The Winner Takes It All: choosing the "best" binarization algorithm for photographed documents

Rafael Dueire Lins[1,2,3], Rodrigo Barros Bernardino[1,3], Ricardo Barboza[3], and Raimundo Oliveira[3]

[1] Universidade Federal de Pernambuco, Recife-PE, 50730-120, Brazil
{rdl,rbb4}@cin.ufpe.br
[2] Universidade Federal Rural de Pernambuco, Recife-PE, 52.171-900, Brazil
[3] Universidade do Estado do Amazonas, Manaus-AM, Brazil
rsbarboza@uea.edu.br, rcorrea.oliveira@gmail.com

Abstract. The recent Time-Quality Binarization Competitions have shown that no single binarization algorithm is good for all kinds of document images and that the time elapsed in binarization varies widely between algorithms and also depends on the document features. On the other hand, document applications for portable devices have space and processing limitations that allow to implement only the "best" algorithm. This paper presents the methodology and assesses the time-quality performance of 61 binarization algorithms to choose the most time-quality efficient one, under two criteria.

Keywords: Smartphone applets · Document binarization · DIB-dataset · Photographed Documents · Binarization competitions.

1 Introduction

Today, half of the population of the world has a smartphone with a built-in digital camera, according with the June 2021 report from the consulting firm Strategy Analytics[4]. Such devices are incredibly versatile and even low cost ones have good quality cameras that allow digitizing document images, that are widely used in a large number of everyday situations that in a recent past photocopying was used.

Binarization, or *thresholding*, is the name given to the conversion process of a color image into its black-and-white (or monochromatic) version. Binary images make most documents more readable and save toner for printing. save storage space [47], communication bandwidth. Binarization also works as a file compression strategy, as the size of binary images is often orders of magnitudes smaller than the original gray or color images. Thresholding is a key preprocessing step for document OCR, classification and indexing. The recent Time-Quality Binarization Competitions [30, 25, 27, 31] have shown that no single binarization

[4] https://www.strategyanalytics.com/

algorithm is good for all kinds of document images and that the time elapsed in binarization varies widely between algorithms and also depends on the document features.

Portable devices are limited in space and users are eager for outputs. Thus, being able to pinpoint which algorithm would fast provide a good quality binary image, capable of being embedded in applications in a certain smartphone model is a valuable information. **This paper assesses 61 binarization algorithms to choose the one that presents the best time-quality trade-off to be implemented in embedded applications in smartphones**. The universe of the tested algorithms is formed by "classical" and recently published binarization algorithms: $Akbari_1$ [1], $Akbari_2$ [1], $Akbari_3$ [1], Bataineh [3], Bernsen [5], Bradley [6], Calvo-Z [7], CLD [41], CNW [40], dSLR [46], DeepOtsu (SL) [12], DiegoPavan (DP) [50], DilatedUNet [27], DocDLink [57], Doc-UNet (WX) [30], ElisaTV [2], $Ergina_G$ [49], $Ergina_L$ [18], Gattal [9], Gosh [4], Howe [13], Huang [14], HuangBCD (AH_1) [28], HuangUnet (AH_2) [28], iNICK (KS_1) [42], Intermodes [38], ISauvola [11], IsoData [53], Jia-Shi [15], Johannsen [16], KSW [17], Li-Tam [22], Lu-Su [32], Mean [10], Mello-Lins [34], Michalak [27], $Michalak21_1$ (MO_1) [28], $Michalak21_2$ (MO_3) [28], $Michalak21_3$ (MO_3) [28], MinError [20], Moments [52], Niblack [36], Nick [19], Otsu [37], Percentile [8], Pun [39], RenyEntropy [43], Sauvola [44], Shanbhag [45], Singh [48], Su-Lu [51], Triangle [56], Vahid (RNB) [28], WAN [35], Wolf [54], Wu-Lu [33], Yen-CC [55], YinYang [27], YinYang21 (JB), [27], Yuleny [30].

The test set used for such an assessment is part of the IAPR DIB dataset (https://dib.cin.ufpe.br) including "real-world" offset, laser, and deskjet printed text documents. Such documents were photographed at two different places, with four different models of smartphones widely used today, with their in-built strobe flash *on* and *off*. The methodology presented here may be used to find the most suitable algorithm for other devices, or the same smartphone models under different setups.

2 Quality-Time Evaluation Methods

Two quality measures were used to evaluate the performance of the 61 binarization algorithms assessed here. The first one, made use of Google Vision to perform Optical Character Recognition (OCR) on the documents and applies the Levenshtein distance ($[L_{dist}]$) to the correct number of characters in the document transcription ($\#\ char$). The error rate is calculated as:

$$([L_{dist}] = (\#char - L_{dist})/\#char.) \tag{1}$$

The second quality measure, P_{err} , compares the proportion between the black-to-white pixels in the scanned and photographed binary documents [24]. One expects that although the photographed and scanned documents have different resolutions, the number of black pixels in a photographed and its scanned version of a document document will be balanced, somehow. Thus,

$$([P_{err}] = 100 * \frac{abs(PW_{GT} - PW_{bin})}{PW_{GT}}) \tag{2}$$

where PW_{GT} and PW_{bin} are the proportion of white pixels in the ground-truth and the binarized image, respectively, and $abs()$ obtains the absolute value of the difference. The quality evaluation was done in the context of each measure separately. They were ranked using the mean value for the whole dataset.

The processing time evaluation provides the order of magnitude of the time elapsed for binarizing the whole datasets. The training-times for the AI-based algorithms were not computed. The processing device was CPU: Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz, with 32GB RAM and a GPU GeForce GTX 1650 4GB The algorithms were implemented using two operating systems and different programming languages, for specific hardware platforms such as GPUs:

- **Windows 10 (version 1909), Matlab:** $Akbari_1$, $Akbari_2$, $Akbari_3$, CLD, CNW, ElisaTV, Ergina-Global, Ergina-Local, Gattal, Ghosh, Howe, iNICK, Jia-Shi, Lu-Su, Michalak, MO_1, MO_2, MO_3
- **Linux Pop!_OS 20.10:**
  - **C++ (GCC 10.3):** Bataineh, Bernsen, ISauvola, Niblack, Nick, Otsu, Sauvola, Singh, Su-Lu, WAN, Wolf
  - **Python 2.7:** SL
  - **Java 14:** YinYang, JB, Bradley, daSilva-Lins-Rocha, Huang, Intermodes, IsoData, Johannsen-Bille, Kapur-SW, Li-Tam, Mean, Mello-Lins, Min-Error, Minimum, Moments, Percentile, Pun, RenyEntropy, Shanbhag, Triangle, Wu-Lu, Yen
  - **Python 3.6:** AH_1, AH_2, Calvo-Z, DP, DilatedUNet, DocDLink, WX, RNB, Yuleny

The algorithms were executed on different operating systems (OS), but on the same hardware. For those that could be executed on both OS types, the processing times for each OS was measured and no significant difference was noticed. This behaviour was also observed and reported in [27]. The mean processing time was used in the analysis. The primary purpose is to provide the order of magnitude time of the processing time elapsed. The **SL** algorithm (DeepOtsu) would take weeks to process the images using a CPU; therefore, a NVIDIA Tesla K80 has been used to accelerate the processing. However, an approximation of the CPU processing time is used as reference in order to compare with the other algorithms, each of which was processed using a CPU on the specific platform.

