УДК 519.67, 519.61

DOI: 10.15827/0236-235X.120.684-689

Дата подачи статьи: 28.09.17

2017. Т. 30. № 4. С. 684–689

О ЗАДАЧЕ РЕТРОКОНВЕРСИИ НЕФОРМАТНЫХ ДОКУМЕНТОВ В СИСТЕМАХ ФОРМИРОВАНИЯ И ПРЕДОСТАВЛЕНИЯ ЦИФРОВОГО КОНТЕНТА

А.Н. Сотников, д.ф.-м.н., профессор, ASotnikov@jscc.ru;

И.Н. Чередниченко, к.т.н., inch@jscc.ru

*(Межведомственный суперкомпьютерный центр РАН,
Ленинский просп., 32а, г. Москва, 119991, Россия)*

Рассматривается математическое обоснование решения задачи ретроконверсии (восстановления) сканированных неформатных документов применительно к задаче наполнения и публикации цифровых фондов электронных библиотек. Рассмотрены алгоритмы автоматической кластеризации и формирования автоадаптивных шрифтов, используемые для ретроконверсии документов.

В качестве механизма работы с кластеризуемыми объектами используется классический подход для формирования вектора признаков на основе модели алгоритма вычисления оценок, описывающего каждый рассматриваемый объект. Алгоритм автоматической кластеризации построен на основе статистического подхода к формированию характеристических векторов каждого кластера и к динамическому построению разделяющих гиперплоскостей.

В связи со спецификой решаемых задач используется контурный подход к обработке графических объектов, который позволяет применить методы частотного анализа на основе преобразования Фурье. Представлены оценки необходимой и достаточной размерности вектора признаков, используемого как для описания графических объектов, так и для их восстановления с минимальными искажениями в результате формирования выходных документов, полученных при выполнении ретроконверсии.

Приведены возможные перспективы развития описанного метода ретроконверсии неформатных документов в области организации контекстного поиска внутри графических документов.

Ключевые слова: ретроконверсия, автоадаптивные шрифты, неформатные документы, автоматическая кластеризация, электронная библиотека.

Базовой компонентой формируемого цифрового пространства знаний являются информационные системы обеспечения потребности общества в структурированном, достоверном и актуальном знании. Такие системы призваны обеспечить формирование, интеграцию, сопровождение и предоставление цифрового контента для информационной поддержки управленческой, научно-образовательной и просветительской деятельности. Эффективной технологической платформой для этих задач является комплекс решений, объединенных понятием электронной библиотеки, обеспечивающий формирование, интеграцию и предоставление широкому кругу пользователей информационных материалов по различным проблемам.

Основные источники формирования цифрового контента электронных библиотек – фонды библиотечного, архивного и музейного хранения. Наиболее значимыми как с точки зрения объемов, так и востребованности являются документы на бумажных носителях (книги, научные журналы, манускрипты и пр.), в том числе редкие и особо ценные издания, имеющие свою «историю» в виде рукописных пометок и комментариев читателей, которые зачастую являются свидетельствами эпохи и имеют историческую ценность.

Большое количество таких работ отпечатано с использованием различных, как правило, нестандартизованных (нерегулярных), шрифтов, выполнено зачастую с низким качеством печати и имеет возникшие со временем дефекты, такие как есте-

ственное старение и высыхание бумаги, осыпание типографской краски, механические повреждения, полученные в результате активного и не всегда аккуратного пользования, а также ненадлежащего хранения.

В качестве примера, демонстрирующего особенности работы с такого рода объектами, является электронная библиотека «Научное наследие России», интегрирующая в своих фондах электронные образы документов по истории развития отечественной науки, основополагающие труды выдающихся ученых, сведения об их жизни и деятельности, созданных ими научных школах и полученных результатах [1, 2].

Постановка задачи

Перевод документов на бумажных носителях в электронный формат осуществляется посредством их сканирования, последующей обработки и форматирования, формирования и сопровождения полученных цифровых образов документов набором дополнительных атрибутов, обеспечивающих возможность доступа к электронным копиям и работы с ними. При этом пользователи предполагают следующее:

1) объект сканирования, его содержание, а также графическое отображение не претерпят существенных изменений в процессе оцифровки;

2) уменьшится нагрузка на первоисточник с одновременным расширением пользовательской

аудитории вследствие работы с его цифровой копией;

3) снизятся затраты на обеспечение сохранности материалов и их доступности для пользователей за счет применения эффективных методов, алгоритмов и программ формирования и предоставления цифровых копий объектов, а также их погружения в современную информационно-коммуникационную среду;

4) будут доступны традиционные возможности работы с текстовыми и графическими объектами: удобное для зрительного восприятия представление, контекстный поиск, редактирование, формирование статистики и пр.

