

IV ANNUAL INTERNATIONAL WORKSHOP ON
COMPUTER SCIENCE AND EDUCATION IN COMPUTER SCIENCE

June 6–9, 2008, Borovets

Words retrieval from text images

Nikolay Kirov

Department Informatics, New Bulgarian University

Institute of Mathematics and Informatics, BAS

Contents

- Introduction
- Segmentation
- Hausdorff type distances
- Point distances
- Experiments
- The software
- Conclusion

Introduction

The main question in this paper is: How to find a word in a text document? When the document is represented as text file, the answer is quite trivial – open the file in any text editor, choose a word and push Find button. But the task is not so easy when the document is a set of graphic images. This is natural situation when we deal with digitization of cultural and scientific heritage and scanner devices produce files in graphic formats.

Optical character recognition (OCR) is the usual way of conducting text retrieval from scanned document images. OCR software converts text images into a text file, recognizing every letter and mapping it to a number, which is called code. This technique is well developed and has high accuracy. And then we apply the previous algorithm.

Here is some obstacles to successful OCR:

- The quality of page images;
- Language dependency (alphabet and coding, unknown language):
 - dictionaries;
 - old grammar, obsolete words, phrases and idioms;
 - old letters, out of the coding tables;
 - multi-lingual documents;
- Errors in automatic OCR, human intervention needed.

de la Terre, &c.

fort peu de sens commun.

Quant à la Terre, si vous la rencontrez bonne, ce vous sera un grand avantage, & une grande épargne ; mais rarement en pourrez-vous trouver, où il n'y ait beaucoup à travailler, d'autant que telle paraîtra passablement bonne au dessus, qui étant ouverte de la profondeur d'un fer de Béche seulement, se trouvera Argileuse dessous ; ce fonds est pire aux Arbres que le Tuf, ou la Roche, à cause qu'il s'y rencontre de petites veines où les Racines peuvent s'étendre & profonder, afin de tirer la fraîcheur de plus bas, & prendre quelque nourriture ; mais l'Argileuse ou Terre franche ou rouge, fait comme un plancher qui par sa dureté & densité, ne peut être percé par aucunes Racines, & qui dans les grandes ardeurs de l'Eté, em-

il)uanrГ¤, la l'erre, ^! vous la rencomrex t,onne,ce vous lera un ^ranГ¤ avanra^e,sc u/B»e ^ranГ¤B« epar^ne ; maiz raremenr en pourrex-vouz trouvex, on il n've air tieaucouri a travailler , claur^nc ^ue teile riarorra rrallГ¤llernenr cionne au clelluz,c^am ouverre cle la rirokonleur Г¤'un iec Г¤e Lecrre ieulemenr , re rrouveral ^r?l-leule cleclouz ; ce fonclz eli riire aux ^r-Gtrez c^ue le l"uf, on la l^t,cke,a caule c^u'il z'^ reuconrre cle r/erirez veuez on lez ^ .acinez peuvenr z'ccenГ¤resc rirofoncler,arlB« cle rirer la ir,ucleurcle^luz baz,sc rirenГ¤re ^uelc^ue nourricure; maiz l^r^ilcule ou 'l'erre francerre ou rou^e, fair comme un plancrer c^ui riar ia Г¤urere sc clenirre , ne peur ecre rierce riar aucunez li.acinez, sc c^ui cl^nz lez ^r,inclez arcleurz cle l'^,te,emB»

We suggest a different approach: instead of applying two steps – OCR and searching in text documents, we try searching words directly in scanned text documents.

We can organize retrieval of words, similar to a given pattern word, searching in the binary text images.

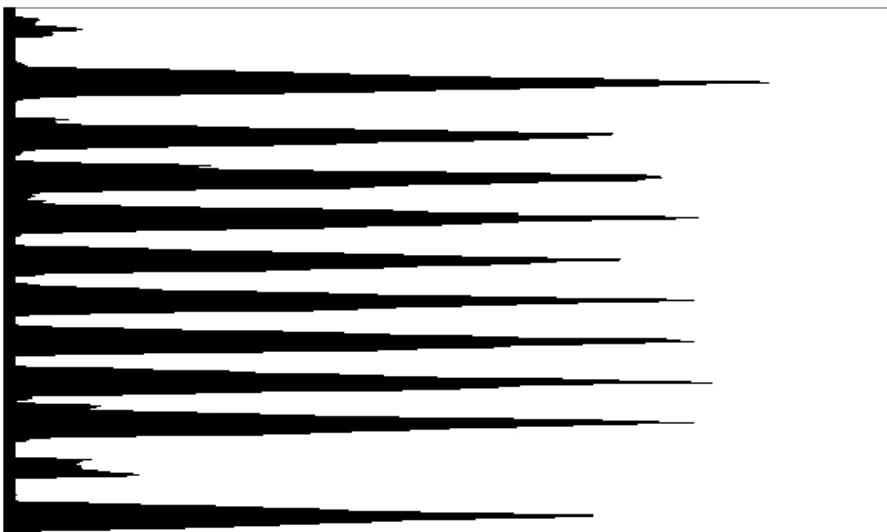
Three main steps in the process: segmentation, search and result representation.

- In the segmentation step we create so-called word images – every word is encompassed by a rectangle, which consist of white and black pixels.
- For measuring similarities between word images we use Hausdorff type distances.
- We produce a sequence of words, sorted by values of similarity measure.

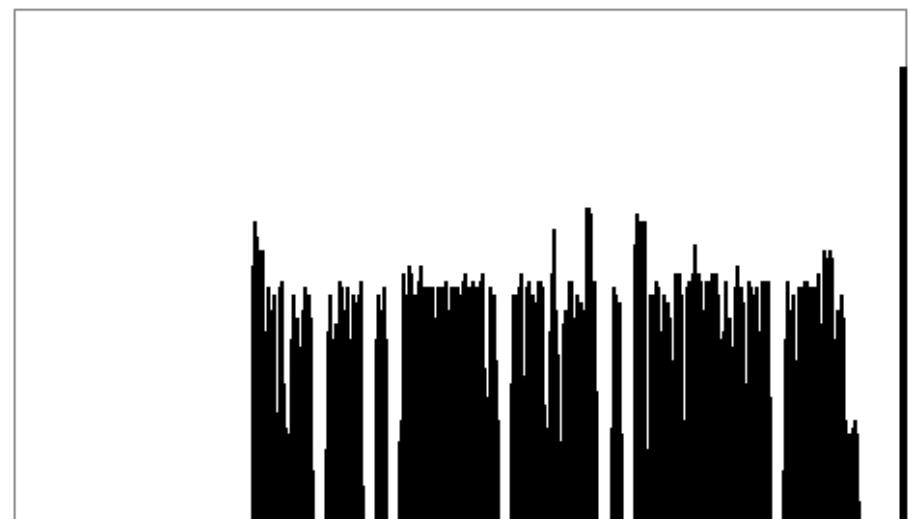
Segmentation

We use horizontal projection for line extraction. If the lines are horizontal, the histogram has near zero values between lines.