3 Test Set

Document images acquired using mobile phones are harder to binarize if compared with the use of scanners. The distance between the document and the capturing device and the illumination may vary significantly. Other external light sources and the activation or not of the strobe flash may interfere in the quality

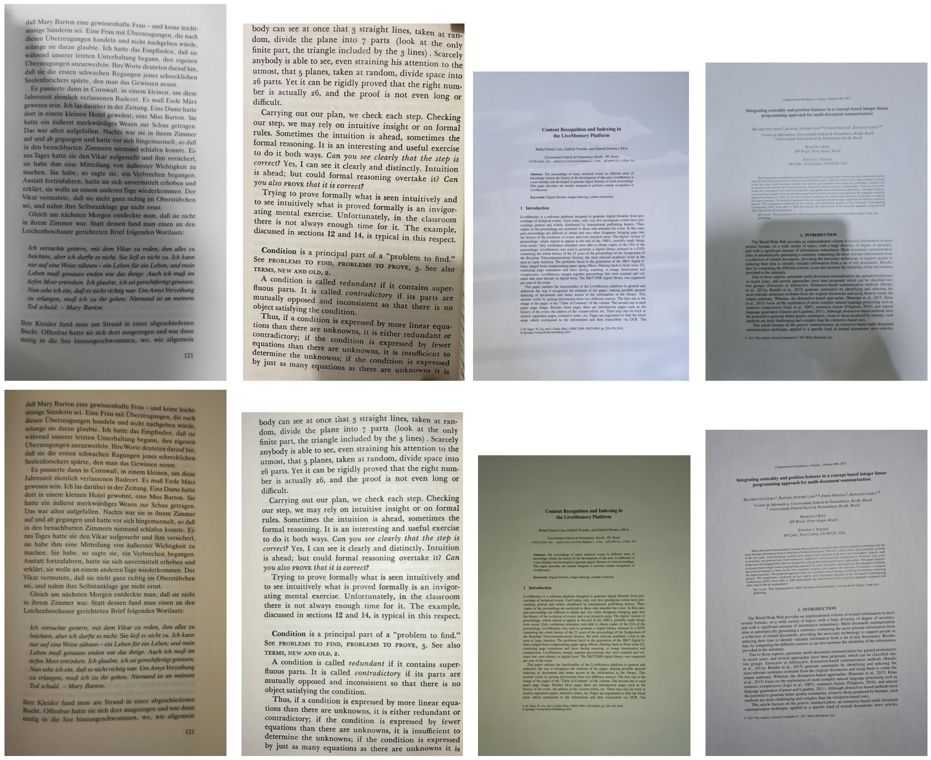


Fig. 1. Samples of the images clustered by device (Motorola G9 Plus, iPhone SE 2, Samsung A10S, Samsung S20) and set-up of the strobe flash (top-line "off", bottom-line "on").

of the obtained image. The kind of document images used here are representative of the kind of images that ordinary people take photos of and correspond to the kinds of documents people often used to take photocopies a few years ago. Typically, such documents are text ones with a plain background, printed in either plain white printer or recycled paper. The test set used here, samples of which are presented in Figure 1, is formed by nine documents offset printed book pages, and deskjet and laser printed documents. Very seldom, people take a photo of a historic documents. If that is the case, such a document image tends not to be binarized in the camera itself as historic documents tend to have a darker background, some show back-to-front interference [29] and physical noises [23]. Such images are part of The IAPR DIB-dataset, which encompasses nine documents obtained from four different models of portable cell-phones, widely used today. Besides the device model, the documents in this set were clustered according to having the in-built strobe-flash set as "*on*" or "*off*".

4 Results

Four models of smartphones of different manufacturers were used is this study. The choice of the devices was made to cover mid-price range models of different manufacturers in such a way to be representative of the smartphones used by the majority of the population. The technical specifications of their front cameras are presented in Table 1.

Table 1. Summary of specifications of the front camera of the devices studied

	Moto G9	**iPhone SE2**	**Galaxy S20**	**Galaxy A10S**
Megapixels	48	12	64	13
Flash	Dual LED	Quad-LED	Dual LED	Dual LED
Aperture	f/1.8	f/1.8	f/2.0	f1.8
Sensor size	1/2 inch	-	1/1.72 inch	-
Pixel size	0.8 m	-	0.8 μm	-

In this paper, for each device tested, three algorithms will be selected:

- (i) the best binarization algorithm for printing, screen reading, or less storage space or claims for less communication bandwidth for transmission.
- (ii) the best for OCR processing binarization algorithm.
- (iii) the overall “winner” - the algorithm that provides the best quality-time trade-off for any sort of binarization application.

The developers of applications for other device models should use the methodology presented here to make a criterious choice of which algorithm(s) to use.

4.1 Motorola Moto G9

The Moto G9 device used in this assessment is an Android 10 smartphone developed by Motorola Mobility[5], a subsidiary of Lenovo. It was first released in August 2020. The analysis of the data presented on table 2 for the document images in the testset acquired with this smartphone, shows that several algorithms perform very well in terms of the quality of the generated monochromatic image. Two of them compete for the podium in analysing the general quality time trade-off: Michalack [27] and MO_1 [25], both developed by Hubert Michalak and Krzysztof Okarma at the West Pomeranian University of Technology, Poland.

- (i) best for printing: **MO_1**: the difference of the P_{err} with flash on and off makes it slightly better than Michalak, as both are top fast among the top quality algorithms.
- (ii) best for OCR: **Michalak**
- (iii) global “winner”: **Michalak**: it is just as fast as MO_1, but has better quality measures than MO_1.

[5] https://www.motorola.com/we/compare-smartphones

Table 2. Best binarization algorithms using Motorola G9 Plus

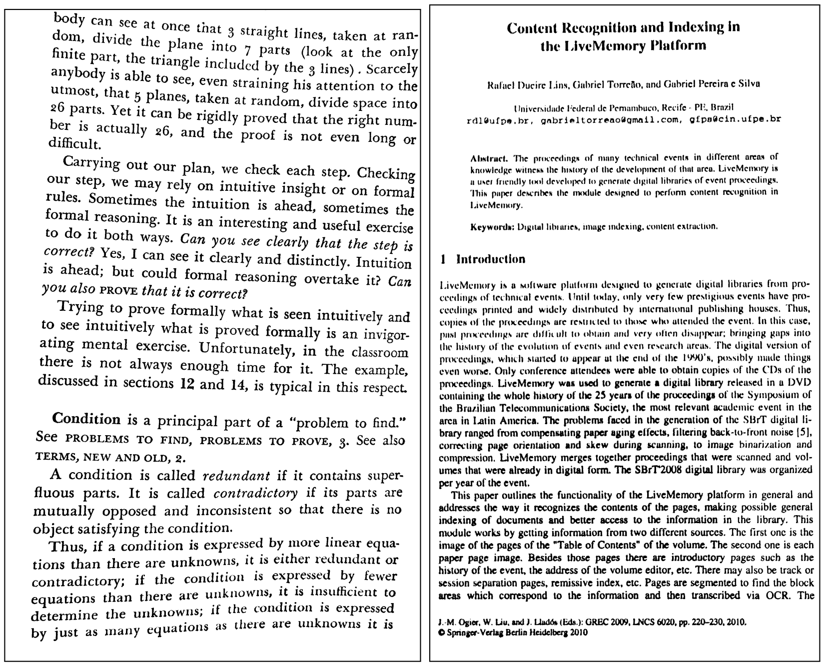
Motolora G9 Plus						
	OFF			ON		
#	**Alg.**	P_{err}	**Time (s)**	**Alg.**	P_{err}	**Time (s)**
1	**Michalak**	0.92	0.06	**KS_1**	0.55	3.42
2	**MO_3**	0.94	1.41	**MO_1**	0.59	0.05
3	**Bradley**	0.95	0.41	**Gosh**	0.70	145.16
4	**MO_1**	0.97	0.06	**Yasin**	0.74	1.75
5	**ElisaTV**	1.06	11.59	**ElisaTV**	0.83	11.2
6	**Yasin**	1.14	2.03	**MO_3**	0.86	1.34
7	**DilatedUNet**	1.17	188.27	**Bradley**	0.91	0.40
8	**MO_2**	1.19	3.09	**Michalak**	0.97	0.05
9	**Gosh**	1.24	143.09	**Singh**	1.00	0.44
10	**WX**	1.25	281.66	**Nick**	1.12	0.21
11	**KS_2**	1.42	3.80	**Su-Lu**	1.22	2.17
12	**DocDLink**	1.43	300.18	**DilatedUNet**	1.24	187.73
13	**KS_1**	1.68	3.72	**Wolf**	1.32	0.29
14	**ISauvola**	1.72	0.53	**WX**	1.64	281.16
15	**Su-Lu**	1.74	2.19	**MO_2**	1.65	3.00
#	**Alg.**	$[L_{dist}]$	**Time (s)**	**Alg.**	$[L_{dist}]$	**Time (s)**
1	**KS_2**	0.98	3.80	**AH_1**	0.98	398.98
2	**MO_3**	0.98	1.41	**AH_2**	0.98	91.2
3	**Bradley**	0.98	0.41	**KS_2**	0.98	3.69
4	**Michalak**	0.98	0.06	**MO_3**	0.98	1.34
5	**RNB**	0.98	46.17	**SL**	0.98	13666.25
6	**WAN**	0.98	1.36	**Michalak**	0.98	0.05
7	**ISauvola**	0.97	0.53	**Bradley**	0.98	0.40
8	**MO_2**	0.97	3.09	**RNB**	0.98	45.58
9	**MO_1**	0.97	0.06	**WAN**	0.97	1.35
10	**ElisaTV**	0.97	11.59	**MO_2**	0.97	3.00
11	**JB**	0.97	1.79	**JB**	0.97	1.73
12	**KS_1**	0.97	3.72	**KS_1**	0.97	3.42
13	**Gosh**	0.97	143.09	**MO_1**	0.97	0.05
14	**YinYang**	0.97	2.08	**ISauvola**	0.97	0.52
15	**Bataineh**	0.97	0.16	**ElisaTV**	0.97	11.2

