

Межгосударственный совет по стандартизации, метрологии и сертификации, 1989. – 13 с.

References

1. Nosova, Z. A. (1973). Zirconium glazes. Leningrad, USSR: Stroyizdat, 192.
2. Lisachuk, G. V., Ryshchenko, M. I., Belostotskaya, L. A. et al. (2008). Glass-ceramic coatings. Kharkov, Ukraine: NTU "KPI", 480.
3. Levitsky, I. A. (1999). Fusible glazes for tiles and household ceramics. Minsk, Belarus: Belarusian State Technological University, 396.

4. Bobkova, N. M., Silich, L. M. (1992) Alkali-free glass-ceramics and glass-crystalline materials. Minsk, Belarus: Science and Technology, 277.

5. Bloemen, L. M. (1954) Glazes. Moscow, USSR: Promstroyizdat, 172.

6. Minerals. Equilibrium phase diagram: a handbook. Issue 1. (1974). Moscow, USSR: Science, 514.

7. Lisachuk, G., Fedorenko, O., Pitak, O. et al. (2013). Theoretical background of alkaline-free tin content coatings on ceramic in the system $RO - SnO_2 - Al_2O_3 - SiO_2$. Chemistry & Chemical technology, 7 (3), 351–354.

8. Porcelain articles. Specifications GOST 28390-89. (1989). Moscow, USSR.: Interstate Council for Standardization, Metrology and certification, 13.

Дата надходження рукопису 10.10.2014

Федоренко Олена Юріївна, доктор технічних наук, професор, кафедра технології кераміки, вогнетривів, скла та емалей, Національний технічний університет «Харківський політехнічний інститут», вул. Фрунзе, 21, м. Харків, Україна, 61002

E-mail: fedorenko_e@ukr.net

Рищенко Михайло Іванович, доктор технічних наук, професор, кафедра технології кераміки, вогнетривів, скла та емалей, Національний технічний університет «Харківський політехнічний інститут», вул. Фрунзе, 21, м. Харків, Україна, 61002

Дайнеко Катерина Борисівна, аспірант, кафедра технології кераміки, вогнетривів, скла та емалей, Національний технічний університет «Харківський політехнічний інститут», вул. Фрунзе, 21, м. Харків, Україна, 61002

E-mail: catherine@i.ua

Пермяков Юрій В'ячеславович, доцент, кафедра технології кераміки, вогнетривів, скла та емалей, Національний технічний університет «Харківський політехнічний інститут», вул. Фрунзе, 21, м. Харків, Україна, 61002

УДК 681.3:655

DOI: 10.15587/2313-8416.2014.28817

ИССЛЕДОВАНИЕ ЗАДАЧИ АВТОМАТИЗАЦИИ ПОДГОТОВКИ К ПЕЧАТИ СТАРИННЫХ ИЗДАНИЙ

© **И. В. Тришина, А. В. Бизюк**

В данной работе исследуется процесс автоматизации подготовки к печати старинных изданий на примере дореволюционной книги с целью усовершенствовать и ускорить процесс. Приведен и подробно рассмотрен этап, занимающий больше всего времени в данном процессе. Предложен метод частичной автоматизации данного этапа.

Ключевые слова: оптическое распознавание символов, текстовая информация, коррекция, макрос.

This paper describes the old publications automation process based on the example of a pre-revolutionary book with the purpose of improving and speeding up this process. The step that takes most of the process time is shown and clarified. A partial automation method for this step is proposed.

Keywords: optical character recognition, text information, correction, macro.

1. Введение

Несмотря на развитие электронных технологий, одним из основных способов получения информации все еще являются книги.

Производство книжной продукции является очень сложным, многоэтапным процессом и важным этапом определяющим изготовление книжной изделий является разработка технологии ее изготовления.

Повышена необходимость оцифровки старин-

ных и ветхих книг, которые находятся на грани исчезновения, а так же их переиздание [1].

Переиздание – каждая новая публикация, отличающаяся от предыдущих минимум номером типографского заказа или же изменениями, как содержания (основного текста), так и оформления [2].

Не маловажной проблемой является адаптация дореволюционной орфографии. Часть книг переиздается с адаптацией со старого правописания на новое, для восприятия информации более широким кругом

читателей. С точки зрения доступности электронная форма публикаций позволяет повысить надежность хранения, удобство распространения, экономичность. Однако печатные издания традиционны, а, следовательно, привычны и удобны в использовании для большого количества людей, поэтому эта форма переиздания не может быть исключена, т. е. современное переиздание должно содержать печатную и электронную форму [3].