Vertical projection histogram is obtained by counting the number of black pixels in each vertical scan at a given horizontal position. If the characters are well separated, the histogram should have zero values between the characters.



The horizontal histogram of a page



The vertical histogram of a row

For segmentation step we use parameters p_1, p_2, p_3, p_4 , which are important for successful word separation.

- The height of every row must be at least p_1 . This helps us to avoid creating (due to noise) rows with small height;
- When the value at a point in row histogram is less than p_2 , we suppose that this point belongs to white space between words.
- The length of white space between words must be greater than p_3 .
- After separation we try to shrink the frame decreasing the height of rectangle if the points of histograms have values less than p_4 .

на пред близки негови клиенти, които много обичали да го
слушат. Тремолирането на дясната му ръка е било неиздадено

на пред близки негови клиенти, които много обичали да го
слушат. Тремолирането на дясната му ръка е било неиздадено

Hausdorff type distances

The Hausdorff distance (HD) between two point sets A and B is defined as

$$H(A, B) = \max\{h(A, B), h(B, A)\}, \quad (1)$$

where $h(A, B)$ and $h(B, A)$ are so-called directed distances between the sets. For the original Hausdorff metric

$$h(A, B) = \max_{a \in A} d(a, B), \text{ where } d(a, B) = \min_{b \in B} \rho(a, b) \quad (2)$$

is the distance from a point a to the set B and $\rho(a, b)$ is any point distance. For image matching a number of modifications of $h(A, B)$ have been introduced by many authors. Dubuisson and Jain consider so-called Modified Hausdorff Distance, one of the best method for word search in the text images. They replaced $h(A, B)$ by

$$h_{\text{MHD}}(A, B) = \frac{1}{N_A} \sum_{a \in A} d(a, B) = \frac{1}{N_A} \sum_{a \in A} \min_{b \in B} \rho(a, b), \quad (3)$$

where N_A is number of points in the set A .

The directed distance $h_M(A, B)$ for M-estimation Hausdorff Distance (M-HD) defined by Dong-Gyu Sim *et al.* is

$$h_M(A, B) = \frac{1}{N_A} \sum_{a \in A} f_\tau(d(a, B)), \quad (4)$$

where the function $f_\tau : R_+ \rightarrow R_+$ can be

$$f_\tau(x) = \begin{cases} x & \text{if } x \leq \tau \\ \tau & \text{if } x > \tau \end{cases}$$

or $f_\tau(x) = \min\{x, \tau\}, \tau > 0$.

This means that we sum the distances $d(a, B)$ which are less than the constant τ and add τ when the distance is greater than τ .

The recommended value of τ is in the interval [3, 5].

Point distances

Let $a = (a_x, a_y)$ and $b = (b_x, b_y)$ be two points on the plane.
The infinity (or l_∞) norm distance is

$$\rho_{max}(a, b) = \max\{|a_x - b_x|, |a_y - b_y|\}.$$

Manhattan (or l_1 norm) distance is

$$\rho_1(a, b) = |a_x - b_x| + |a_y - b_y|.$$

5	5	5	5	5	5	5	5	5	5	5	5
5	4	4	4	4	4	4	4	4	4	4	5
5	4	3	3	3	3	3	3	3	3	4	5
5	4	3	2	2	2	2	2	3	4	5	
5	4	3	2	1	1	1	1	2	3	4	5
5	4	3	2	1	0	1	2	3	4	5	
5	4	3	2	1	1	1	1	2	3	4	5
5	4	3	2	2	2	2	2	3	4	5	
5	4	3	3	3	3	3	3	3	3	4	5
5	4	4	4	4	4	4	4	4	4	4	5
5	5	5	5	5	5	5	5	5	5	5	5

ρ_{max}

					5											
						5	4	5								
						5	4	3	4	5						
						5	4	3	2	3	4	5				
						5	4	3	2	1	2	3	4	5		
						5	4	3	2	1	0	1	2	3	4	5
						5	4	3	2	1	2	3	4	5		
						5	4	3	2	3	4	5				
						5	4	3	4	5						
							5	4	5							
								5								

ρ_1

Well known Euclidean (or l_2 norm) distance is

$$\rho_2(a, b) = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2}.$$

Because $\rho_{max}(a, b) \leq \rho_2(a, b) \leq \rho_1(a, b)$ it is useful to define the following combined distance

$$\rho_c(a, b) = \frac{\rho_1(a, b) + \rho_{max}(a, b)}{2}.$$

					5						
	5.00	4.47	4.12	4	4.12	4.47	5.00				
5.00	4.24	3.61	3.16	3	3.16	3.61	4.24	5.00			
4.47	3.61	2.83	2.24	2	2.24	2.83	3.61	4.47			
4.12	3.16	2.24	1.41	1	1.41	2.24	3.16	4.12			
5	4	3	2	1	0	1	2	3	4	5	
4.12	3.16	2.24	1.41	1	1.41	2.24	3.16	4.12			
4.47	3.61	2.83	2.24	2	2.24	2.83	3.61	4.47			
5.00	4.24	3.61	3.16	3	3.16	3.61	4.24	5.00			
	5.00	4.47	4.12	4	4.12	4.47	5.00				
				5							

ρ_2

					5						
			5.0	4.5	4	4.5	5.0				
		4.5	4.0	3.5	3	3.5	4.0	4.5			
	5.0	4.0	3.0	2.5	2	2.5	3.0	4.0	5.0		
	4.5	3.5	2.5	1.5	1	1.5	2.5	3.5	4.5		
5	4	3	2	1	0	1	2	3	4	5	
	4.5	3.5	2.5	1.5	1	1.5	2.5	3.5	4.5		
	5.0	4.0	3.0	2.5	2	2.5	3.0	4.0	5.0		
	4.5	4.0	3.5	3	3.5	4.0	4.5				
		5.0	4.5	4	4.5	5.0					
				5							

ρ_c

We define 0-1 distance

$$\rho_{01}(a, b) = \begin{cases} 0 & \text{if } a \equiv b \\ 1 & \text{otherwise} \end{cases} \quad (5)$$

is also a metric in the plane. This distance is called bounded because $\rho_{01}(a, b) \leq 1 < \infty$ for every $a, b \in R^2$.

We define also bounded variants of the other distances

$$\rho^{(\tau)}(a, b) = f_\tau(\rho(a, b)) = \min\{\rho(a, b), \tau\},$$

where τ is a fixed positive number.

MHD with any bounded point distance coincides with M-HD.

Experiments

We carried out our experiments using an old book (1884) – Иван Вазов и Константин Величков, Българска христоматия или сборник от избрани образци по всичките родове съчинения.

The quality of scanned images are quite bad because this was one of the first books, processing in the digitization center and operators' qualification was not on appropriate level. Many pages have slopes in lines, there are significant variations in gray levels, etc.

There is no text version till now of this book, which may be produced using OCR software. The first reason is the quality of images. The second reason is the absence of appropriate OCR software because the text contains old and abandoned Bulgarian letters. Also spelling and grammar are quite different in modern Bulgarian language.

We used 200 pages from 953 book's pages scanned at a resolution of 200 DPI. The images are about 2300×3600 pixels (8.28 MPixels), 14.8 x 23.3 cm, grayscale 256 (8 BitsPerPixel).