Figure 2 presents the results of the binarization produced by the top two algorithms for two of the document images produced by the Moto G9 smartphone.

4.2 Samsung A10S

The smartphone Samsung Galaxy A10S [6] was released around August 2019 and became the second top selling device worldwide in December 2019 and it is still

[6] https://www.gsmarena.com/samsung_galaxy_a10s-9793.php



body can see at once that 3 straight lines, taken at random, divide the plane into 7 parts (look at the only finite part, the triangle included by the 3 lines). Scarcely anybody is able to see, even straining his attention to the utmost, that 5 planes, taken at random, divide space into 26 parts. Yet it can be rigidly proved that the right number is actually 26, and the proof is not even long or difficult.

Carrying out our plan, we check each step. Checking our step, we may rely on intuitive insight or on formal rules. Sometimes the intuition is ahead, sometimes the formal reasoning. It is an interesting and useful exercise to do it both ways. *Can you see clearly that the step is correct?* Yes, I can see it clearly and distinctly. Intuition is ahead; but could formal reasoning overtake it? *Can you also* PROVE *that it is correct?*

Trying to prove formally what is seen intuitively and to see intuitively what is proved formally is an invigorating mental exercise. Unfortunately, in the classroom there is not always enough time for it. The example, discussed in sections 12 and 14, is typical in this respect.

Condition is a principal part of a "problem to find." See PROBLEMS TO FIND, PROBLEMS TO PROVE, 3. See also TERMS, NEW AND OLD, 2.

A condition is called *redundant* if it contains superfluous parts. It is called *contradictory* if its parts are mutually opposed and inconsistent so that there is no object satisfying the condition.

Thus, if a condition is expressed by more linear equations than there are unknowns, it is either redundant or contradictory; if the condition is expressed by fewer equations than there are unknowns, it is insufficient to determine the unknowns; if the condition is expressed by just as many equations as there are unknowns it is

Content Recognition and Indexing in the LiveMemory Platform

Rafael Dueire Lins, Gabriel Torreão, and Gabriel Pereira e Silva

Universidade Federal de Pernambuco, Recife - PE, Brazil
rdl@ufpe.br, gabrieltorreao@gmail.com, gfps@cin.ufpe.br

Abstract. The proceedings of many technical events in different areas of knowledge witness the history of the development of that area. LiveMemory is a user friendly tool developed to generate digital libraries of event proceedings. This paper describes the module designed to perform content recognition in LiveMemory.

Keywords: Digital libraries, image indexing, content extraction.

1 Introduction

LiveMemory is a software platform designed to generate digital libraries from proceedings of technical events. Until today, only very few prestigious events have proceedings printed and widely distributed by international publishing houses. Thus, copies of the proceedings are restricted to those who attended the event. In this case, past proceedings are difficult to obtain and very often disappear; bringing gaps into the history of the evolution of events and even research areas. The digital version of proceedings, which started to appear at the end of the 1990's, possibly made things even worse. Only conference attendees were able to obtain copies of the CDs of the proceedings. LiveMemory was used to generate a digital library released in a DVD containing the whole history of the 25 years of the proceedings of the Symposium of the Brazilian Telecommunications Society, the most relevant academic event in the area in Latin America. The problems faced in the generation of the SBrT digital library ranged from compensating paper aging effects, filtering back-to-front noise [5], correcting page orientation and skew during scanning, to image binarization and compression. LiveMemory merges together proceedings that were scanned and volumes that were already in digital form. The SBrT'2008 digital library was organized per year of the event.

This paper outlines the functionality of the LiveMemory platform in general and addresses the way it recognizes the contents of the pages, making possible general indexing of documents and better access to the information in the library. This module works by getting information from two different sources. The first one is the image of the pages of the "Table of Contents" of the volume. The second one is each paper page image. Besides those pages there are introductory pages such as the history of the event, the address of the volume editor, etc. There may also be track or session separation pages, remissive index, etc. Pages are segmented to find the block areas which correspond to the information and then transcribed via OCR. The

J.-M. Ogier, W. Liu, and J. Lladós (Eds.): GREC 2009, LNCS 6020, pp. 220–230, 2010.
© Springer-Verlag Berlin Heidelberg 2010

Fig. 2. Result of the binarization with $\mathbf{MO}_1$ algorithm of an offset printed **book** page with the strobeflash **on** (left) and, using **Michalak**, a **deskjet** printed document with the stobeflash **off** (right), both acquired using the **Motorola Moto G9**

on sale today [7] It originally runs an Android 9.0 (Pie), upgradable to Android 11, One UI 3.1. The two assessments made here with the 61 binarization algorithms yielded the data shown on Table 3 for the top 15 algorithms and allow to point as global results:

- (i) best for printing and transmitting: **Michalak** – it has the best P_{err} either with flash on or off
- (ii) best for OCR: **Michalak** – it is the fastest among the smallest $[L_{dist}]$, with value 0.98
- (iii) overall winner: **Michalak** – it is the best either for OCR or printing and transmitting applications

The result of the binarization of two of the test images in the dataset used here processed by Michalak algorithm may be seen in Figure 3.