Если требования 1–3 достаточно просто обеспечиваются использованием современного сканирующего оборудования и погружением создаваемого цифрового контента в информационные фонды электронной библиотеки в сочетании с возможностями современного Интернета, то реализация последнего пункта требует поиска новых решений, что и является предметом исследования, отраженного в настоящей статье. Наряду с данной проблемой исследуется возможность сокращения вычислительных и сетевых ресурсов, необходимых для хранения и работы с такого рода цифровыми объектами.

Очевидно, что традиционные и достаточно развитые средства работы с текстами основаны на методах манипулирования обособленными графическими объектами, например, рисунками или буквами, являющимися элементами какого-либо шрифта. Для использования таких инструментов при работе со сканированными документами необходимо решить задачу выделения каждого графического объекта из сканированного образа, их объединения в группы по принципу «похожести», формирования из представителей этих групп элементов алфавита (шрифта), который будет источником для восстановления сканированного образа документа, но уже с использованием элементов построенного шрифта. При этом необходимо обеспечить максимально возможное визуальное соответствие восстановленного образа объекта исходному документу.

В случае, когда исходный документ выполнен с высоким качеством печати и с использованием стандартных (регулярных) шрифтов, процесс восстановления (ретроконверсии) документа после сканирования не вызывает затруднений. Проблемы возникают, когда исходный объект выполнен с низким качеством и/или с использованием нерегулярных шрифтов. Документы, созданные с использованием нерегулярных шрифтов, будем называть неформатными.

Под ретроконверсией будем понимать процедуру получения электронной копии сканированного документа, визуальное представление которой создано с помощью специально сформирован-

ного набора графических объектов. Полученное при этом графическое изображение должно быть максимально приближено к первоисточнику. Такой набор упорядоченных графических объектов, сформированный на основе анализа сканированного изображения, будем называть автоадаптивным шрифтом [3]. При этом будем стремиться к тому, чтобы созданный таким образом шрифт обладал основными признаками регулярного шрифта, а сформированный с его использованием документ открывал возможность его обработки стандартными процедурами работы с электронными документами.

Таким образом, решение задачи ретроконверсии неформатных документов требует последовательного выполнения следующих процедур:

- анализ входного потока сканированных документов, выделение графических объектов, формирование автоадаптивного шрифта;
- восстановление цифрового образа сканированного документа с использованием элементов автоадаптивного шрифта.

Формирование автоадаптивного шрифта

В основу технологии формирования автоадаптивного шрифта положены процедуры анализа графических объектов, полученных в процессе сканирования, их кластеризации с учетом степени схожести, а также динамическое формирование характеристических представителей кластеров – суть элементов автоадаптивного шрифта. Исходной для построения и исследования данных процедур выступает модель *алгоритма вычисления оценок* (АВО), предложенная в [4], в рамках которой графические объекты определяются векторами признаков фиксированной размерности, в совокупности образующими конечномерное векторное пространство. Каждому кластеру ставится в соответствие характеристический вектор U той же размерности, компоненты которого отражают усредненную характеристику близости объединенных в кластер объектов. В качестве меры близости (функции оценок) для них может быть выбрана приемлемая для предметной области метрика, отвечающая интуитивному понятию схожести. Решающее правило о принятии решения об отнесении объекта к тому или иному кластеру имеет в своей основе оценку отклонения значений соответствующих компонент вектора признаков и характеристического вектора, которое не должно превышать некоторого заданного порога. Таким образом, основными составляющими модели распознавания являются векторы признаков, характеристические векторы классов объектов, функция оценок и пороги.

Так как в рассматриваемом случае множество подлежащих кластеризации объектов (графические объекты – суть элементы шрифтов) пополняется

случайным образом, предлагается использовать адаптивную модель АВО [5], которая предполагает наличие динамической процедуры формирования компонентов характеристического вектора, функции оценок и порогов на основе статистической обработки случайного входного потока сканированных объектов, подлежащих кластеризации.