поетъ, сатирикъ и публицистъ. Първо-то нѣшто, което е издалъ е книжка стихотворения „Басненникъ“ и по-послѣ „Смѣсна Китка“ (Букурештъ 1852 г.), съ които той доби първа-та си известност у насъ, като български писателъ. Отъ 1857 год. се почева негова та многополезна дѣятелност въ борба-та ни съ Грѣци-тѣ за черковна независимостъ. Той дохожда въ Цариградъ и издава свои-тѣ „Смѣшни Календари“ сатирически книги, въ които бичува съ единъ искусенъ и ядовитъ сарказмъ пороци-тѣ и недостатки-тѣ на тогавашно-то българско общество, и грѣцко-то високо духовенство (1857—1863). На 1863 год. той прѣдприе издаваніе-то на сатирический вѣстникъ „Гайда,“ който не трая много време. Доста хубави статии все въ полемично-сатирический духъ, напечата той тамъ. Слѣдъ двѣ години Славейковъ прѣдприе издание-то на политический вѣстникъ „Македония“ (1867—1870). Тамъ при разискваніе-то на разни въпроси отъ общественъ и черковенъ интересъ Славейковъ се стараеше да разбуди народно-то чувство у Македонски-тѣ Българе, които душеше нетърпимо-то влияние на грѣкоманство-то и фанариотство-то. Най-послѣ подиръ нѣколко временни спирания и конфискации на вѣстникъ-тѣ, правительство-то съ всѣмъ го унищожи и запрѣти на Славейкова да издава вече какъвъ-да-е вѣстникъ, а и него самаго тури въ тѣмница, по обвинение, че въ послѣдни-тѣ броеве на „Македония“ явно проповѣдавъ революционни идеи между Българе-тѣ.

- The goal of our experiments is to compare practically the efficiency of searching, counting the number of correctly retrieved words in a sequence of words, sorted by values of similarity measure.
- For all experiments the same segmentation parameters are used.
- We choose a pattern word and then measure similarities between it and the words with approximately same length.

For our experiments we choose a pattern word **всички**. There are two relative words (derivatives) of this word **всичка** and **всичко**. We count as correct all three of them.

Dependance of resolution:

Освѣнь това, еъставителе-тѣ
въ нея всички-тѣ добри :

Original resolution, grayscale

Освѣнь това, еъставителе-тѣ
въ нея всички-тѣ добри :

Half resolution

Освѣнь това, еъставителе-тѣ
въ нея всички-тѣ добри :

Original resolution, b/w

Освѣнь това, еъставителе-тѣ
въ нея всички-тѣ добри :

Quarter resolution

	$n =$	100	200	300	400	500
Original	2340×3672 Pixels	100	177	205	213	220
Half	1170×1836 Pixels	99	178	197	206	206
Quarter	585×918 Pixels	93	139	157	168	174

Resolution, $\rho_1^{(5)}$

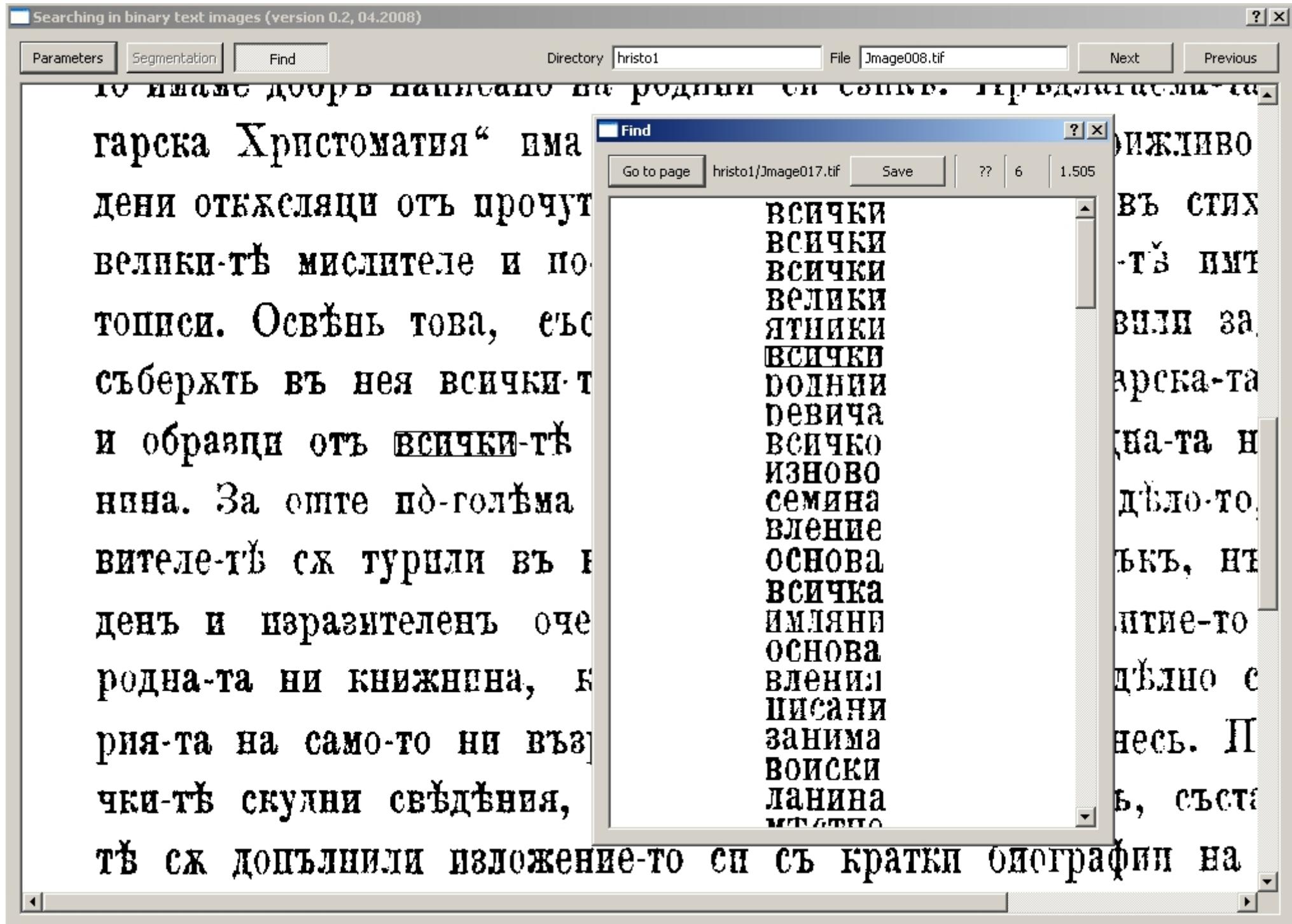
In the Table we count the number of correctly retrieved words among first 100, 200, 300, 400, 500 words with approximately same length.