4.3 Samsung S20

The Samsung Galaxy S20 is another Android-based smartphone designed and manufactured by Samsung. It is the successor model to the successful Galaxy S10

[7] https://www.91mobiles.com/hub/best-selling-phone-q3-2019-iphone-xr-11-samsung-galaxy-a10-a50/?pid=33347

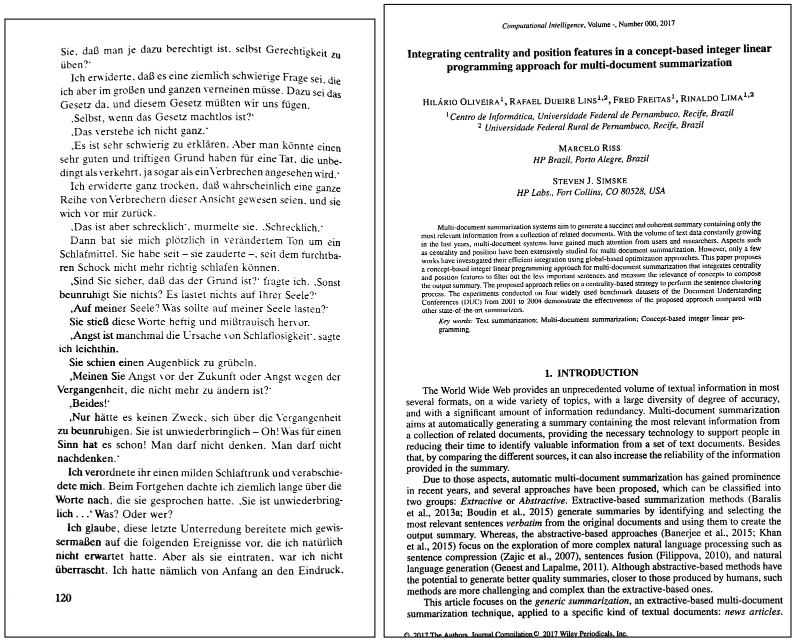
Table 3. Best binarization algorithms using Samsung A10S

Samsung A10S						
	OFF			ON		
#	**Alg.**	P_{err}	**Time (s)**	**Alg.**	P_{err}	**Time (s)**
1	**Michalak**	0.76	0.05	**Michalak**	0.76	0.03
2	**MO_2**	0.91	1.95	**MO_2**	0.91	1.86
3	**MO_1**	0.92	0.04	**MO_1**	0.92	0.03
4	**MO_3**	0.92	0.87	**MO_3**	0.92	0.8
5	**Bradley**	0.94	0.24	**Bradley**	0.94	0.24
6	**Bernsen**	1.06	1.98	**Bernsen**	1.06	1.96
7	**ElisaTV**	1.16	6.13	**ElisaTV**	1.16	6.09
8	**DocDLink**	1.24	173.78	**Yasin**	1.24	1.29
9	**Yasin**	1.24	1.46	**DocDLink**	1.24	173.34
10	**ISauvola**	1.25	0.31	**ISauvola**	1.25	0.31
11	**Gosh**	1.27	80.84	**Gosh**	1.27	80.66
12	**Howe**	1.32	37.38	**Howe**	1.32	37.27
13	**WX**	1.35	174.81	**WX**	1.35	174.31
14	**Wolf**	1.38	0.18	**Wolf**	1.38	0.18
15	**KS_2**	1.4	3.26	**KS_2**	1.4	3.31
#	**Alg.**	$[L_{dist}]$	**Time (s)**	**Alg.**	$[L_{dist}]$	**Time (s)**
1	**RNB**	0.98	27.77	**RNB**	0.98	27.86
2	**KS_2**	0.98	3.26	**AH_2**	0.98	56.78
3	**ElisaTV**	0.98	6.13	**KS_2**	0.98	3.31
4	**JB**	0.98	1.24	**ElisaTV**	0.98	6.09
5	**ISauvola**	0.98	0.31	**JB**	0.98	1.23
6	**Bradley**	0.98	0.24	**ISauvola**	0.98	0.31
7	**AH_2**	0.98	59.22	**AH_1**	0.98	257.38
8	**$Akbari_1$**	0.98	15.27	**Bradley**	0.98	0.24
9	**Jia-Shi**	0.98	15.19	**$Akbari_1$**	0.98	15.18
10	**MO_3**	0.98	0.87	**Jia-Shi**	0.98	15.22
11	**Michalak**	0.98	0.05	**MO_3**	0.98	0.8
12	**WAN**	0.98	0.82	**Michalak**	0.98	0.03
13	**KS_1**	0.97	3.49	**WAN**	0.98	0.83
14	**YinYang**	0.97	1.41	**KS_1**	0.97	3.38
15	**Gosh**	0.97	80.84	**SL**	0.97	11627.4

and it was released on 11 February 2020 [21]. The analysis of the data presented on table 4 allows one to pinpoint the "best" algorithms in the in terms of image quality-time and OCR-performance and time, and the overall "winner" as:

- (i) best for printing: **MO_1**.
- (ii) best for OCR: **Michalak**
- (iii) the overall "winner": **MO_1**, P_{err} and the L_{dist} are reasonably small.

Figure 4 shows the monochromatic version of two of the images in this dataset.



Sie, daß man je dazu berechtigt ist, selbst Gerechtigkeit zu üben?'

Ich erwiderte, daß es eine ziemlich schwierige Frage sei, die ich aber im großen und ganzen verneinen müsse. Dazu sei das Gesetz da, und diesem Gesetz müßten wir uns fügen.

‚Selbst, wenn das Gesetz machtlos ist?'

‚Das verstehe ich nicht ganz.'

‚Es ist sehr schwierig zu erklären. Aber man könnte einen sehr guten und triftigen Grund haben für eine Tat, die unbedingt als verkehrt, ja sogar als ein Verbrechen angesehen wird.'

Ich erwiderte ganz trocken, daß wahrscheinlich eine ganze Reihe von Verbrechern dieser Ansicht gewesen seien, und sie wich vor mir zurück.

‚Das ist aber schrecklich', murmelte sie. ‚Schrecklich.'

Dann bat sie mich plötzlich in verändertem Ton um ein Schlafmittel. Sie habe seit – sie zauderte –, seit dem furchtbaren Schock nicht mehr richtig schlafen können.

‚Sind Sie sicher, daß das der Grund ist?' fragte ich. ‚Sonst beunruhigt Sie nichts? Es lastet nichts auf Ihrer Seele?'

‚Auf meiner Seele? Was sollte auf meiner Seele lasten?'

Sie stieß diese Worte heftig und mißtrauisch hervor.

‚Angst ist manchmal die Ursache von Schlaflosigkeit', sagte ich leichthin.

Sie schien einen Augenblick zu grübeln.

‚Meinen Sie Angst vor der Zukunft oder Angst wegen der Vergangenheit, die nicht mehr zu ändern ist?'

‚Beides!'

‚Nur hätte es keinen Zweck, sich über die Vergangenheit zu beunruhigen. Sie ist unwiederbringlich – Oh! Was für einen Sinn hat es schon! Man darf nicht denken. Man darf nicht nachdenken.'

Ich verordnete ihr einen milden Schlaftrunk und verabschiedete mich. Beim Fortgehen dachte ich ziemlich lange über die Worte nach, die sie gesprochen hatte. ‚Sie ist unwiederbringlich ...' Was? Oder wer?

Ich glaube, diese letzte Unterredung bereitete mich gewissermaßen auf die folgenden Ereignisse vor, die ich natürlich nicht erwartet hatte. Aber als sie eintraten, war ich nicht überrascht. Ich hatte nämlich von Anfang an den Eindruck,

120

Computational Intelligence, Volume -, Number 000, 2017

Integrating centrality and position features in a concept-based integer linear programming approach for multi-document summarization

Hilário Oliveira[1], Rafael Dueire Lins[1,2], Fred Freitas[1], Rinaldo Lima[1,2]

[1] *Centro de Informática, Universidade Federal de Pernambuco, Recife, Brazil*
[2] *Universidade Federal Rural de Pernambuco, Recife, Brazil*

Marcelo Riss
HP Brazil, Porto Alegre, Brazil

Steven J. Simske
HP Labs., Fort Collins, CO 80528, USA

Multi-document summarization systems aim to generate a succinct and coherent summary containing only the most relevant information from a collection of related documents. With the volume of text data constantly growing in the last years, multi-document systems have gained much attention from users and researchers. Aspects such as centrality and position have been extensively studied for multi-document summarization. However, only a few works have investigated their efficient integration using global-based optimization approaches. This paper proposes a concept-based integer linear programming approach for multi-document summarization that integrates centrality and position features to filter out the less important sentences and measure the relevance of concepts to compose the output summary. The proposed approach relies on a centrality-based strategy to perform the sentence clustering process. The experiments conducted on four widely used benchmark datasets of the Document Understanding Conferences (DUC) from 2001 to 2004 demonstrate the effectiveness of the proposed approach compared with other state-of-the-art summarizers.

