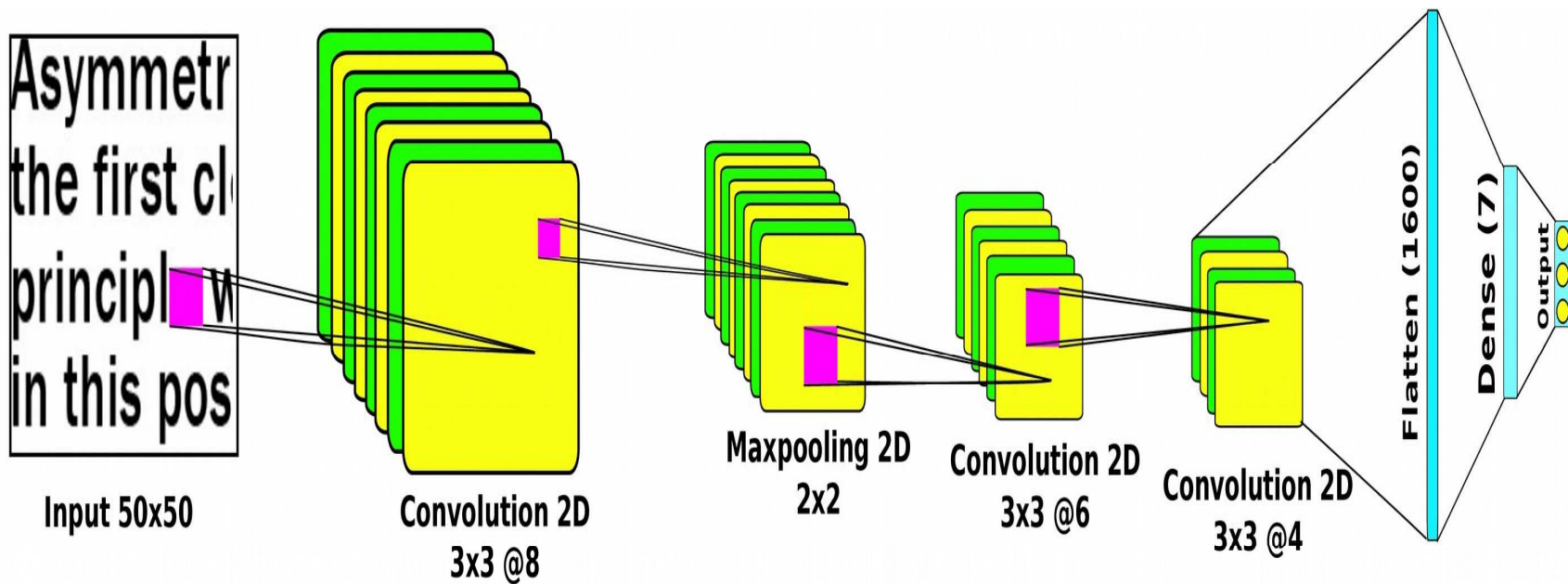
Label: as described before.

Model description

Input: 50x50 Patch of gray scale.

Layer (type)	Output Shape	Param #
Convolution2D(3x3 @8)	(8, 48, 48)	80
MaxPooling2D(2x2)	(8, 24, 24)	0
Convolution2D(3x3 @6)	(6, 22, 22)	438
Convolution2D(3x3 @4)	(4, 20, 20)	220
Flatten	(1600)	0
Dense(7)	(7)	11207
Activation(Sigmoid)	(7)	0
Dense(3)	(3)	24
Activation(Softmax)	(3)	0
=====		
Total parameters: 11969		

Model description



Training the model

- Number of epoch: 200
- Batch size: 100
- Learning Rate: 0.01
- Learning weight decay: 0.95
- Optimizer: Stochastic gradient descent
- Loss function: Mean squared error

Testing

- Input: A test image
- Take 50x50 patch and submit it to the trained model
- If predicted class is text, color that patch as pink.
- If predicted class is non-text, color the patch as white.
- If predicted class is ambiguous, then
 - Divide that patch into 4 patches of size 25x25 and resize to 50x50 and submitted to the model.
 - If that 50x50 patch is again ambiguous, then color that patch as yellow (Ideally it should be done recursively until we get no ambiguous patch)
 - Else color the patch as according to text or non-text class.