Dependance of point distance:

$n =$	100	200	300	400	500
$\rho_{max}^{(5)}$	96	165	188	198	201
$\rho_1^{(5)}$	97	165	189	199	202
$\rho_2^{(5)}$	98	165	189	200	202
ρ_{01}	97	165	189	199	201

Point distances, 1170 × 1836 Pixels

In the Table we count the number of correctly retrieved words among first 100, 200, 300, 400, 500 words with approximately same length.



Conclusions

- Experiments give us the certainty in practical benefit of our approach. It is quite universal and does not required any specific features of the concrete language.
- Higher resolution does not involve essentially better results. In our example the middle row of the table ensure good retrieval combined with relatively small execution time.
- Our results for using a number of point distances manifest that all cases are practically identical.

Bulgarian typewritten document (about 1940), 335 pages, tif (2400×3200), 1 BPP

I РАЗДЕЛ

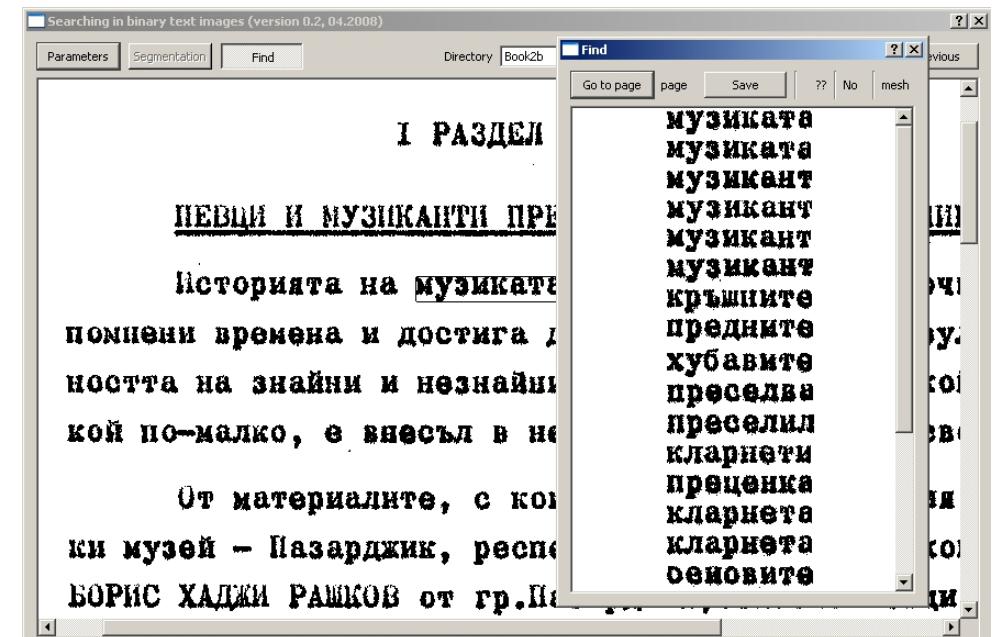
ПЕВЦИ И МУЗИКАНТИ ПРЕДИ И СЛЕД ОСВОБОЖДЕНИЕТО

Историята на музиката в гр.Пазарджик започва от незапомнени времена и достига до наши дни, като резултат от дейността на знайни и незнайни труженици, които, кой повече, кой по-малко, е внесъл в нейната съкровищница своя дал.

От материалите, с които разполага Окръжния исторически музей – Пазарджик, респективно сведенията, които е събрал БОРИС ХАДЖИ РАШКОВ от гр.Пазарджик, относно певци и музиканти преди и след Освобождението се установява, че битовите нужди свързани с годежи, сватби, занаятчийско-еснафски сбирки, хора, вечеринки и пр. са били задоволявани от музиканти – професионалисти и любители.

Професионалисти били онези музиканти-инструменталисти или певци, като най-често инструменталиста е бил и певец, които са свирели и пеели срещу възнаграждение, а любители-онези, които със своето пеене и свирене са радвали душите и сърдата на хората по сборове, хорища и др., без да получават възнаграждение.

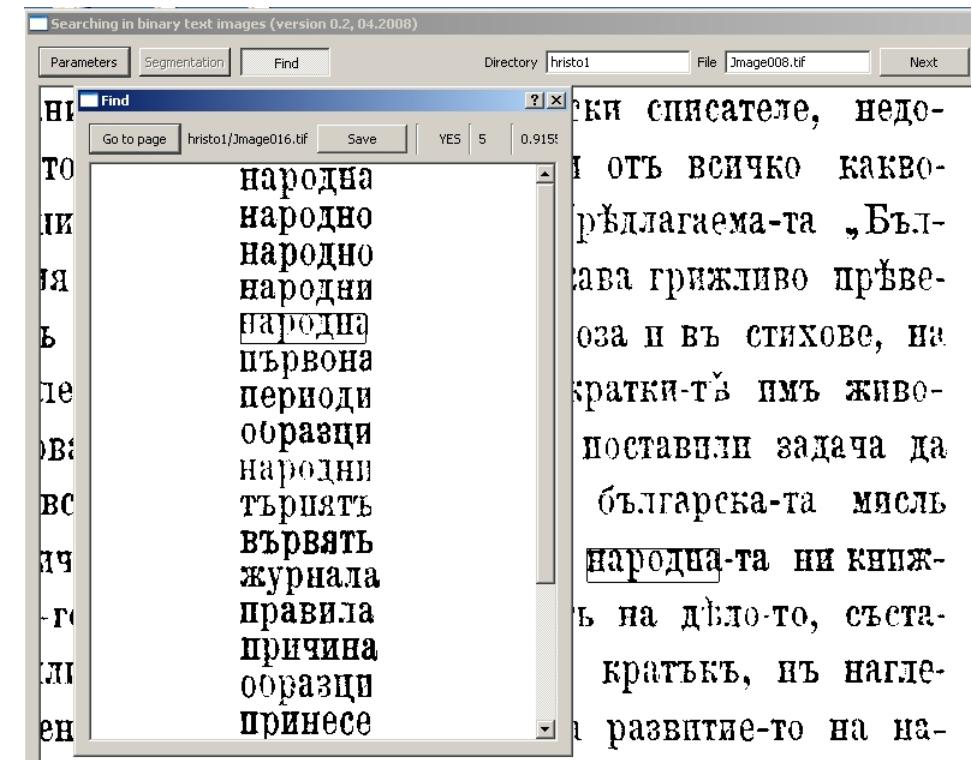
Но и едните и другите със своето майсторство са пренасяли душите на своите слушатели в друг мир. С живите хороводни мелодии те са карали непринудено хората да играят хора и ръченици, карали са със своите гласови възможности да забравят делничните трудности, като същевременно са поддържали будно националното съзнание посредством българските напеви и словесно съдържание.