Key words: Text summarization; Multi-document summarization; Concept-based integer linear programming.

1. INTRODUCTION

The World Wide Web provides an unprecedented volume of textual information in most several formats, on a wide variety of topics, with a large diversity of degree of accuracy, and with a significant amount of information redundancy. Multi-document summarization aims at automatically generating a summary containing the most relevant information from a collection of related documents, providing the necessary technology to support people in reducing their time to identify valuable information from a set of text documents. Besides that, by comparing the different sources, it can also increase the reliability of the information provided in the summary.

Due to those aspects, automatic multi-document summarization has gained prominence in recent years, and several approaches have been proposed, which can be classified into two groups: *Extractive* or *Abstractive*. Extractive-based summarization methods (Baralis et al., 2013a; Boudin et al., 2015) generate summaries by identifying and selecting the most relevant sentences *verbatim* from the original documents and using them to create the output summary. Whereas, the abstractive-based approaches (Banerjee et al., 2015; Khan et al., 2015) focus on the exploration of more complex natural language processing such as sentence compression (Zajic et al., 2007), sentences fusion (Filippova, 2010), and natural language generation (Genest and Lapalme, 2011). Although abstractive-based methods have the potential to generate better quality summaries, closer to those produced by humans, such methods are more challenging and complex than the extractive-based ones.

This article focuses on the *generic summarization*, an extractive-based multi-document summarization technique, applied to a specific kind of textual documents: *news articles*.

© 2017 The Authors. Journal Compilation © 2017 Wiley Periodicals, Inc.

Fig. 3. Result of the binarization with **Michalak** algorithm of an offset printed **book** page with the strobeflash **on** (left) and a **deskjet** printed document with the strobeflash **off** (right), adcquired using the **Samsung A10S**

4.4 Apple iPhone SE

The second-generation iPhone SE (also known as the iPhone SE 2 or the iPhone SE 2020) is a smartphone designed and developed by Apple Inc. It was released on April, 2020 and became one of the top selling smartphone models in 2020 (24.2 million devices sold).[8] It continues today as one of the top sold mid-price devices. Table 5 presents the results for the assessment of this dataset. As one can see, several "classical" binarization algorithms appear high-up in the quality P_{err} rank, with very efficient time figures. Taking the formula $Perr_{off}\ x\ T_{off} + Perr_{on}\ x\ T_{on}$ as the way to decide the best algorithm for printing, MO_1 appears top with 0.0972, closely followed by Michalak (0.1194) and Otsu (0,1316). Thus, in this category the winner is MO_1. The global results are:

- (i) best for printing and transmitting: **MO_1**
- (ii) the best for OCR: **MO_1**
- (iii) the overall winner: **MO_1**

Figure 5 presents two of the imagesin this dataset binarized with the overall "winner".

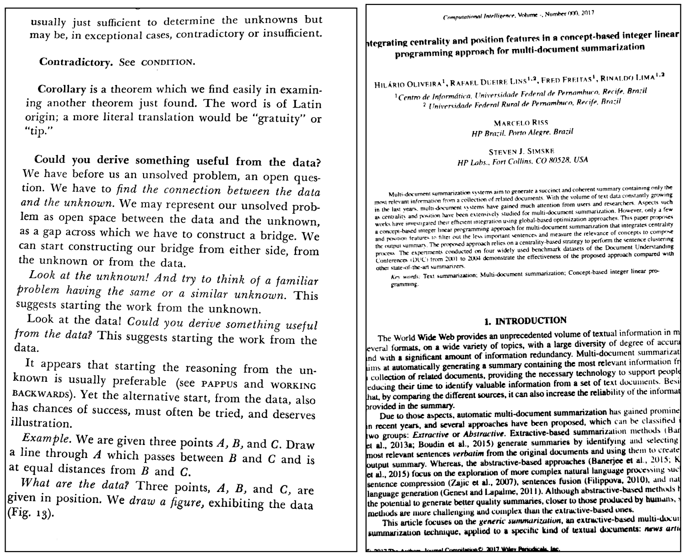
[8] https://www.gizmochina.com/2021/02/25/most-shipped-smartphones-2020-omdia/

Table 4. Best binarization algorithms using Samsung S20

Samsung S20						
	OFF			ON		
#	Alg.	P_{err}	Time (s)	Alg.	P_{err}	Time (s)
1	MO_1	0.91	0.05	Gattal	0.66	55.68
2	MO_3	0.92	1.09	IsoData	0.72	0.13
3	Bradley	0.96	0.31	Otsu	0.74	0.02
4	Michalak	0.99	0.05	MO_1	0.79	0.04
5	DilatedUNet	1.06	151.65	Li-Tam	0.84	0.13
6	WX	1.13	279.6	Yasin	0.92	1.47
7	Howe	1.26	49.79	Gosh	0.95	102.95
8	DocDLink	1.27	228.22	MO_3	0.96	0.98
9	Gosh	1.28	120.9	ElisaTV	0.97	7.46
10	KS_1	1.28	3.79	Wolf	1.02	0.22
11	Wolf	1.28	0.23	KS_1	1.05	3.39
12	Yasin	1.28	1.75	Michalak	1.05	0.04
13	Singh	1.29	0.34	Bradley	1.05	0.29
14	MO_2	1.33	2.49	Singh	1.06	0.32
15	Nick	1.37	0.16	$Ergina_L$	1.06	0.62
#	Alg.	$[L_{dist}]$	Time (s)	Alg.	$[L_{dist}]$	Time (s)
1	MO_3	0.98	1.09	$Ergina_G$	0.98	0.44
2	RNB	0.98	36.34	KSW	0.98	0.13
3	KS_2	0.98	3.47	Yen-CC	0.98	0.13
4	Michalak	0.98	0.05	Bradley	0.98	0.29
5	ISauvola	0.98	0.41	MO_3	0.98	0.98
6	JB	0.98	1.43	SL	0.98	10319.87
7	Bradley	0.98	0.31	ElisaTV	0.98	7.46
8	WAN	0.98	1.07	IsoData	0.98	0.13
9	ElisaTV	0.98	7.68	Wolf	0.98	0.22
10	Bataineh	0.98	0.12	Su-Lu	0.98	1.62
11	YinYang	0.98	1.64	AH_2	0.98	72.09
12	DocDLink	0.97	228.22	RNB	0.98	34.71
13	MO_1	0.97	0.05	AH_1	0.98	319.31
14	MO_2	0.97	2.49	RenyEntropy	0.98	0.13
15	AH_2	0.97	75.01	MO_1/ Michalak	0.98	0.04

5 Conclusions

Smartphones have drastically changed the way of life of people worldwide with their omnipresence, growing computational power and high-quality embedded cameras. Photographing documents is now a simple way of digitizing everyday documents and book pages for later referencing and even meeting legal requirements in many countries. Document binarization plays a key role in many document processing pipelines, besides yielding smaller documents for storing and sending via networks better readable and more economic to print. Recent document binarization competitions [30] [27] [27] [31] show that no single algorithm is the best for all kinds of documents. Each smartphone model has a camera with



usually just sufficient to determine the unknowns but may be, in exceptional cases, contradictory or insufficient.

Contradictory. See CONDITION.

Corollary is a theorem which we find easily in examining another theorem just found. The word is of Latin origin; a more literal translation would be "gratuity" or "tip."

Could you derive something useful from the data? We have before us an unsolved problem, an open question. We have to *find the connection between the data and the unknown*. We may represent our unsolved problem as open space between the data and the unknown, as a gap across which we have to construct a bridge. We can start constructing our bridge from either side, from the unknown or from the data.

Look at the unknown! And try to think of a familiar problem having the same or a similar unknown. This suggests starting the work from the unknown.

Look at the data! *Could you derive something useful from the data?* This suggests starting the work from the data.