Results

Input test image

Resultant (labeled) image



Results

Input test image



Resultant (labeled) image



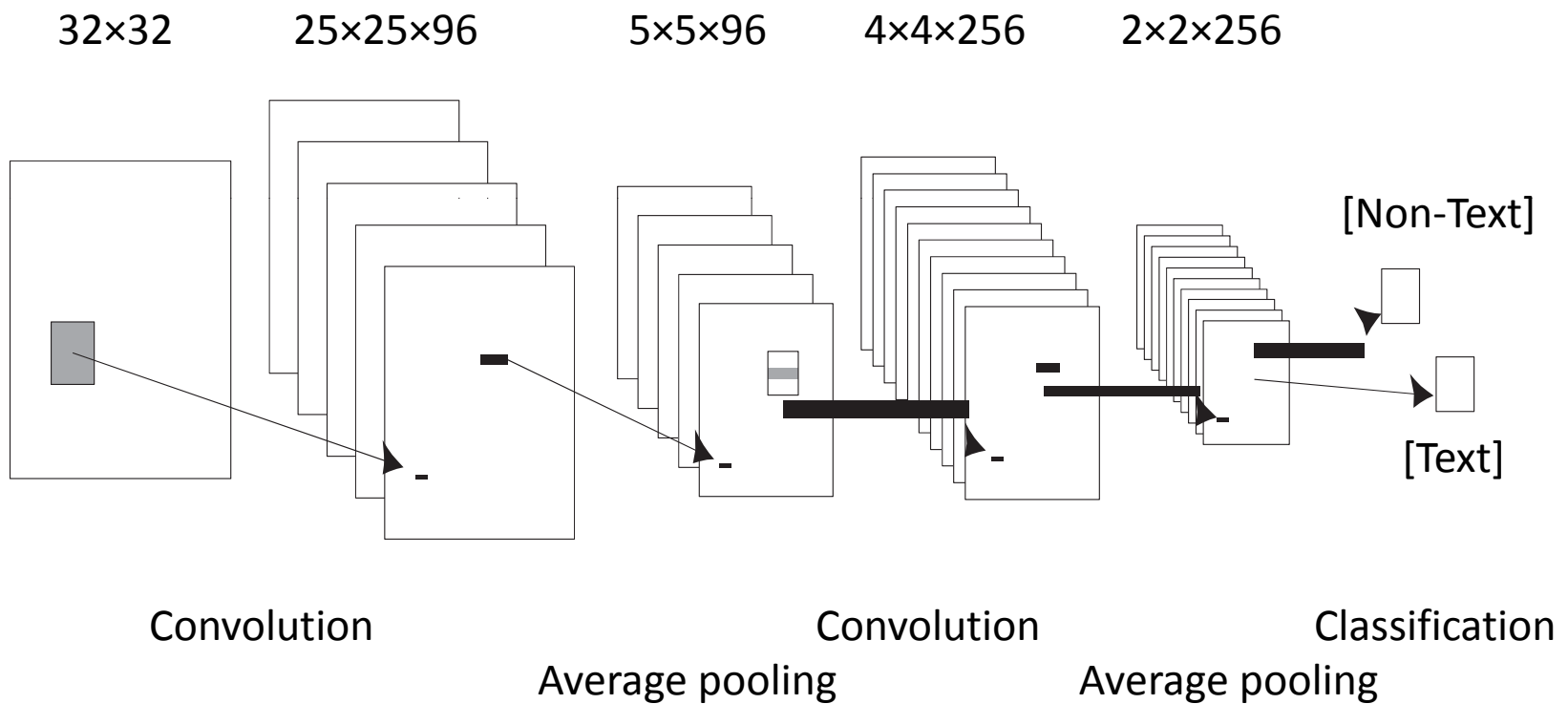
Results

Input test image

Resultant (labeled) image



An improved network



Wang, Wu, Coates and Ng, **End-to-End Text Recognition with Convolutional Neural Networks**, *ICPR 2012*.

Comparative results

Simpler system

HERBERT MATTER BIBLIOGRAPHY & WORK

1. America calling 1923
Look closer at the poster. It can be seen that the main color used is red and blue. The background is black and gray along with elements which greatly emphasize the 3D shape of the eagle. We can see that there is a perfect combination between these three colors. Each block has a color. However, they are grouped perfectly that it is hard to find the contact point between blocks. Even though the lines in this work are not used that other poster before, they appear on the top and bottom that create sense which relate to sky and earth. Moreover, we can see that those lines enhance the energy of the movement of the eagle.

2. America calling 1923
Like the other posters, the Bauhaus color exists in the work of Herbert Matter. Red and blue are applied abundantly which are main colors of national flag of American, along with shades which greatly emphasize the 3D shape of the eagle. We can see that the main object is a perfect combination between three other blocks. Each block has a color. However, they are grouped perfectly that it is hard to find the contact point between blocks. Even though they are over lapping together, it is hard to point out where the touch is. They seem to be meeting at the touch. Furthermore, the contrast, size, and visual weight between blocks are equal. It can be seen that the unity is applied wisely with a 2-Dimensional poster. The typography which is located in the poster does not be dominant in the background, but they stand out.