Bulgarian book, Christomatia (1884), 1000 pages, tif (2300×3800), 8 BPP

Повече-то отъ ранни-тъ му стихотворения сж любовни пѣсни, по подражание на грѣцки-тѣ, и не прѣставляватъ лигературна стойностъ; стихотворения-та му въ „Смѣсна Китка“ при всичко, че повечето сж слаби подражания на руски-тѣ, нѣ свидѣтелствоватъ вече за поетическо-то дарование на г. Славейкова; най-добри-тѣ му стихотворения сж обнародвани-тѣ пд-послѣ въ „Читалиште,“ отъ които „Не пѣй ми се,“ „Жестокостъ-та ми се сломи“ и „Тогасть понѣ“ джхатъ съ истински лиризъмъ и заслужено привлѣкоха внимание-то на читателе-тѣ. Славейковъ, който е вѣшътъ въ бѣлгарски езикъ, прѣвъ доказа гѣвкостъ-та му въ поезия-та. Като се числи между първи-тѣ борци по черковниятъ вѣпросъ, той захващата въ сѫшто-то врѣме почтенно място въ редъ-тѣ на малко-то ни добри литератори.

Велико влияние е упражнила вѣзъ пробужданіе-то духъ-тѣ камъ свободата на независимостъ-та у бѣлгарски народъ доста обширна-та литературунна дѣятелностъ на *Георгий Саса Раковски* (род. въ Котелъ 1818, умр. въ Букурешть 1868 г.). Въ личностъ-та и въ дѣла-та на Раковски се отрази най-нагледно тогавашно-то състояние на умове-тѣ, нужди-тѣ, стремления-та и идеали-тѣ на народъ-тѣ ии. Тако-речи единичкъ дѣнецъ по онова врѣме, той писувѣ, работї всичко. Той искаше да обгърне въ своя-та широка дѣятеличество всички-тѣ нужди на народъ-тѣ ии, да удовлетвори всички-тѣ национални купнѣянія, да осѫществи най-смѣтни-тѣ и въжделени мечти. Той вѣзъздаде съ фанатический вѣсторгъ минжло-то и приготви бѫдже-ште-то. Бѣше въ сѫшто-то врѣме поетъ, историкъ, етнографъ, публицистъ, агитаторъ и хайдутинъ. Нито на единъ бѣлгарски дѣя-тель животъ-тѣ не е билъ напълненъ съ толкова неутолима и раз-нообразна дѣятелностъ и напъстренъ съ толкова бѣди, приключения и странности. Той се бѣше училъ въ Атина, Парижъ, Цариградъ и въ Русия. Знаеше руский, срѣбский, румънский, турский, грѣцкий, староєлинский, французскей, арабский и дори отъ части санскрит-

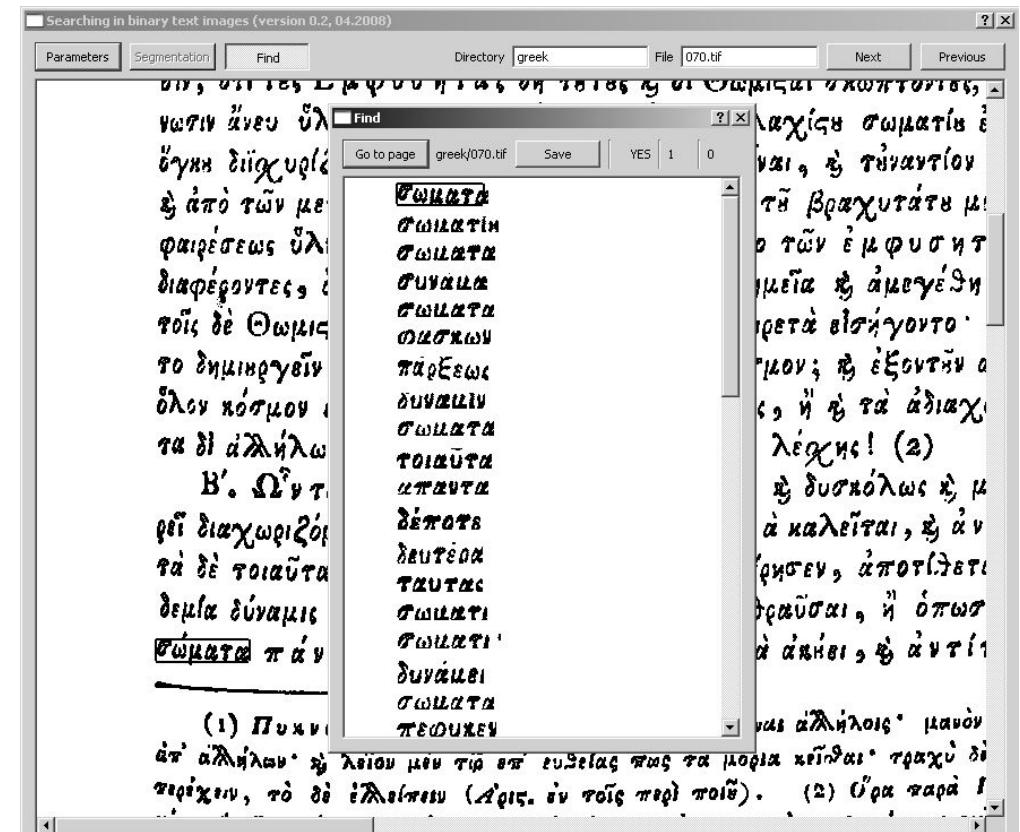


Old Greek text (approx. in the third century BC), 50 pages, jpg (1077×1416), 8 BPP

τοιαῦται κατὰ συχῆμα πάντη εἰσὶν ἄτρεπται, οἷα μ
Σῶμα δὲ σύνθετον ἐν τῇ φύσει ύδεν τοιάτον· ὅπ
δλίψει, τομῇ ὀπωσθὲν ὑπείκει, καὶ λίθαντες οἱ σερρ
τητι διαφέρων ἀδάμας, μηδενὸς ὅλως ἐξαιρυμένη,
μοιρον, ὡς δέδειπται. Ω̄ν δὲ σωμάτων τὰ μέρη
καὶ, καὶ ῥᾶσαι διαχωρίζεσθαι πέφυκεν ἀπολαὸ ταῦτ
μέλι, ἄργιλλος, κτ.: ὅσῳ δὲ ἡττονι καὶ ἀτονωτέρῳ
πλοκῆς ἀποσείχει, τοτήτῳ καὶ ἀπαλώτερῳ, ἐς ἕκρ
ύδενὶ γάρ ἐσιν ἐντυχεῖν, οὗ τὰ μέρη μὴ ὀπωσθε
γγύμενα.