It appears that starting the reasoning from the unknown is usually preferable (see PAPPUS and WORKING BACKWARDS). Yet the alternative start, from the data, also has chances of success, must often be tried, and deserves illustration.

Example. We are given three points *A*, *B*, and *C*. Draw a line through *A* which passes between *B* and *C* and is at equal distances from *B* and *C*.

What are the data? Three points, *A*, *B*, and *C*, are given in position. We *draw a figure*, exhibiting the data (Fig. 13).

Computational Intelligence, Volume -, Number 000, 2017

ntegrating centrality and position features in a concept-based integer linear programming approach for multi-document summarization

HILÁRIO OLIVEIRA[1], RAFAEL DUEIRE LINS[1,2], FRED FREITAS[1], RINALDO LIMA[1,2]
[1]*Centro de Informática, Universidade Federal de Pernambuco, Recife, Brazil*
[2] *Universidade Federal Rural de Pernambuco, Recife, Brazil*

MARCELO RISS
HP Brazil, Porto Alegre, Brazil

STEVEN J. SIMSKE
HP Labs., Fort Collins, CO 80528, USA

Multi-document summarization systems aim to generate a succinct and coherent summary containing only the most relevant information from a collection of related documents. With the volume of text data constantly growing in the last years, multi-document systems have gained much attention from users and researchers. Aspects such as centrality and position have been extensively studied for multi-document summarization. However, only a few works have investigated their efficient integration using global-based optimization approaches. This paper proposes a concept-based integer linear programming approach for multi-document summarization that integrates centrality and position features to filter out the less important sentences and measure the relevance of concepts to compose the output summary. The proposed approach relies on a centrality-based strategy to perform the sentence clustering process. The experiments conducted on four widely used benchmark datasets of the Document Understanding Conferences (DUC) from 2001 to 2004 demonstrate the effectiveness of the proposed approach compared with other state-of-the-art summarizers.

Key words: Text summarization; Multi-document summarization; Concept-based integer linear programming.

1. INTRODUCTION

The World Wide Web provides an unprecedented volume of textual information in m
everal formats, on a wide variety of topics, with a large diversity of degree of accura
nd with a significant amount of information redundancy. Multi-document summarizat
ims at automatically generating a summary containing the most relevant information fr
collection of related documents, providing the necessary technology to support people
educing their time to identify valuable information from a set of text documents. Besi
hat, by comparing the different sources, it can also increase the reliability of the informat
provided in the summary.

Due to those aspects, automatic multi-document summarization has gained promine
n recent years, and several approaches have been proposed, which can be classified
wo groups: *Extractive* or *Abstractive*. Extractive-based summarization methods (Bar
et al., 2013a; Boudin et al., 2015) generate summaries by identifying and selecting
most relevant sentences *verbatim* from the original documents and using them to create
output summary. Whereas, the abstractive-based approaches (Banerjee et al., 2015; K
et al., 2015) focus on the exploration of more complex natural language processing suc
sentence compression (Zajic et al., 2007), sentences fusion (Filippova, 2010), and nat
language generation (Genest and Lapalme, 2011). Although abstractive-based methods h
the potential to generate better quality summaries, closer to those produced by humans,
methods are more challenging and complex than the extractive-based ones.

This article focuses on the *generic summarization*, an extractive-based multi-docu
summarization technique, applied to a specific kind of textual documents: *news arti*

© 2017 The Authors. Journal Compilation © 2017 Wiley Periodicals, Inc.

Fig. 4. Result of the binarization with $\mathbf{MO}_1$ algorithm of an offset printed **book** page with the strobe flash **on** (left) and a **deskjet** printed document with the stobeflash **off** (right), acquired using the **Samsung S20**

different features making the binarization of photographed document images a challenging task.

Applications that run on smartphones need to be light due to the hardware limitations of a device that needs to execute several processes simultaneously. Thus a binarization algorithm to be used in an embedded smartphone application must have an excellent quality-time balance. This paper presents a methodology to choose such an algorithm. Four popular smartphone models of three different manufacturers were quality-time assessed using 61 of the possibly best binarization algorithms of today, pointing out the "best" algorithm for printing, the "best" algorithm for OCR applications, and the global "winner" for each of those devices.

The recent paper [26] shows that feeding binarization algorithms with the image, their RGB-components or the grayscale converted image yield to differences in their quality-time performance. That analysis would multiply the number of the assessed algorithms by five, thus is left as one of the lines for further work.

Acknowledgements

The authors grateful to all those who made their binarization code available.

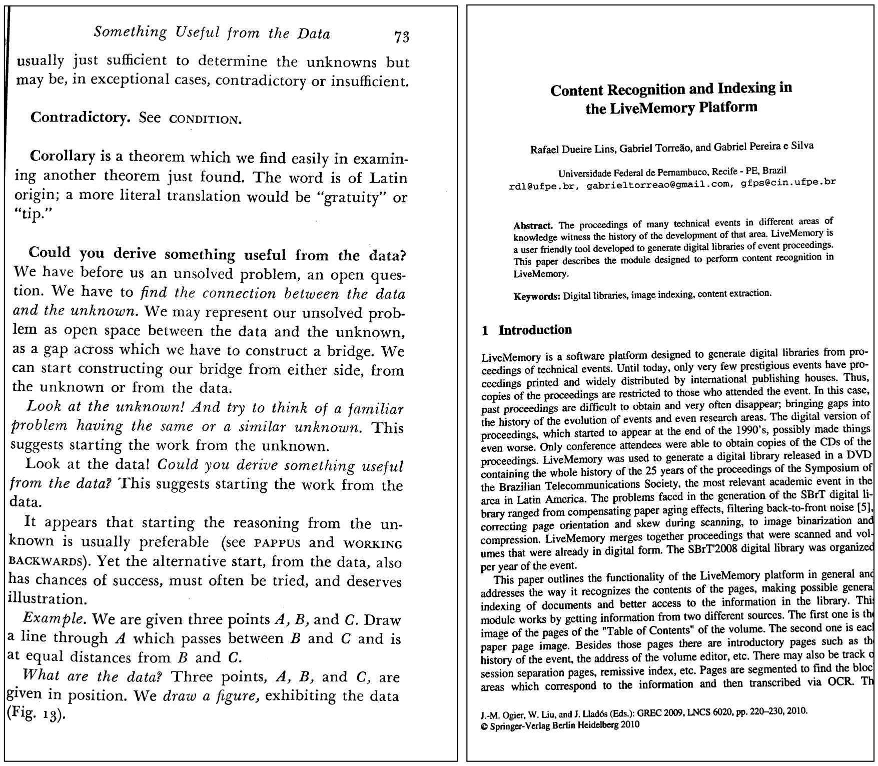
The research reported in this paper was mainly sponsored by the RD&I project between the Universidade do Estado do Amazonas and Transire Eletrônicos and Tec Toy S.A. through the Lei de Informática/SUFRAMA. Rafael Dueire Lins was also partly sponsored by CNPq – Brazil.