3. Container Corporation of America 1962
Look closer at the poster. It can be seen that the main color used is red and blue. The background is black and gray along with elements which greatly emphasize the 3D shape of the eagle. We can see that there is a perfect combination between these three colors. Each block has a color. However, they are grouped perfectly that it is hard to find the contact point between blocks. Even though they are over lapping together, it is hard to point out where the touch is. They seem to be meeting at the touch. Furthermore, the contrast, size, and visual weight between blocks are equal. It can be seen that the unity is applied wisely with a 2-Dimensional poster. The typography which is located in the poster does not be dominant in the background, but they stand out.

Biography
Born in Switzerland in 1907, Herbert Matter was a famous graphic designer and typographer of Swiss. He was the place where he studied until 1927. After that, he went to study at the Bauhaus in Weimar, Germany. He studied typography and design. He worked for the Swiss National Tourist Office in 1934. He died in 1981.

Analysis of famous works
Swiss National Tourist Office 1934

With move forward, elements existed in the use of basic principles, the combination of many color and the Swiss typewriter that make the design. The all works of Herbert Matter show a rational style in his design. Strong, expressive, Herbert Matter's success in the design industry was due to his graphic work.

“Herbert had a strong feeling for minute details, and this was exemplified by the distinguished typography he did for the Knoll catalogues”

January 30, 2017

Wang et al.

HERBERT MATTER BIBLIOGRAPHY & WORK

1. America calling 1923
Look closer at the poster. It can be seen that the main color used is red and blue. The background is black and gray along with elements which greatly emphasize the 3D shape of the eagle. We can see that there is a perfect combination between these three colors. Each block has a color. However, they are grouped perfectly that it is hard to find the contact point between blocks. Even though they are over lapping together, it is hard to point out where the touch is. They seem to be meeting at the touch. Furthermore, the contrast, size, and visual weight between blocks are equal. It can be seen that the unity is applied wisely with a 2-Dimensional poster. The typography which is located in the poster does not be dominant in the background, but they stand out.

2. America calling 1923
Like the other posters, the Bauhaus color exists in the work of Herbert Matter. Red and blue are applied abundantly which are main colors of national flag of American, along with shades which greatly emphasize the 3D shape of the eagle. We can see that the main object is a perfect combination between three other blocks. Each block has a color. However, they are grouped perfectly that it is hard to find the contact point between blocks. Even though they are over lapping together, it is hard to point out where the touch is. They seem to be meeting at the touch. Furthermore, the contrast, size, and visual weight between blocks are equal. It can be seen that the unity is applied wisely with a 2-Dimensional poster. The typography which is located in the poster does not be dominant in the background, but they stand out.

3. Container Corporation of America 1962
Look closer at the poster. It can be seen that the main color used is red and blue. The background is black and gray along with elements which greatly emphasize the 3D shape of the eagle. We can see that there is a perfect combination between these three colors. Each block has a color. However, they are grouped perfectly that it is hard to find the contact point between blocks. Even though they are over lapping together, it is hard to point out where the touch is. They seem to be meeting at the touch. Furthermore, the contrast, size, and visual weight between blocks are equal. It can be seen that the unity is applied wisely with a 2-Dimensional poster. The typography which is located in the poster does not be dominant in the background, but they stand out.

Biography
Born in Switzerland in 1907, Herbert Matter was a famous graphic designer and typographer of Swiss. He was the place where he studied until 1927. After that, he went to study at the Bauhaus in Weimar, Germany. He studied typography and design. He worked for the Swiss National Tourist Office in 1934. He died in 1981.

Analysis of famous works
Swiss National Tourist Office 1934

With move forward, elements existed in the use of basic principles, the combination of many color and the Swiss typewriter that make the design. The all works of Herbert Matter show a rational style in his design. Strong, expressive, Herbert Matter's success in the design industry was due to his graphic work.

“Herbert had a strong feeling for minute details, and this was exemplified by the distinguished typography he did for the Knoll catalogues”

Indian Statistical Institute

61

Comparative results

Simpler system

when it comes to building brands, we don't have rules, but we do have beliefs.

Little Deviant Tales 77

76

Wang et al.

when it comes to building brands, we don't have rules, but we do have beliefs.

Little Deviant Tales 77

76

Benchmark database

- UW-I, II, III databases, Developed at University of Washington, Seattle, USA in 1996.
- Widely used earliest database with 1620 pages
- Zones contain text, non-text such as halftone, line drawing, math and chemical equation.
- The database also contains
 - *Page condition file* : skew angle, noise.
 - *Page attribute file* : dominant font and other content.
 - *Page bounding box file* : location and size of zones.

<http://isis-data.science.uva.nl/events/dlia//datasets/uwash3.html>

Benchmark database

- Mediateam document database
- Developed at University of Oulu, Finland in 1998.
- One of the early databases containing

Pattern type	Samples
Text	4811
Graphics	735
Image	161
Composite	219

Duong, Emptoz, Côté: Features for Printed Document Image Analysis, ICPR 2002.