Γ'. Τὸ σκληρὸν σῶμα ὑπὸ κάφιστε καὶ ἀσθενεῖ
πεψυκός, εἴ τι δέρασις οὐ ἀκύει· τοιαῦτα χάλυψ
τὰ κεράμεια σκεύη· Τύτων τὰς ερεεὰς μέρη ἐχότων
ἔληλοις, διὸ καὶ δῆται τῆς ἀμοιβαίας παρῆις ἀφίσσεται

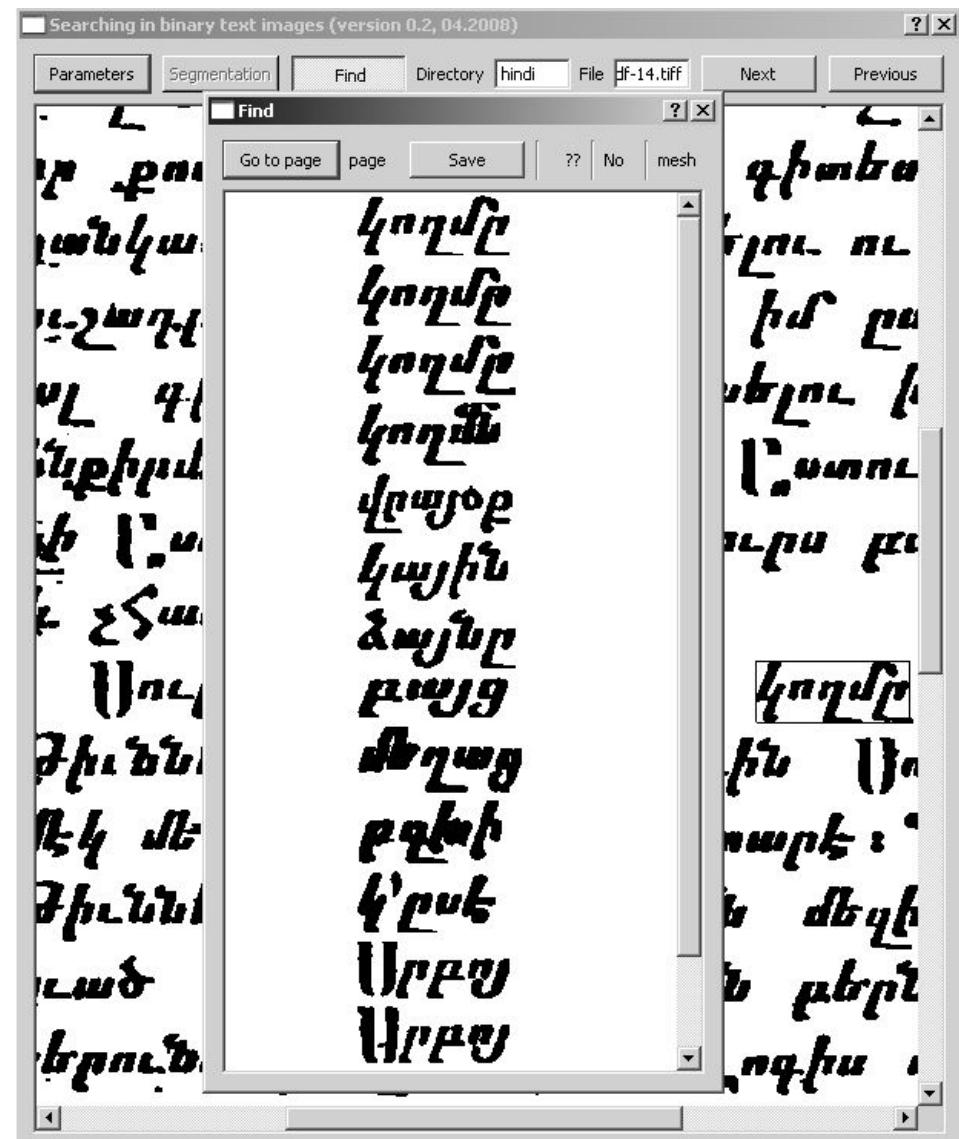
Δ'. Τὸ ἐκ πολῶν οἵουεὶ λεπίδων πάνυ λεπτῶν ἔ
εάμενον, εὗσχισον σῶμα προσείρηται· τότε τὰ μόρια εὐρύτερον προσκεκόληται ἀλλήλοις, ἥ λεπίς ἔ
ἄριτ τὰ τοιαῦτα σώματα εἰς λεπίδας ἀναλυόμενα·
καλέμενοι λίθοι οἱ ἐξ Ἰβηρίας, καὶ Καππαδοκίας,



Armenian book (1858), 178 pages, tiff (2800×5000), 1 BPP

կառակութեան օգին ամէն աստիճանի մարդոց արրտին մէջ տարածուեցաւ, և սորվեցուց անոնց մէկզմէկ ատել ու մէկզմէկէ գարշիլ. մինչև անգամ մէկ խանութի մէջ գործող արհեստաւորը սորվեցաւ նզովք տալ իր քովի խանութին մէջ բանող դրացին, ան պատճառով որ անիկայ Հոգւցին Արքոյ բղխումը իրեն համաձայն չդաւանիր. ոչ մէկը և ոչ մէկալը հասկրնալով թէ ինչ կ'ըսեն, կամ ինչ բան կ'ուզեն հաստատել:

Ուստի այսպիսի պնտեղի վիճաբանութիւնները պատճառ տուին որ երբոր մարդիկ Հոգւցին Արքոյ վրայցը խորհին, գրեթէ միայն աս մէկ նիւթիս ուղղեն միտքերնին, այսինքն թէ՝ Հոգւցին բղխումը միայն Հորմէն է, կամ Հորմէն ու Արդիէն ։ Ամէնն ալ կը դաւանին թէ Հոգին Այուրբ՝ Երրորդութեան մէկ անձն է. բայց ովկ կրնայ ըսել թէ անիկայ ինչ ներգործութիւն կ'ընէ մարդուս Հոգւցին փրկութեանը համար, կամ ինչ է իր մասնաւոր պաշտօնը մարդս երկինքը բարձրացընելու համար։ Ահա աս մեծ և ամենահարկաւոր նպատակիս համար է որ Երրորդութեան վարդապետութիւնը յայտնուած է։ Հայրը Խրկեց Արդին աշխարհը փրկելու։ Կինչու որ Աստուած անանկ սիրեց աշխարհը մինչև որ իր միածին Ար-



Text in Spanish (1901), 30 + 57 pages, gif (1400×2500), 4 BPP

Alonso, Rogelio M., Cartilla histyrico-descriptiva del türmino municipal de Macuriges. Habana:
Impr. La Propagandista, 1901, HOLLIS Catalog, Harvard University