Следовательно, анализ и усовершенствование технологий переиздания является актуальной задачей, которая позволит сократить время и принести экономический эффект на базе повышения качества переиздания.

2. Постановка проблемы

Под объектом исследования подразумевают процесс или явление, что порождает проблемную ситуацию и избранное для изучения. Понятие объекта исследования является общим, относительно понятия предмета исследования. В объекте выделяют ту его часть, которая служит предметом исследования и определяет тему магистерской аттестационной работы [4].

В данной работе объектом исследования является процесс подготовки к печати старинных изданий. А предметом исследования автоматизация данного процесса.

Рассмотрим начало технологической цепочки переиздания ветхих книг, так как в этих этапах и заключается сложность данного процесса. При подготовке к переизданию старинную книгу изначально сканируют или же фотографируют. В каждом конкретном случае вопрос решается индивидуально. Если возможности сканирования нет (ведь не всегда старую книгу можно вынести из книгохранилища и расшить ее), то снимки делаются с помощью цифровых фотоаппаратов новейших моделей [5].

Если задачей стоит репродуцировать оригинал (создать репринт), следующим этапом идет цветокоррекция. С ее помощью добиваются эффекта однородности страниц. Основным недостатком этого способа переиздания так же является плохое качество оттисков. Поэтому не всегда останавливаются на репринте, а вместо цветокоррекции распознают текстовую и графическую информацию [6].

По большому счету распознавание текстовой информации являет собой более сложный процесс, чем кажется в начале. Истертая от времени бумага, плохой оттиск, пометки карандашом на полях усложняют распознавание. Корректур такого текста занимает много времени и особого внимания из-за громадного количества ошибок. В результате возникает проблема следующего вида – некоторые буквы упрямо распознаются как строчные, а другие, как прописные. Например: СоглаСитЕсь, что таКой теКст нужно исправить [7]. Следующим этапом текст переводят с дореформенного правописания на новое, затем так же идет коррекция адаптированного текста.

Самый долгий и ёмкий из рассмотренных этапов – это процесс корректуры текста после распознавания. Предлагается исследовать данный этап на предмет необходимости и возможности автоматизации, что явно уменьшило бы временные затраты и облегчило работу на данном этапе допечатной подготовки.

3. Литературный обзор

Ёмкость данного процесса заключается в распознавании текстовой информации с большим количеством ошибок. Данная тема уже была рассмотрена и исследована в диссертации Филиппович Анны Юрьевны, которая называется «Автоматизированная технология переиздания лексикографических источников XVIII-нач. XIX вв.: на материале Словаря Академии Российской 1789–1794 гг.» В данной работе сравнивалось две технологии корректуры – традиционная (ручная) и автоматизированная (с помощью словаря-спеллера) [8].

Сравнение позволяет сделать вывод об эффективности технологий корректуры с использованием словаря-спеллера по сравнению с традиционной технологией. В случае использования словаря-спеллера количество слов, сравниваемых корректором, уменьшается и по мере пополнения словаря на последней странице достигает ~ 20 % общего объема [9]. Существует два недостатка в данной работе.

Первое – это то, что такие буквы как «ять», «фита» были рассмотрены как ошибки, которые так же нужно исправить. Вместо того что бы выбрать языком распознавания в программе OCR языком распознавания «Русский (Старая Орфография)». Что позволило бы программе OCR распознавать данные символы правильно и в дальнейшем к отредактированному тексту применить переводчик с «петровского правописания» на «современный русский». Если выбрать языком распознавания «Русский (Старая Орфография)» количество ошибок уменьшится приблизительно на 73 %.

Второе – это недостаток словаря-спеллера. Когда в словарь добавляют, к примеру, слово «защитить», где программа OCR не упрямо распознает букву «щ» как «іц», то встретив в дальнейшем слово «защита» с такой же проблемной программой не исправит букву «щ», так как не склоняет слова. Далее я расскажу вам автоматизировать этот процесс. И добиться таких же результатов (20 %) другим способом, который можно применять и к работе с другими оригиналами.

4. Исследование задачи автоматизации корректуры текста

Исследуем процесс автоматизации подготовки к печати старинных изданий на примере дореволюционной детской книги «Моя первая священная история» П. Н. Воздвиженского 1899 г.

Выходные сведения издания представлены в табл. 1.

Таблица 1

Выходные сведения рассматриваемого издания

Наименование издания	Формат издания, см	Объем издания в физ. печ. л.	Тираж издания, тыс. экз.	Красочность издания		Внешнее оформление	
				Лицо	Оборот	Переплетная крышка	
						Тип	Красочность
Научно-популярное	84x108/16	7	100	1	1	7Б	4+0

Недостатки процесса допечатной подготовки переиздания старинных книг:

- это емкая и долгая работа;
- самые долгий и ёмкий процесс это 1-я корректура.