Benchmark database

- Pattern Recognition and Image Analysis (PRImA) Layout Analysis dataset
- Developed at University of Salford, Manchester
- 1240 ground-truthed pages from magazines (1085 pages) and technical journals (155 pages)
- Used in following contests
 - ICDAR 2015 Recognition of Documents with Complex Layouts (RDCL2015)
 - ICDAR2013 Historical Newspaper Layout Analysis (HNLA2013)
 - ICDAR2011 Historical Document Layout Analysis (HDLAC 2011)

http://www.primaresearch.org/datasets/Layout_Analysis

Benchmark database

- Historical Newspaper dataset (ENP dataset)
- Developed at University of Salford, Manchester in *Europeana Newspapers Project*
- 500 ground-truthed pages covering
 - 13 languages (German, french, English, Estonian, etc.)
 - 17th, 18th, 19th and 20th centuries
- Contains (total regions 61,619) including
 - 1,497 image zones
 - 208 table zones
 - 46,889 text zones

Clausner et. Al, The ENP Image and Ground Truth Dataset of historical newspaper, ICDAR 2015.

Performance evaluation

- A document page D may be represented as a m tuple.

$$D = (E_1, E_2, \dots, E_m)$$

where E_i s are entities such as text, tables, half-tone, etc.

- Each entity has a unique property denoted by $\text{Prop.}(E_i)$.
- Document page image domain X has n bounding boxes B_j ($j=1, \dots, n$) with such that:

(i) $\bigcup_1^n B_j \subseteq X$

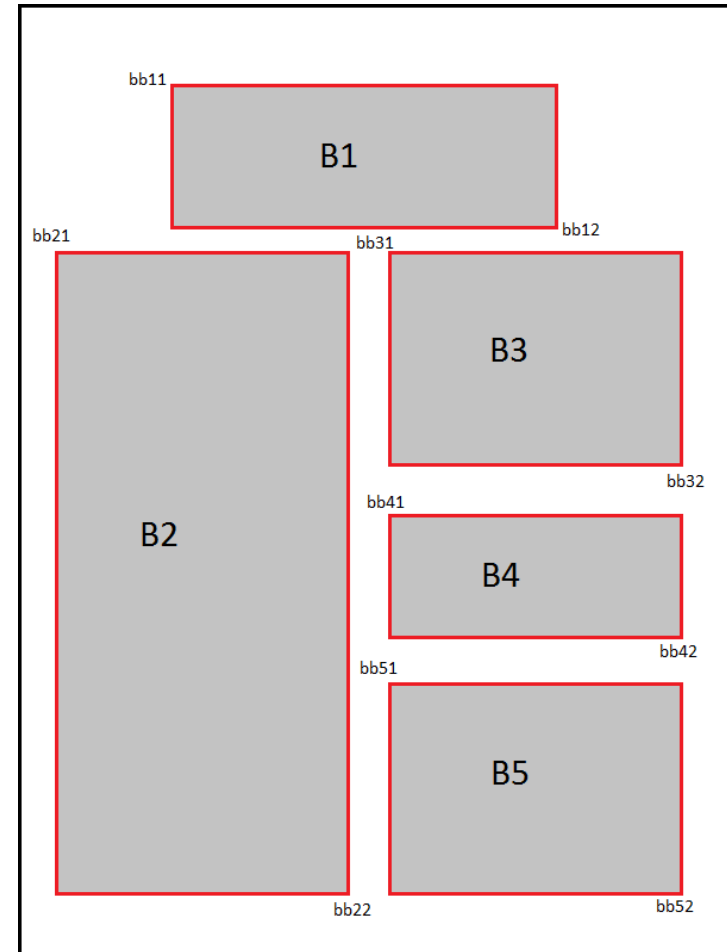
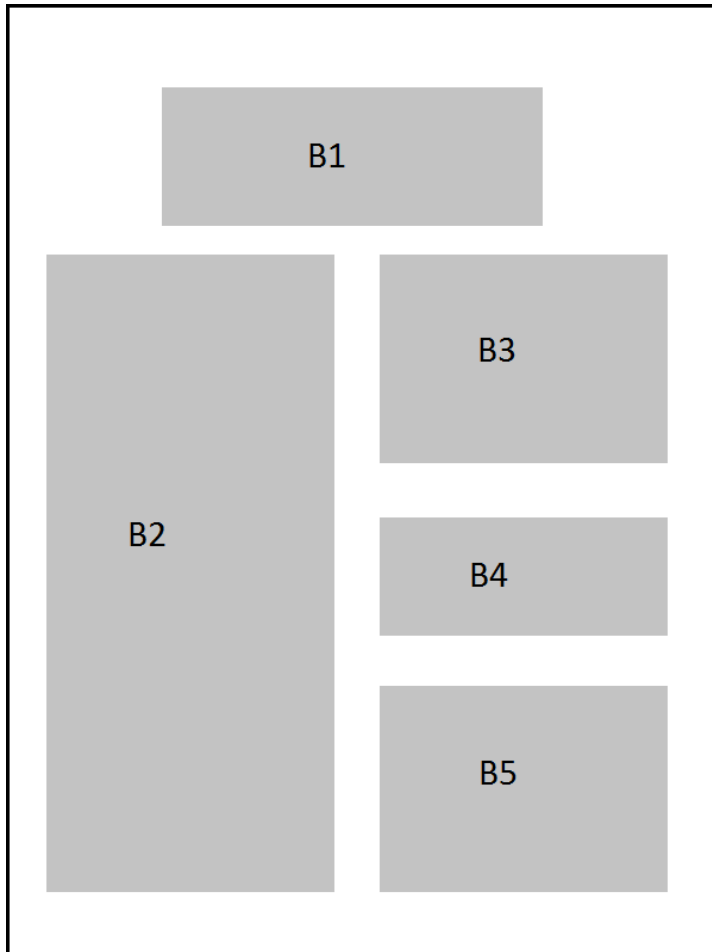
(ii) $B_j \cap B_k = \phi$ for $j \neq k$