Кластеризация сканированных графических объектов. Будем считать, что каждому графическому объекту поставлен в соответствие вектор признаков

$$\vec{v} = (v_1, \dots, v_n), v = \{v_i^R\}, i = 1, 2, \dots, n \quad (1)$$

(процедура формирования вектора признаков V рассмотрена далее).

Естественно предположить, что каждый вектор признаков является случайной реализацией входного графического объекта. Таким образом, приходим к рассмотрению случайной величины $V = (v_1, v_n)$. Каждому кластеру (k -му классу объектов) поставим в соответствие характеристический вектор U той же размерности, что и вектор признаков V .

При добавлении объектов в кластер компоненты его характеристического вектора U будем определять значением математического ожидания для каждой из компонент вектора на множестве всех включенных в кластер объектов:

$$U^k = M[v^k(t)] = \frac{1}{m} \sum_{t=1}^m v_i^k(t), k = \overline{1, m}, \quad (2)$$

m – количество образцов в кластере.

В качестве решающего правила, определяющего построение граничащей гиперплоскости, разделяющей различные кластеры, выступает функция от дисперсии, возникающей в процессе статистической обработки образцов графических объектов:

$$\xi^k = D[v^k(t)] = M[(v^k(t))^2] - (M[v^k(t)])^2, \quad (3)$$

$k = \overline{1, m}$.

Зная параметры распределения, можно корректировать границы гиперплоскости, разделяющей графические объекты, варьируя всего один параметр – например, точность распознавания, которая определяется значением порога включения объекта в кластер.

Определение значения порога включения объекта в кластер. Зададим некоторое малое число $\alpha > 0$ и для скалярной случайной величины $x = v_i$ (компонента вектора признаков) определим радиус ε интервала (пороговое значение), в который величина x попадает с вероятностью $p = 1 - \alpha$. Воспользовавшись неравенством Чебышева [6], имеем

$$p = 1 - \alpha = P\{(x - \xi) < \varepsilon\} = 1 - P\{(x - \xi) \geq \varepsilon\} \geq 1 - \frac{\sigma^2}{\varepsilon^2}, \quad (4)$$

где ξ – математическое ожидание; σ – среднее квадратичное отклонение случайной величины x . По-

скольку неравенство Чебышева дает оценку сверху для соответствующей вероятности, естественно положить

$$\varepsilon = \frac{\sigma}{\sqrt{\alpha}}. \quad (5)$$

Таким образом, получаем верхнюю оценку принадлежности объекта кластеру:

$$p = \max \left(1 - \frac{(\sigma_j^k)^2}{|x_j - \xi_j^k|^2} \right), \quad (6)$$

$j = \overline{1, n}, k = \overline{1, m}$.

Последовательно применяя данное соотношение ко всем компонентам исследуемого вектора признаков, получаем наихудшую оценку истинности объекта и на основании этой оценки принимаем решение о принадлежности образца графического объекта рассматриваемому кластеру или о необходимости создания нового кластера (элемента автоадаптивного шрифта).

Таким образом, оценка для порога включения объекта в кластер ε определяется как

$$\varepsilon = \sigma \cdot F^{-1} \left(\frac{1 - \alpha}{2} \right). \quad (7)$$

Динамическое формирование характеристических векторов. В случае выполнения условия включения исследуемого объекта в кластер происходит рекуррентный пересчет компонент его характеристического вектора:

$$U_j^k(m+1) = \frac{1}{m+1} [mU_j^k(m) + v_j^k(m+1)], \quad (8)$$

$j = \overline{1, n}$,

и вектора дисперсий:

$$\sigma_j^2(m+1) = \frac{1}{m+1} \{ m[\sigma_j^2(m) + \xi_j^2(m)] + v_j^2(m+1) - (m+1)\xi_j^2(m+1) \}. \quad (9)$$

Таким образом, на основе статистического подхода и управляемого параметра α , определяющего качество распознавания, построено решающее правило для автоматической классификации входного графического объекта в качестве одного из элементов автоадаптивного шрифта.