Table 5. Best binarization algorithms using Apple iPhone SE

Apple iPhone SE 2						
	OFF			ON		
#	Alg.	P_{err}	Time (s)	Alg.	P_{err}	Time (s)
1	Yasin	0.72	1.96	IsoData	0.60	0.12
2	Nick	0.79	0.17	Otsu	0.60	0.02
3	Sauvola	0.79	0.17	Sauvola	0.73	0.18
4	Singh	0.79	0.30	Gattal	0.74	54.59
5	Gosh	0.79	88.74	Gosh	0.77	85.64
6	JB	0.88	1.27	Yasin	0.81	1.55
7	YinYang	0.94	1.70	MO_1	0.81	0.04
8	Wolf	0.95	0.23	Singh	0.81	0.29
9	KS_1	0.96	4.23	Wolf	0.84	0.24
10	ElisaTV	1.04	5.00	Nick	0.84	0.17
11	Su-Lu	1.04	1.77	JB	0.85	1.27
12	MO_1	1.08	0.06	ElisaTV	0.90	3.44
13	KS_3	1.21	4.70	YinYang	0.94	1.78
14	Michalak	1.31	0.06	Michalak	1.02	0.04
15	Bradley	1.36	0.34	KS_1	1.03	3.30
#	Alg.	$[L_{dist}]$	Time (s)	Alg.	$[L_{dist}]$	Time (s)
1	KS_1	0.98	4.23	YinYang	0.98	1.78
2	$Akbari_1$	0.98	21.76	SL	0.98	10,310.89
3	Jia-Shi	0.98	20.74	Yasin	0.97	1.55
4	Singh	0.98	0.30	KS_2	0.97	3.39
5	Wolf	0.98	0.23	Singh	0.97	0.29
6	Wu-Lu	0.98	0.13	Nick	0.97	0.17
7	Bataineh	0.98	0.13	KS_3	0.97	4.65
8	AH_1	0.98	277.31	Bataineh	0.97	0.13
9	ElisaTV	0.98	5.00	RNB	0.97	33.9
10	Calvo-Z	0.98	9.83	$Ergina_G$	0.97	0.43
11	MO_2	0.98	2.56	Howe	0.97	55.39
12	RNB	0.98	33.45	Li-Tam	0.97	0.13
13	Nick	0.98	0.17	MO_2	0.97	2.28
14	MO_1	0.98	0.06	$Ergina_L$	0.97	0.59
15	Bradley	0.98	0.34	DocDLink	0.97	191.72
37	Yen-CC	0.97	0.13	MO_1	0.97	0.04

References

1. Akbari, Y., Britto Jr., A.S., Al-Maadeed, S., Oliveira, L.S.: Binarization of Degraded Document Images using Convolutional Neural Networks based on predicted Two-Channel Images. In: ICDAR (2019)
2. Barney Smith, E.H., Likforman-Sulem, L., Darbon, J.: Effect of Pre-processing on Binarization. In: Document Recognition and Retrieval XVII. p. 75340H (2010)
3. Bataineh, B., Abdullah, S.N.H.S., Omar, K.: An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows. Pattern Recognition Letters **32**(14), 1805–1813 (2011)



Something Useful from the Data 73

usually just sufficient to determine the unknowns but may be, in exceptional cases, contradictory or insufficient.

Contradictory. See CONDITION.

Corollary is a theorem which we find easily in examining another theorem just found. The word is of Latin origin; a more literal translation would be "gratuity" or "tip."

Could you derive something useful from the data? We have before us an unsolved problem, an open question. We have to *find the connection between the data and the unknown.* We may represent our unsolved problem as open space between the data and the unknown, as a gap across which we have to construct a bridge. We can start constructing our bridge from either side, from the unknown or from the data.

Look at the unknown! And try to think of a familiar problem having the same or a similar unknown. This suggests starting the work from the unknown.

Look at the data! *Could you derive something useful from the data?* This suggests starting the work from the data.

It appears that starting the reasoning from the unknown is usually preferable (see PAPPUS and WORKING BACKWARDS). Yet the alternative start, from the data, also has chances of success, must often be tried, and deserves illustration.

Example. We are given three points *A, B,* and *C.* Draw a line through *A* which passes between *B* and *C* and is at equal distances from *B* and *C.*

What are the data? Three points, *A, B,* and *C,* are given in position. We *draw a figure,* exhibiting the data (Fig. 13).

Content Recognition and Indexing in the LiveMemory Platform

Rafael Dueire Lins, Gabriel Torreão, and Gabriel Pereira e Silva

Universidade Federal de Pernambuco, Recife - PE, Brazil
rdl@ufpe.br, gabrieltorreao@gmail.com, gfps@cin.ufpe.br

Abstract. The proceedings of many technical events in different areas of knowledge witness the history of the development of that area. LiveMemory is a user friendly tool developed to generate digital libraries of event proceedings. This paper describes the module designed to perform content recognition in LiveMemory.

Keywords: Digital libraries, image indexing, content extraction.

1 Introduction

LiveMemory is a software platform designed to generate digital libraries from proceedings of technical events. Until today, only very few prestigious events have proceedings printed and widely distributed by international publishing houses. Thus, copies of the proceedings are restricted to those who attended the event. In this case, past proceedings are difficult to obtain and very often disappear; bringing gaps into the history of the evolution of events and even research areas. The digital version of proceedings, which started to appear at the end of the 1990's, possibly made things even worse. Only conference attendees were able to obtain copies of the CDs of the proceedings. LiveMemory was used to generate a digital library released in a DVD containing the whole history of the 25 years of the proceedings of the Symposium of the Brazilian Telecommunications Society, the most relevant academic event in the area in Latin America. The problems faced in the generation of the SBrT digital library ranged from compensating paper aging effects, filtering back-to-front noise [5], correcting page orientation and skew during scanning, to image binarization and compression. LiveMemory merges together proceedings that were scanned and volumes that were already in digital form. The SBrT'2008 digital library was organized per year of the event.

This paper outlines the functionality of the LiveMemory platform in general and addresses the way it recognizes the contents of the pages, making possible genera indexing of documents and better access to the information in the library. Thi module works by getting information from two different sources. The first one is th image of the pages of the "Table of Contents" of the volume. The second one is eac paper page image. Besides those pages there are introductory pages such as th history of the event, the address of the volume editor, etc. There may also be track o session separation pages, remissive index, etc. Pages are segmented to find the bloc areas which correspond to the information and then transcribed via OCR. Th

J.-M. Ogier, W. Liu, and J. Lladós (Eds.): GREC 2009, LNCS 6020, pp. 220–230, 2010.
© Springer-Verlag Berlin Heidelberg 2010

Fig. 5. Result of the binarization with $\mathbf{MO}_1$ algorithm of an offset printed **book** page with the strobeflash **on** (left) and, a **deskjet** printed document with the stobeflash **off** (right), acquired using the **iPhone SE 2**

4. Bera, S.K., Ghosh, S., Bhowmik, S., Sarkar, R., Nasipuri, M.: A non-parametric binarization method based on ensemble of clustering algorithms. Multimedia Tools and Applications **80**(5), 7653–7673 (2021)
5. Bernsen, J.: Dynamic thresholding of gray-level images. In: International Conference on Pattern Recognition. pp. 1251–1255 (1986)
6. Bradley, D., Roth, G.: Adaptive Thresholding using the Integral Image. Journal of Graphics Tools **12**(2), 13–21 (2007)
7. Calvo-Zaragoza, J., Gallego, A.J.: A selectional auto-encoder approach for document image binarization. Pattern Recognition **86**, 37–47 (2019)
8. Doyle, W.: Operations Useful for Similarity-Invariant Pattern Recognition. Journal of the ACM **9**(2), 259–267 (1962)
9. Gattal, A., Abbas, F., Laouar, M.R.: Automatic Parameter Tuning of K-Means Algorithm for Document Binarization. In: 7th ICSENT. pp. 1–4. ACM Press (2018)
10. Glasbey, C.: An Analysis of Histogram-Based Thresholding Algorithms. Graphical Models and Image Processing **55**(6), 532–537 (1993)
11. Hadjadj, Z., Meziane, A., Cherfa, Y., Cheriet, M., Setitra, I.: ISauvola: Improved Sauvola's Algorithm for Document Image Binarization. pp. 737–745. LNCS (3212)