(iii) For every j there exists one and only one i such that

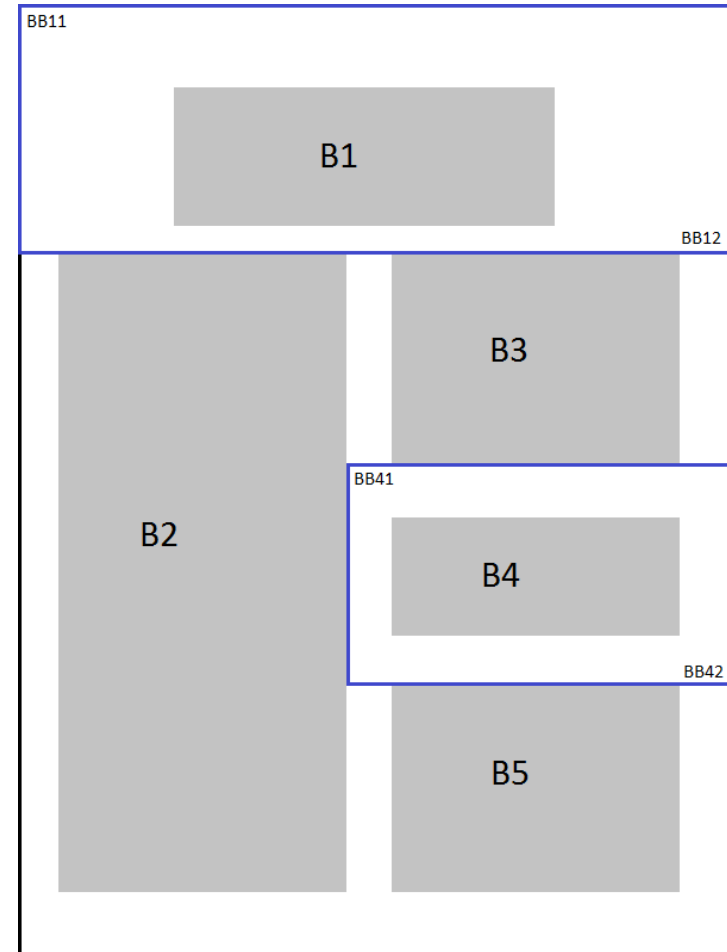
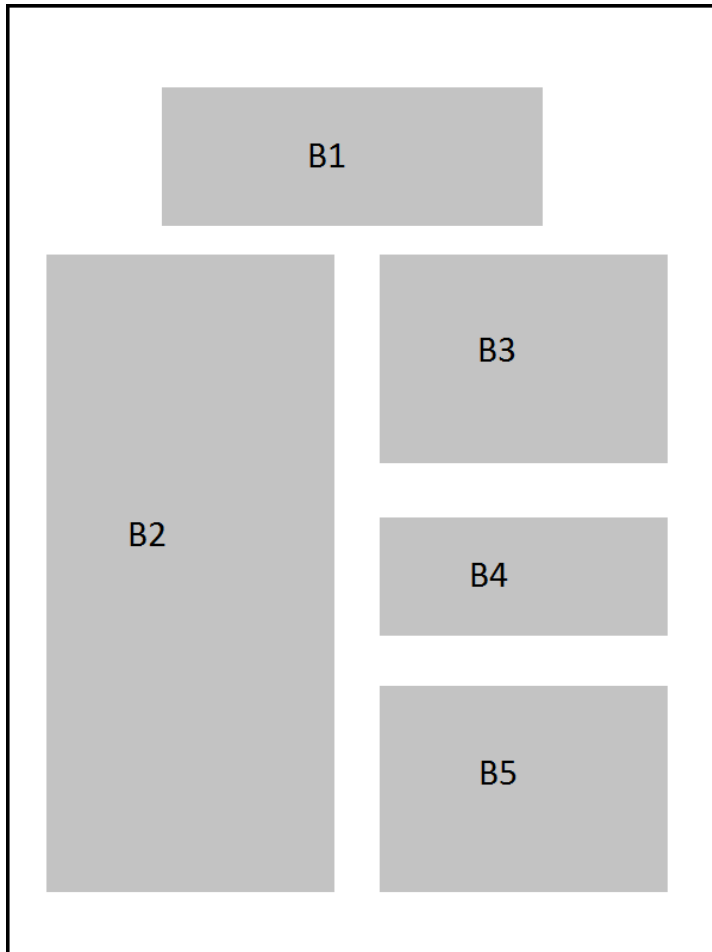
$$\text{Prop.}(B_j) = \text{Prop.}(E_i)$$