Известно, что математическое ожидание и дисперсия полностью определяют параметры нормального распределения. Однако при малом количестве обработанных образцов (на начальном этапе) дисперсия не может быть точно подсчитана. Для оценки значения ξ^k можно использовать теорему Чебышева. В ней утверждается, что выборочное среднее случайной величины с конечными моментами первого и второго порядков сходится в среднем к математическому ожиданию при неограниченном увеличении числа опытов m , то есть

$$M[\xi^k(m)] \rightarrow \xi^k, m \rightarrow \infty. \quad (10)$$

Следовательно, чем больше образцов шрифта будет обработано, тем точнее знание о реальных параметрах распределения.

Алгоритм формирования автоадаптивного шрифта. На основе вышеизложенного метода автоматической кластеризации был реализован алгоритм формирования автоадаптивного шрифта, описание которого приведено далее. Его выполнение можно условно разделить на две смысловые части:

1) проверка принадлежности входного графического объекта элементу автоадаптивного шрифта и/или образование нового кластера:

- в случае несовпадения входного графического объекта ни с одним из элементов автоадаптивного шрифта формируется новый элемент шрифта с нулевой начальной дисперсией и фиксированным порогом;

- включение нового элемента в кластер, формирующий элемент автоадаптивного шрифта;

2) коррекция

- параметров элемента автоадаптивного шрифта;

- порогов принадлежности к элементу автоадаптивного шрифта.

Процедура ретроконверсии графического объекта

Основная задача ретроконверсии – восстановление исходного вида неформатного документа с возможностью управления качеством восстановления.

Формирование вектора признаков. Так как шрифты – это особый класс изображений, которые полностью могут быть представлены контурной информацией, был выбран следующий подход к формированию вектора признаков. Каждое изображение элемента шрифта (буквы) рассматривалось как набор внешних и вложенных замкнутых контуров. Каждый контур представляет собой набор точек, принадлежащих параметрической кривой (набору точек на плоскости), заданной вектор-функцией

$$W(t) = (x(t), y(t)), t = \overline{1, m}. \quad (11)$$

Таким образом, анализ графического объекта сводится к анализу функций $x(t)$ и $y(t)$, заданных дискретно на некотором отрезке. Как правило, для этого выполняют приближение $x(t)$ и $y(t)$ элементарными функциями и анализируют полученные приближения.

Для корректной работы с вектором признаков графических объектов к его формированию предъявляются требования инвариантности относительно преобразований типа сдвига и масштабирования.

Требование инвариантности к преобразованию сдвига (независимости от координат расположения объекта) можно обеспечить переходом в локальную систему координат, а ее начало использовать

как точку привязки рассматриваемого графического объекта:

$$W'(t) = (x(t) - x(1), y(t) - y(1)), t = \overline{1, m}. \quad (12)$$

Требование инвариантности относительно преобразования типа масштабирования выполнить несколько сложнее. Так, существуют методы прямого преобразования контурной информации в почти готовый вектор признаков, например, формирование цепного кода Фримана [7]. Однако для формирования вектора признаков напрямую эти данные использовать нельзя, так как сразу же возникает вопрос о нормализации размерности из-за того, что количество точек в каждом контуре различное. Задача может быть решена преобразованием шага сетки, но это требует существенных дополнительных вычислений.

В настоящей работе поставленная задача решается с использованием Фурье-преобразования для кусочно-линейных функций. Формой представления графического объекта $W(t)$ являются коэффициенты разложения функций $x(t)$ и $y(t)$ в ряд Фурье [8]. Поскольку компоненты x и y вектор-функции $W'(t)$ ортогональны и независимы, можно проводить Фурье-преобразование отдельно для каждой составляющей: $x'(t) = x(t) - x(1)$ и $y'(t) = y(t) - y(1)$. Так как анализ функции $x'(t)$ не отличается от анализа функции $y'(t)$, все дальнейшие рассуждения можно проводить относительно некоторой непрерывной функции $f(t)$.

Применив линейные преобразования сдвига и сжатия, получаем функцию $f(t')$, определенную на отрезке $[0, \pi]$. Таким образом, m точек контура отображаются на отрезок $[0, \pi]$ с различными угловыми значениями $f_i = f(t'_i)$ кусочно-линейной функции $f(t)$, где $t'_i = i\delta$, $i = \overline{0, m}$, $\delta = \pi/m$. В силу линейности преобразования количество исходных точек не меняется и информация не теряется.