12. He, S., Schomaker, L.: DeepOtsu: Document Enhancement and Binarization using Iterative Deep Learning. Pattern Recognition **91**, 379–390 (jan 2019)
13. Howe, N.R.: Document binarization with automatic parameter tuning. IJDAR **16**(3), 247–258 (2013)
14. Huang, L.K., Wang, M.J.J.: Image thresholding by minimizing the measures of fuzziness. Pattern Recognition **28**(1), 41–51 (1995)
15. Jia, F., Shi, C., He, K., Wang, C., Xiao, B.: Degraded document image binarization using structural symmetry of strokes. Pattern Recognition **74**, 225–240 (2018)
16. Johannsen, G and Bille, J.: A threshold selection method using information measures. In: Int'l Conf. Pattern Recognition. pp. 140–143 (1982)
17. Kapur, J., Sahoo, P., Wong, A.: A new method for gray-level picture thresholding using the entropy of the histogram. C. Vision, Graphics, I. Processing **29**(1), 140
18. Kavallieratou, E., Stathis, S.: Adaptive binarization of historical document images. Proceedings - International Conference on Pattern Recognition **3**, 742–745 (2006)
19. Khurshid, K., Siddiqi, I., Faure, C., Vincent, N.: Comparison of Niblack inspired binarization methods for ancient documents. In: SPIE. p. 72470U (2009)
20. Kittler, J., Illingworth, J.: Minimum error thresholding. Pattern Recognition **19**(1), 41–47 (1986)
21. Knapp, M.: Samsung Galaxy S20 price and deals: here's where you can get it in the US. In: TechRadar. pp. Retrieved 2020–02–12 (2020)
22. Li, C., Tam, P.: An iterative algorithm for minimum cross entropy thresholding. Pattern Recognition Letters **19**(8), 771–776 (1998)
23. Lins, R.D.: A Taxonomy for Noise in Images of Paper Documents - The Physical Noises. In: Lecture Notes in Computer Science, vol. 5627 LNCS, pp. 844–854 (2009)
24. Lins, R.D., Almeida, M.M.D., Bernardino, R.B., Jesus, D., Oliveira, J.M.: Assessing binarization techniques for document images. In: DocEng'17. pp. 183–192
25. Lins, R.D., Bernardino, R.B., Barney Smith, E., Kavallieratou, E.: ICDAR 2019 Time-Quality Binarization Competition. In: ICDAR. pp. 1539–1546
26. Lins, R.D., Bernardino, R.B., da Silva Barboza, R., Lins, Z.D.: Direct binarization a quality-and-time efficient binarization strategy. In: Proceedings of the 21st ACM Symposium on Document Engineering. DocEng '21, ACM (2021)
27. Lins, R.D., Bernardino, R.B., Simske, S.J.: DocEng'2020 Time-Quality Competition on Binarizing Photographed Documents. In: DocEng'20. pp. 1–4. ACM (2020)
28. Lins, R.D., Bernardino, R.B., Smith, E.B., Kavallieratou, E.: ICDAR 2021 Competition on Time-Quality Document Image Binarization. In: ICDAR 2021 Competition on Time-Quality Document Image Binarization. pp. 708–722 (2021)
29. Lins, R.D., Guimarães Neto, M., França Neto, L., Galdino Rosa, L.: An environment for processing images of historical documents. Microprocessing and Microprogramming **40**(10-12), 939–942 (1994)
30. Lins, R.D., Kavallieratou, E., Barney Smith, E., Bernardino, R.B., de Jesus, D.M.: ICDAR 2019 Time-Quality Binarization Competition. In: ICDAR. pp. 1539–1546
31. Lins, R.D., Simske, S.J., Bernardino, R.B.: DocEng'2021 Time-Quality Competition on Binarizing Photographed Documents. In: DocEng'21. pp. 1–4. ACM (2020)
32. Lu, S., Su, B., Tan, C.L.: Document image binarization using background estimation and stroke edges. IJDAR **13**(4), 303–314 (2010)
33. Lu, W., Songde, M., Lu, H.: An effective entropic thresholding for ultrasonic images. 14th ICPR pp. 1552–1554, vol. 2 (1998)
34. Mello, C.A.B., Lins, R.D.: Image segmentation of historical documents. Visual 2000 (2000)
35. Mustafa, W.A., Abdul Kader, M.M.M.: Binarization of Document Image Using Optimum Threshold Modification. J. Physics: C. Series **1019**(1), 012022 (2018)

36. Niblack, W.: An introduction to digital image processing. Strandberg Publishing Company (1985)
37. Otsu, N.: A threshold selection method from gray-level histograms. IEEE T. on Systems, Man, and Cybernetics **9**(1), 62–66 (1979)
38. Prewitt, J.M.S., Mendelsohn, M.L.: The Analysis of Cell Images. Annals of the New York Academy of Sciences **128**(3), 1035–1053 (2006)
39. Pun, T.: Entropic thresholding, a new approach. Computer Graphics and Image Processing **16**(3), 210–239 (1981)
40. Saddami, K., Afrah, P., Mutiawani, V., Arnia, F.: A New Adaptive Thresholding Technique for Binarizing Ancient Document. In: INAPR. pp. 57–61. IEEE (2018)
41. Saddami, K., Munadi, K., Away, Y., Arnia, F.: Effective and fast binarization method for combined degradation on ancient documents. Heliyon (2019)
42. Saddami, K., Munadi, K., Muchallil, S., Arnia, F.: Improved Thresholding Method for Enhancing Jawi Binarization Performance. In: ICDAR. vol. 1, pp. 1108–1113. IEEE (2017)
43. Sahoo, P., Wilkins, C., Yeager, J.: Threshold selection using Renyi's entropy. Pattern Recognition **30**(1), 71–84 (1997)
44. Sauvola, J., Pietikäinen, M., Pietikainem, M.: Adaptive document image binarization. Pattern Recognition **33**(2), 225–236 (2000)
45. Shanbhag, A.G.: Utilization of Information Measure as a Means of Image Thresholding. CVGIP: Graphical Models and Image Processing **56**(5), 414–419 (1994)
46. Silva, J.M.M., Lins, R.D., Rocha, V.C.: Binarizing and Filtering Historical Documents with Back-to-Front Interference. In: ACM SAC 2006. pp. 853–858 (2006)
47. da Silva, J.M.M., Lins, R.D.: Color Document Synthesis as a Compression Strategy. Icdar (Icdar), 466–470 (2007)
48. Singh, T.R., Roy, S., Singh, O.I., Sinam, T., Singh, K.M.: A New Local Adaptive Thresholding Technique in Binarization. IJCSI International Journal of Computer Science Issues **08**(6), 271–277 (2011)
49. Sokratis, V., Kavallieratou, E., Paredes, R., Sotiropoulos, K.: A Hybrid Binarization Technique for Document Images. In: Studies in Computational Intelligence, pp. 165–179 (2011)
50. Souibgui, M.A., Kessentini, Y.: DE-GAN: A Conditional Generative Adversarial Network for Document Enhancement. T. on Pattern Analysis and Machine Intelligence pp. 1–1 (2021)
51. Su, B., Lu, S., Tan, C.L.: Binarization of historical document images using the local maximum and minimum. In: 8th IAPR DAS. pp. 159–166. ACM Press (2010)
52. Tsai, W.H.: Moment-preserving thresolding: A new approach. Computer Vision, Graphics, and Image Processing **29**(3), 377–393 (1985)
53. Velasco, F.R.: Thresholding Using the Isodata Clustering Algorithm. Tech. rep., OSD or Non-Service DoD Agency (1979)
54. Wolf, C., Jolion, J.M., Chassaing, F.: Text localization, enhancement and binarization in multimedia documents. In: Object recognition supported by user interaction for service robots. vol. 2, pp. 1037–1040. IEEE Comput. Soc (2003)
55. Yen J. C.; Chang, F.J.C.S., Yen, J.C., Chang, F.J., Chang, S.: A New Criterion for Automatic Multilevel Thresholding. T. on Image Processing **4**(3), 370–378 (1995)
56. Zack, G.W., Rogers, W.E., Latt, S.A.: Automatic measurement of sister chromatid exchange frequency. J. Histochemistry and Cytochemistry **25**(7), 741–753 (1977)
57. Zhou, L., Zhang, C., Wu, M.: D-linknet: Linknet with pretrained encoder and dilated convolution for high resolution satellite imagery road extraction. In: Computer Vision and Pattern Recognition (2018)