В книге 112 страниц. Изначально языком распознавания было взято «Русский». После распознавания количество ошибок достигало 2000. После применения языком распознавания «Русский (Старая Орфография)» выявлено 1093 ошибки. Количество ошибок уменьшилось на 73 %. Анализ ошибок показал их повторяемость и формализуемость. Вид повторяемых и формализуемых ошибок представлен в табл. 2.

Таблица 2

Классификация существующих ошибок после распознавания текстовой информации в программе ABBY Y FineReader 11

Классификация ошибок	Примеры	Исправленные ошибки
Переносы	на- звали	назвали
Переносы	встрѣ- тится пер- выми \ сло- вомъ допол- няя уяс- нятъ без- плоднымъ размыш- леній по- звали покуп- кою ка- мышѣ ко- торый на- чали по- шель ПРЕ- СВЯТОЙ Да- вида Оби- дѣтъ	встрѣтится первыми словомъ дополняя уяснять безплоднымъ размышлений позвали покупкою камышѣ который начали пошелъ ПРЕСВЯТОЙ Давида Обидѣтъ
Шипящие	Священное Священная Начинаючихъ Заіцитить Помѣцаются Услыніаль	Священное Священная Начинающихъ Защитить Помѣщаются Услышалъ
Большие буквы в середине слова	Иосифу Да здрав- СТВУеТЬ ГОС- подь счаСтдивы	Иосифу Да здравствуетъ Господь счастливы
Написание «іі» вместо «и».	животныхъ іі растений замѣ- чанія іі совѣты дѣтеіі родіітелѣй ріісунками изданій іі въ	животныхъ и растений замѣчания и совѣты дѣтей родителей рисунками изданий и въ

5. Аprobация результатов исследований

Методом проведения исследования является создание макросов, которые уменьшит количество ошибок после распознавания текста старого оригинала.

Макрос – это набор команд и инструкций, выполняемых как одна команда [10].

Для решения поставленной задачи предположительно необходимо выполнить два задания.

Задание 1:

- создать макрос, который будет находить и удалять переносы в тексте.

- создать макрос, который будет находить и исправлять ошибки такого рода: помѣцаются – помѣщаются.

- создать макрос, который будет находить и исправлять ошибки такого рода: ріісунк – рисунок.

Задание 2:

Создать макрос, который будет находить и исправлять прописные буквы на строчные в середине текста, что бы заглавные буквы в начале предложений остались заглавными, а так же имена и названия.

В результате использования языком распознавания «Русский (Старая Орфография)» количество ошибок уменьшилось приблизительно на 73 %, как и говорилось ранее.

В результате проверки работы макросов и их применения количество ошибок уменьшилось на 20 %.

5. Выводы

Количество ошибок после применения языком распознавания «Русский (Старая Орфография)» уменьшилось приблизительно на 73 %. После применения созданных макросов были исправлены ошибки такого рода, как представлено в табл. 1. Все остальные ошибки, обнаруженные в распознанном тексте оригинала, были не повторяющиеся. Большинство из них являлось следствием потертостей оригинала от старости, что уменьшало вероятность правильного распознавания символов в таких местах. Количество ошибок после применения макросов уменьшилось ещё на 20 %.

Литература

1. Компьютерная лингвистика орфография [Электронный ресурс] / Режим доступа: http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/med_mat/med_mat.htm – 15.10.2014. – Загл. с экрана.

2. Антонова, С. Г. Редакторская подготовка изданий [Текст]: уч. / С. Г. Антонова, В. И. Васильев, И. А. Жарков

и др.; под общ. ред. С. Г. Антоновой. – М.: Издательство МГУП, 2002. – 468 с.

3. Квин, О. Д. Допечатная подготовка. Руководство дизайнера [Текст] / О. Д. Квин, пер. с англ. – М. : Издательский дом “Вильямс”, 2002. – 592 с.

4. Кулишова Н. Е. Методические указания для магистерской аттестации [Электронный ресурс] / Н. Е. Кулишова. – Локальная сеть кафедры МСТ, ХНУРЭ.

5. Репринт. Салон-мастерская подарочной книги [Электронный ресурс] / Режим доступа: <http://oldsalon.ru/uslugi/reprint.php> – 15.10.2014 – Загл. с экрана.

6. Маргулис, Д. Классическое руководство по цветокоррекции [Текст] / Д. Маргулис; пер. с англ. – М. : ООО “РТВ – Медиа”, 2001. – 400 с.