Параллельно использование значений компонент Фурье-разложения автоматически решает и проблему ортогональности компонент вектора признаков.

Для построения компонент вектора признаков в качестве системы ортогональных базисных функций воспользуемся Фурье-преобразованием:

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt' + b_k \sin kt'). \quad (13)$$

Используя центрально-симметричное отображение исходной функции, можно избавиться от компонент $a_k \cos kt$.

В [9] показано, что коэффициенты разложения b_k для кусочно-линейной функции с фиксированным количеством точек N , которой является контур, имеют вид:

$$b_k = \frac{(-1)^k 2f(N)}{k\pi} + \frac{2N}{k^2 \pi^2} \sum_{i=0}^{N-1} f_i (\sin kt'_{i+1} - \sin kt'_i). \quad (14)$$

Из равномерной сходимости компонент Фурье-преобразования как $1/x$ следует необходимый для

практических расчетов вывод: при использовании коэффициентов разложения b_k в качестве компонентов вектора признаков их количество определяется точностью поставленной задачи [10].

Такой подход к формированию вектора признаков решает проблему нормализации размерности и параллельно имеет еще одно замечательное свойство – позволяет управлять качеством восстановленных изображений графических объектов. То есть, если нет необходимости получить максимальное качество восстановления изображений исходных объектов, достаточно просто уменьшить количество Фурье-компонент N , участвующих в процедуре восстановления контуров:

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^N b_k \sin kt. \quad (15)$$

В случае веб-публикации электронного документа сам процесс восстановления элементов шрифта происходит на стороне клиента. В процессе ретроконверсии в качестве значимой информации передаются набор Фурье-компонент векторов признаков – основная содержательная часть автоадаптивного шрифта, координаты расположения элементов шрифта и ссылка на номер элемента автоадаптивного шрифта.

Заключение

Построенная с использованием данного подхода технология ретроконверсии неформатных сканированных графических документов подтвердила свою работоспособность и высокое качество при обработке информационных фондов электронной библиотеки «Научное наследие России». В качестве существенного результата можно указать сокращение размеров хранимых электронных документов, возможность регулировать качество восстановления (для печати или только для просмотра на веб-сайте).

Дальнейшие перспективы развития изложенного подхода к обработке неформатных докумен-

тов представляются в организации контекстного поиска по содержимому сканированных графических файлов, что может иметь практическое значение при организации рабочего места специалиста по наполнению фондов электронных библиотек. Также представляется интересным развитие данной технологии в направлении генерации документа произвольного содержания в стиле обработанного исходного документа.

Исследования выполнены в МСЦ РАН – филиале ФГУ ФНЦ НИИСИ РАН при поддержке РФФИ (проект № 17-07-00400).

Литература

1. Научное наследие России. URL: <http://e-heritage.ru> (дата обращения: 27.09.2017).
2. Каленов Н.Е., Савин Г.И., Серебряков В.А., Сотников А.Н. Принципы построения и формирования электронной библиотеки «Научное наследие России» // Программные продукты и системы. 2012. № 4. С. 30–40.
3. Сотников А.Н., Чердниченко И.Н. Построение автоадаптивного фонта в документах электронных библиотек // Программные продукты и системы. 2008. № 2. С. 16–19.
4. Журавлев Ю.И. Об алгебраическом подходе к решению задач распознавания и классификации // Проблемы кибернетики. 1978. Вып. 33. С. 5–68.
5. Bereznev V.A., Sotnikov A.N., Cherednichenko I.N. A probabilistic criterion and an integral method in graphic-object recognition and information retrieval. Pattern Recognition and Image Analysis. Interperiodika, 1994, vol. 4, no. 1, pp. 32–35.
6. Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М.: Наука, 1976. 544 с.
7. Freeman H. On the encoding of arbitrary geometric configurations. IRE Transactions on Electronic Computers EC. 1961, vol. 10, pp. 260–268.
8. Жук В.В., Натансон Г.И. Тригонометрические ряды Фурье и элементы теории аппроксимации. Л.: Изд-во Ленингр. ун-та, 1983. 188 с.
9. Березнев В.А., Волков А.Ю., Чердниченко И.Н. Об использовании преобразования Фурье в задаче распознавания рукописного текста // Вопросы моделирования и анализа в задачах принятия решений: сб. М.: Изд-во ВЦ РАН, 2003. С. 153–159.
10. Березнев В.А., Волков А.Ю., Чердниченко И.Н. О выборе параметров в алгоритме распознавания раздельного рукописного текста // Вопросы моделирования и анализа в задачах принятия решений: сб. М.: Изд-во ВЦ РАН, 2004. С. 136–143.