(iv) $W = X \setminus \bigcup_1^n B_j$ is called background and $\text{Prop.}(W) \neq \text{Prop.}(E_i)$

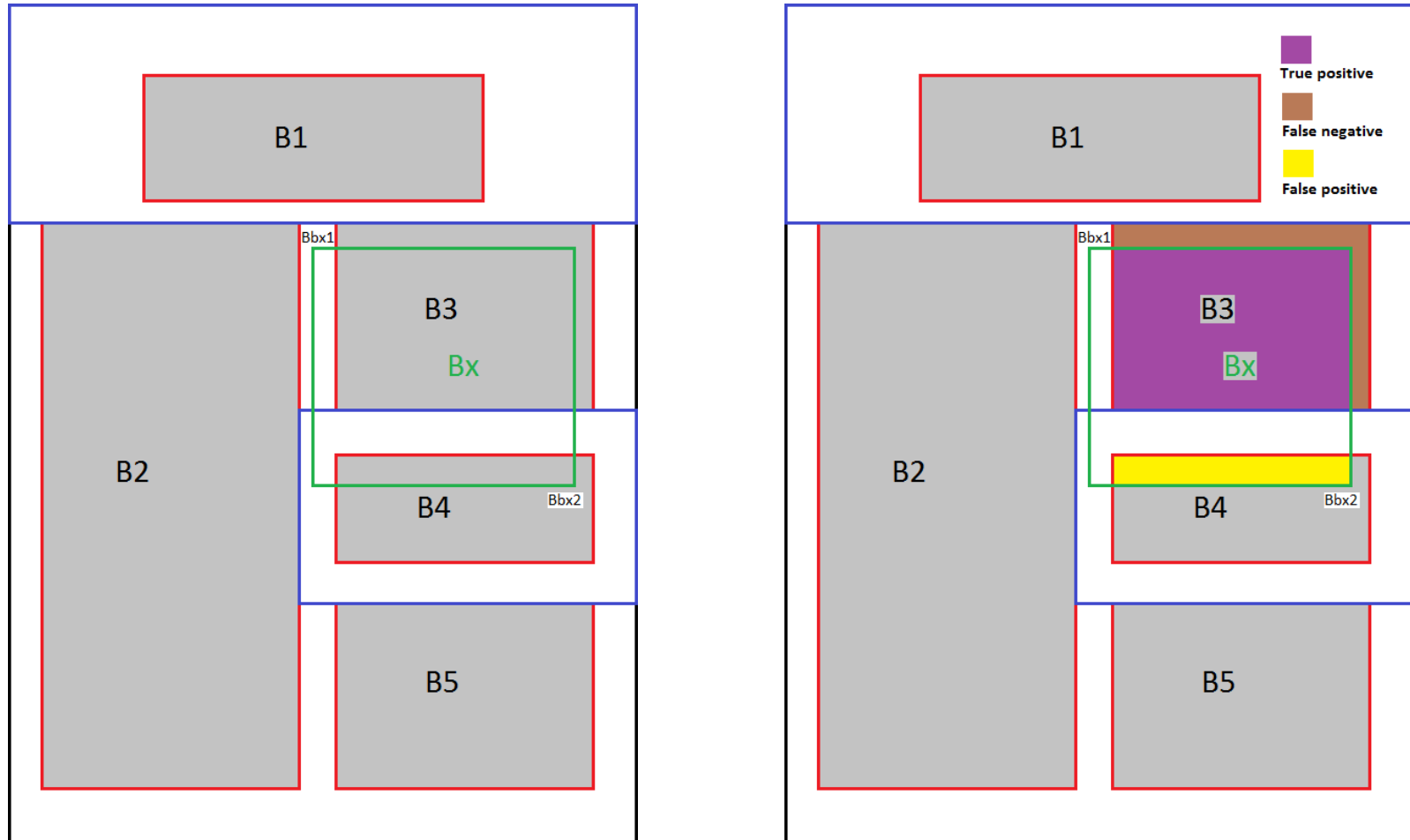