7. Сканирование книг на дореформенной русском [Электронный ресурс] / Режим доступа: <http://rtr24.ru/sekrety/skanirovanie-knig-na-doreformennom-russkom> – 15.10.2014 – Загл. с экрана.

8. Филиппович, А. Ю. Автоматизированная технология переиздания лексикографических источников XVIII-XIX вв.: на материале Словаря Академии Российской 1789–1794 гг. [Электронный ресурс] / А. Ю. Филиппович. – Режим доступа: <http://www.dissercat.com/content/avtomatizirovannaya-tekhnologiya-pereizdaniya-leksikograficheskikh-istochnikov-xviii-nach-xi> – 26.10.2014 – Загл. с экрана.

9. Филиппович, А. Ю. Исследование эффективности автоматизации корректурных процессов с помощью словаря spellera при подготовке переиздания Словаря Академии Российской 1789–1794 гг. [Электронный ресурс] / А. Ю. Филиппович. – Режим доступа: http://it-claim.ru/Anna/public/Public_Statya/AutoCorrect/AutoCorrect.pdf – 26.10.2014 – Загл. с экрана.

10. Установка Office [Электронный ресурс] / О макросах. – Режим доступа: <http://office.microsoft.com/ru-ru/word-help/HP005189421.aspx>

ru/word-help/HP005189421.aspx – 15.10.2014 – Загл. с экрана.

References

1. Computational linguistics and spelling. Available at: http://it-claim.ru/Library/Books/ITS/wwwbook/ist6/med_mat/med_mat.htm (Last access: 15.10.2014).

2. Antonova, V., Vasilyev, V. I., Zharkov, I. A. (2002). Preparation of editorial publications. Moscow: Publisher MGUP, 468.

3. Quinn, O. D. (2002). Prepress. Designer's Guide. Publishing House "Williams", 592.

4. Kulishova, N. E. Methodological instructions for master's certification. Local chair of MST KHNURE.

5. Reprint. Showroom-workshop of gift books. Available at: <http://oldsalon.ru/uslugi/reprint.php> (Last access: 15.10.2014).

6. Margulis, D. (2001). The classic guide to color correction. Moscow: ООО "РТВ – Медиа", 400.

7. Scanning books in the pre-reform Russia. Available at: <http://rtr24.ru/sekrety/skanirovanie-knig-na-doreformennom-russkom> (Last access: 15.10.2014).

8. Computer-aided reprint technology of lexicographical sources of XVIII - early XIX centuries. Available at: <http://www.dissercat.com/content/avtomatizirovannaya-tekhnologiya-pereizdaniya-leksikograficheskikh-istochnikov-xviii-nach-xi> (Last access: 15.10.2014).

9. Philippovich, A. Y. Research the effectiveness of automation of the proof-processes using vocabulary speller in the preparation reprint of the Dictionary of the Russian Academy of 1789–1794. Available at: http://it-claim.ru/Anna/public/Public_Statya/AutoCorrect/AutoCorrect.pdf (Last access: 26.10.2014).

10. MS Office installation. About macros. Available at: <http://office.microsoft.com/ru-ru/word-help/HP005189421.aspx> (Last access: 15.10.2014).

*Рекомендовано до публікації д-р техн. наук Гребенник І. В.
Дата надходження рукопису 27.10.2014*

Тришина Ирина Владимировна, студентка, кафедра медиасистем и технологий, Харьковский национальный университет радиоэлектроники, пр. Ленина, 16, г. Харьков, 61166

E-mail: trishina.irina.v@gmail.com

Бизюк Андрей Валериевич, кандидат технических наук, доцент, кафедра медиасистем и технологий, Харьковский национальный университет радиоэлектроники, пр. Ленина, 16, г. Харьков, 61166

E-mail: abizuk@mail.ru

УДК 628.355

DOI: 10.15587/2313-8416.2014.28768

МОДЕЛИРОВАНИЕ ПРОЦЕССОВ ПЕРЕНОСА В АНАЭРОБНОМ БИОРЕАКТОРЕ С ИММОБИЛИЗИРОВАННОЙ МИКРОФЛОРОЙ

© Л. И. Ружинская, А. А. Фоменкова

Предложена методика численного решения математической модели массопереноса в анаэробном биореакторе с иммобилизированной на плоских вертикальных носителях микрофлорой с учетом двустадийности процессов очистки сточной воды от загрязнений органическими веществами. Для численного решения модели использован метод сеток. Приведены результаты расчета с помощью пакета MATLAB.

Ключевые слова: биореактор, математическая модель, биопленка, иммобилизованная микрофлора, очистка сточных вод, биомасса.