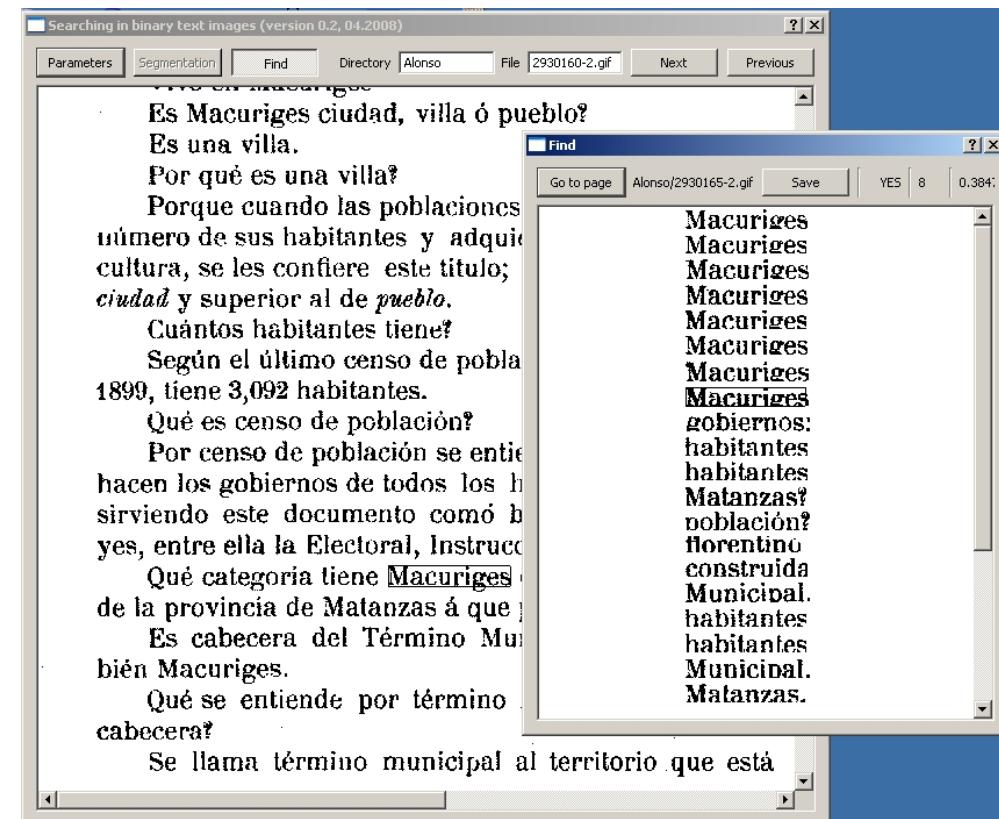
No señor; todas las fincas azucareras tienen sus chuchos que conectan con las líneas del ferrocarril y hay adcmás caminos reales, trasversales y vecinales, estos en estado natural. (1)

Qué entiende V. por caminos reales, trasversales y vecinales?

Caminos reales, son los caminos abiertos por el gobierno Español desde los tiempos primeros de la colonización de la Isla de Cuba y tienen de ancho 24 varas; caminos trasversales son los que solo tienen de ancho 12 varas y vecinales los pasos permitidos por los propietarios de fincas, para acortar distancias de un lugar á otro y salvar lo mal que pudieran estar los caminos por el fango, las piedras ó la yerba.

Cuántos ingenios para la fabricación de azúcar tiene en la actualidad el Término todo?

Los siguientes: «Santa Filomena» en el barrio de Navajas propiedad del Sr. Leandro Soler, «Elizalde» del Sr. Alberto Broch en el Ciego y «Santa Catalina» del señor Enrique Heedigg en el mismo barrio; «Carmen» del Sr. Alexander en Navajas, «Socorro» del Sr. Pedro Arenal en Tramojos y «Dolores» del Sr. Francisco Rosell en Platanal, todos centrales y con magnificos aparatos.

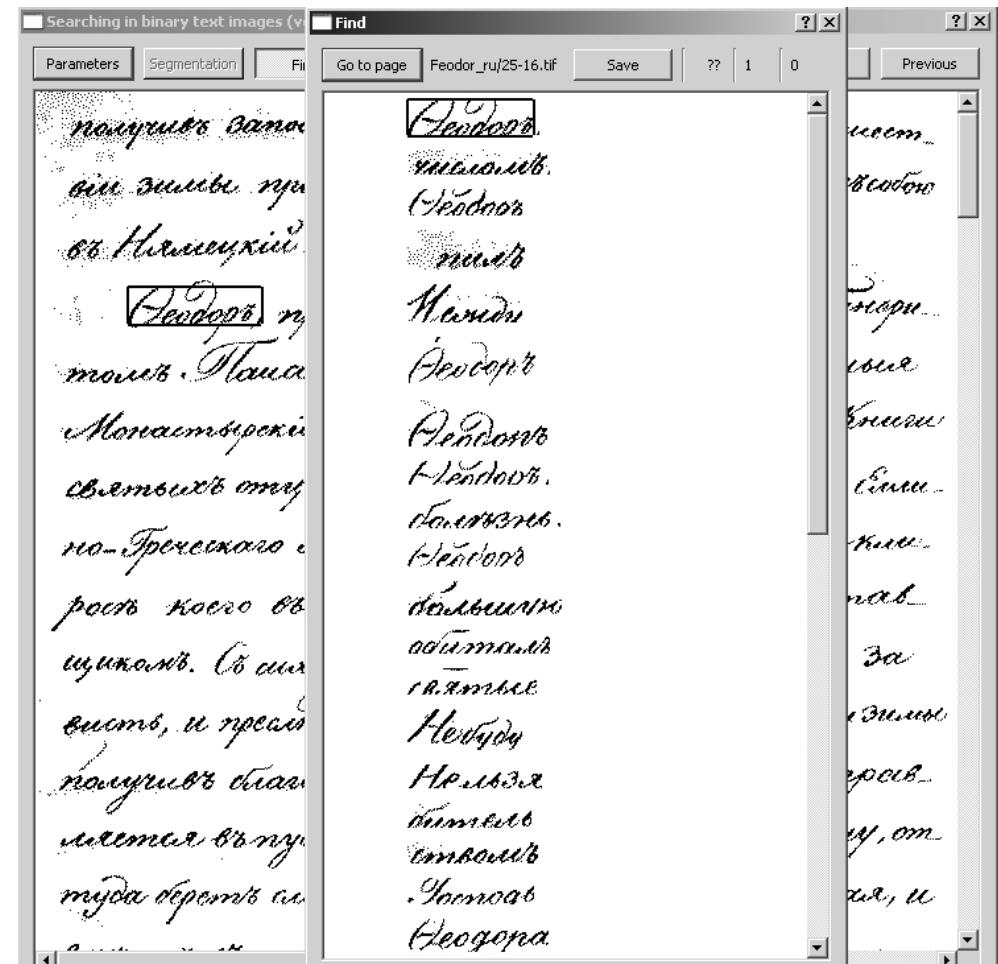


Handwritten document in Russian (1840), 44 pages, jpg (700×900), 24 BPP

Дом живоначальной Троицы, Свято-Троицкая Сергиева Лавра, Собрание славянских рукописей, 43: Житие схимонаха Феодора

Уже нога его не двигалась от боля болезней; ходь амортный оглохший его члены. Феодоръ растворилъ свои икода, инарошилъ собственного тела, согрѣвалъ ощущеніемъ тепла душевнаго его друга, покрывалъ горячими лобзаніями его щеки, освященные чистотою дѣствства, и обильными огнемъ Божественныхъ благословій. На рукахъ Феодора склонился великий Николай, и мощей его не дерзнуло прикоснуться тигоны.