Performance evaluation



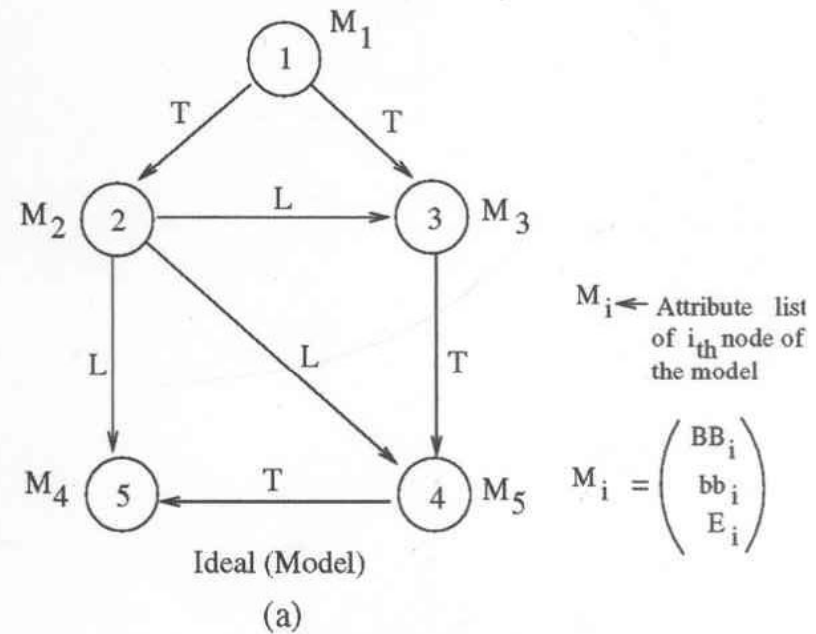
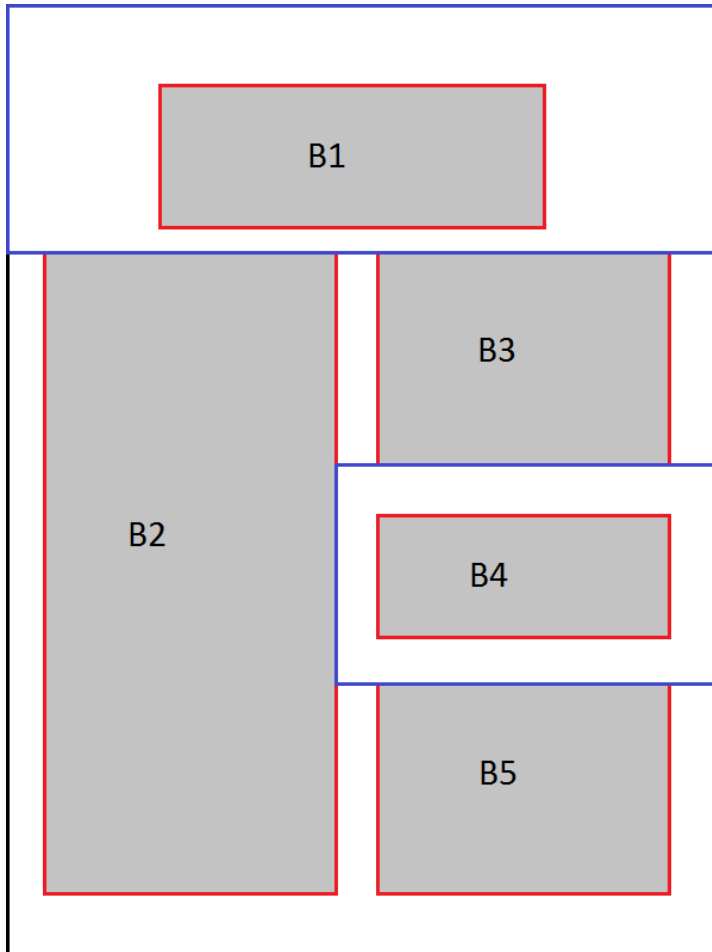
Performance evaluation