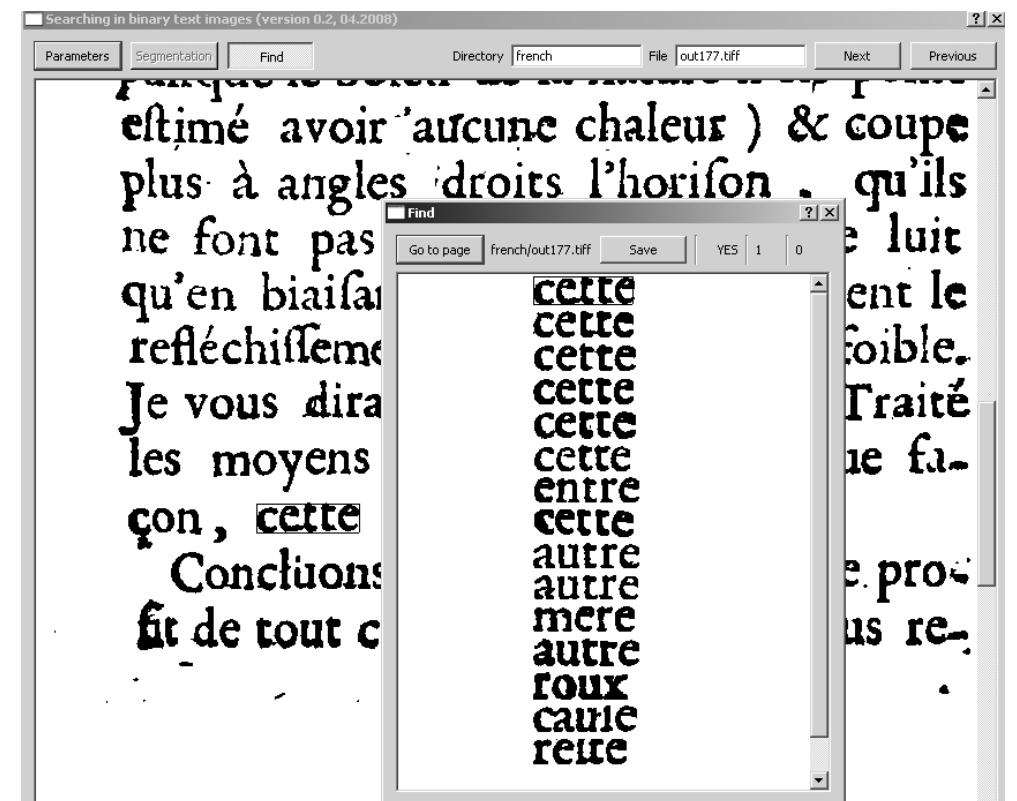
Феодоръ пребывалъ въ Нижнемъ до 1801^{го} года. Въ продолженіи сего времени уединившися кончику, высокаго митіемъ Николая, увидѣлъ кончикъ изнанкинаго Панциря. Проехавши сего



Text in French (1692), 388 pages, jpg (2048×3550), 8 BPP

Nicolas de Bonnefons, Ch. de Sergy, (1692), University of Gent, Digitized by Google (2007)

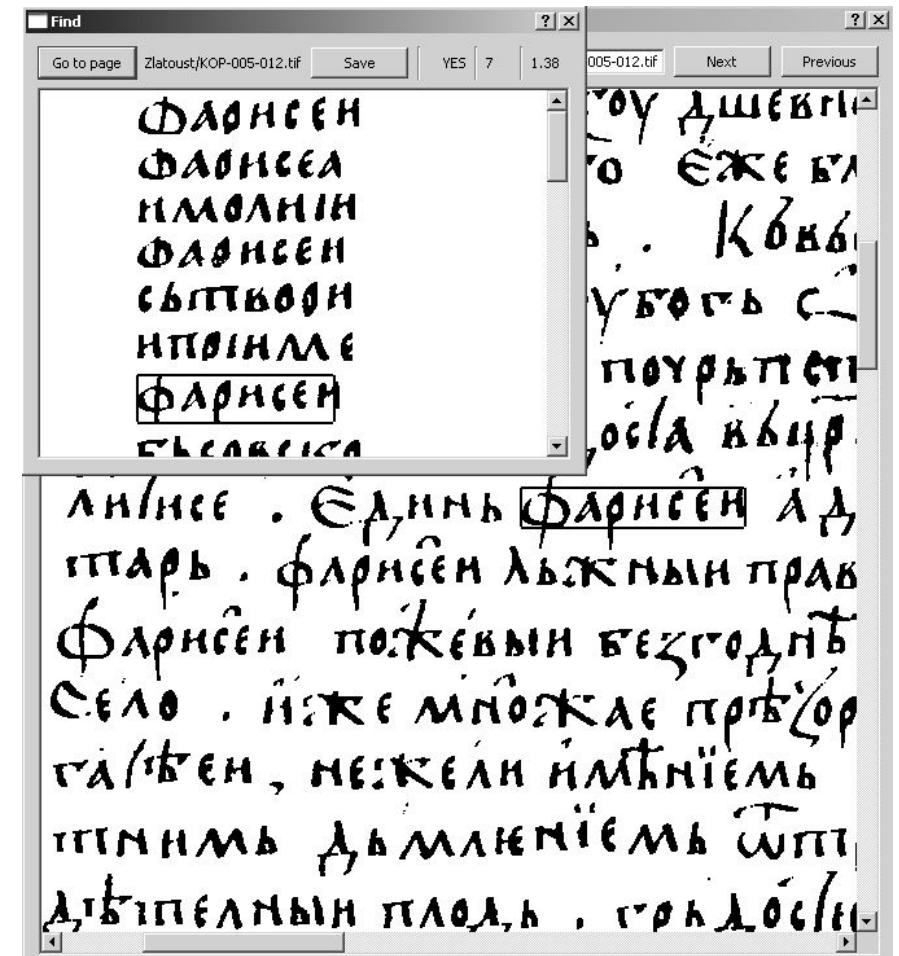
Quant à la Terre, si vous la rencontrez bonne, ce vous sera un grand avantage, & une grande épargne ; mais rarement en pourrez-vous trouver, où il n'y ait beaucoup à travailler, d'autant que telle paraîtra passablement bonne au dessus, qui étant ouverte de la profondeur d'un fer de Béche seulement, se trouvera Argileuse dessous ; ce fonds est pire aux Arbres que le Tuf, ou la Roche, à cause qu'il



Slavonic manuscript, (1574), 747 pages, jpg (1249×1890), 24 ВРР.

Дигитална Народна библиотека Србије, Џирилски рукописи, Збирка словенских рукописа Јернеја Копитара, Зборник "Златоуст"

съмръти и мъхъ. яко да не надѣющесе
богудемъ насъ, пъ на ба въскръшающаго
мътвые. иже ѿполнкы съмръти и
зблви на изблви. пакже оу повахъмы
яко и ёще изблви. Уто оу бо ръцимъ
възпосиши се яко ѿ добрыи свои. Фмъ
споже благодъть исповѣдовали да въшомъ
у тобо имаши єже непрѣель єси. аще
ли прїель єси, упосе хвалиши яко непрѣ
мъ. не ты ба позналъ єси правдою, пъ
вътебѣ благобію позна. вѣщє рече ба,
пакже познанн бывше ѿ ба. не ты ба
прїель єси добродѣтелю. пъ тебѣ хсъ
пришъствїемъ прїеть. гонеко рѣ аще
и постнгну имже и постнжанъ быхъ
ѡхъ. неви мене избрасте рѣ гъ. пъ
азъ избралъ. пъ ли зане поуѣтень єси
веле мудроствоуещи; имѣть вѣнч



Thank you for the attention.