Performance evaluation



Performance evaluation



Performance evaluation

- Both model and object graphs are directed acyclic graph.
- Let us represent the model graph by

$$G_M = (V_M, L_M)$$

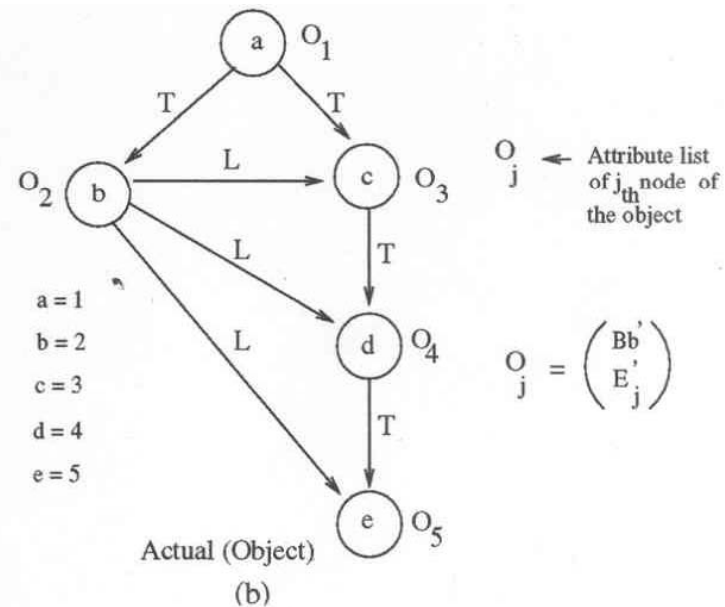
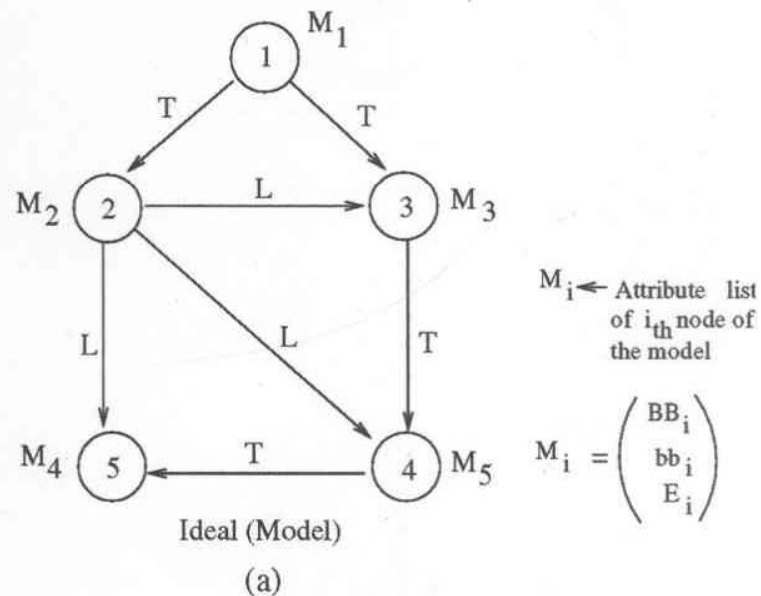
where $V_M = \{M_0, M_1, M_2, \dots, M_n\}$ represents the set of nodes or vertices and L_M represents set of links.

- Note that $M_j = (BB_j, bb_j, E_j)$ and $L_{jk} = (M_j, M_k)$.
- Similarly the object graph is represented by

$$G_o = (V_o, \mathcal{L}_o)$$

- And $O_j = (Bb_j, \mathcal{E}_j)$ and $\mathcal{L}_{jk} = (O_j, O_k)$.
- Finally, graph matching algorithm is employed.

Performance evaluation



Das, Saha and Chanda, An empirical measure of performance of document image segmentation algorithm, *IJDAR*, Vol. 4(3), 2002.

Performance evaluation

- Relation between BB_j and bb_j in model (groundtruth):

$$BB_j \supseteq bb_j \text{ and } BB_j \cap bb_k \text{ for } j \neq k$$

- For good segmentation of object node:

$$bb_j \subseteq Bb_i \subseteq BB_j \text{ if node } M_j \text{ matches node } O_i$$

- The error measure:

(i) Correct classification (*True positive*) = $\#(Bb_j \cap BB_i)$.

(ii) False alarm (*False positive*) = $\#(Bb_j \setminus BB_i)$.

(iii) Mis-classification (*False negative*) = $\#(bb_i \setminus Bb_j)$.

Conclusion

- Presented a document image segmentation method based on shape features
- Used mathematical morphological operators
- Necessary for OCR and data compression
- System is useful for development of digital library providing facilities for electronic storage, searching, navigation

References

- B. Chanda and D. Dutta Majumder, **Digital Image Processing and Analysis**, Prentice Hall of India, New Delhi, 2000.
- A. K. Das and B. Chanda, **A fast algorithm for skew detection of document images using morphology**, Intl. J. Of Document Analysis and Recognition}, Vol.4, pp.109-114, 2001.
- A. K. Das, S. K. Saha and B. Chanda, **An empirical measure of performance of document image segmentation algorithm**, Intl. J. on Document Analysis and Recognition, Vol.4, pp.183-190, 2002.
- S. Mandal, A. K. Das and B. Chanda, **A Simple and Effective Table Detection System from Document Images**, Int. J. on Document Analysis and Recognition, Vol.8, pp.172-182, 2006.

Thank you