ON THE PROBLEM OF NON-FORMAT DOCUMENTS RETRO-CONVERSION IN DIGITAL CONTENT FORMATION AND PROVISION SYSTEMS

*A.N. Sotnikov*¹, Dr.Sc. (Physics and Mathematics), Professor, ASotnikov@jssc.ru
*I.N. Cherednichenko*¹, Ph.D. (Engineering), inch@jssc.ru

¹ Joint Supercomputer Center of RAS, Lenin Ave. 32a, Moscow, 119991, Russian Federation

Abstract. The paper considers mathematical reasoning of solving the problem of scanned non-format documents retro-conversion (documents recovery), as applied to the task of filling and publishing collections of digital libraries. It also describes the algorithms of automatic clustering and forming auto-adaptive fonts for documents retro-conversion.

As a mechanism for working with clustered objects, the authors use a classical approach to form a feature vector based on the estimates calculating algorithm model that describes each object under consideration. The algorithm for automatic clustering is based on a statistical approach to forming characteristic vectors of each cluster and to the dynamic correction of separating hyperplanes.

Due to the specifics of the tasks being solved, the authors use a contour approach to graphic objects processing, which allows applying frequency analysis methods based on the Fourier transform. The paper presents estimates of the necessary and sufficient dimension of a feature vector used both for describing graphic objects and for their recovery with minimal distortion after generating output documents during a retro-conversion procedure.

Finally, the article shows possible prospects for development of the described method of non-format documents retro-conversion in the field of organization of contextual search within graphic documents.

Keywords: retro-conversion, auto-adaptive fonts, non-format documents, automatic clustering, electronic libraries.

Acknowledgements. The research on the problem of non-format documents retro-conversion in digital content formation and provision systems have been performed in JSCC RAS (a branch of FGA Federal Research Center Research Institute of System Studies of RAS and supported by the Russian Foundation for Basic Research (project no. 17-07-00400).

References

1. *Nauchnoe nasledie Rossii* [The Scientific Heritage of Russia]. Available at: <http://e-heritage.ru> (accessed September 27, 2017).
2. Kalenov N.E., Savin G.I., Serebryakov V.A., Sotnikov A.N. The principles of building and forming the electronic library "The Scientific Heritage of Russia". *Programmnye produkty i sistemy* [Software & Systems]. 2012, no. 4, pp. 30–40 (in Russ.).
3. Sotnikov A.N., Cherednichenko I.N. Building an auto-adaptive font in documents of electronic libraries. *Programmnye produkty i sistemy* [Software & Systems]. 2008, no. 2 (82), pp. 16–19 (in Russ.).
4. Zhuravlev Yu.I. On the algebraic approach to solving recognition and classification problems. *Problemy kibernetiki* [The Problems of Cybernetics]. 1978, iss. 33, pp. 5–68 (in Russ.).
5. Bereznev V.A., Sotnikov A.N., Cherednichenko I.N. A probabilistic criterion and an integral method in graphic-object recognition and information retrieval. *Pattern Recognition and Image Analysis*. Moscow, Interperiodika Publ., 1994, vol. 4, no. 1, pp. 32–35.
6. Kolmogorov A.N., Fomin S.V. *Elementy teorii funktsiy i funktsionalnogo analiza* [Elements of the Theory of Functions and Functional Analysis]. 4th ed. Moscow, Nauka Publ., 1976, 544 p.
7. Freeman H. On the encoding of arbitrary geometric configurations. *IRE Trans. on Electronic Computers EC*. 1961, vol. 10, pp. 260–268.
8. Zhuk V.V., Natanson G.I. *Trigonometricheskie ryady Fur'ye i elementy teorii approksimatsii* [Trigonometric Fourier Series and Elements of the Approximation Theory]. Leningrad, Leningr. Univ. Publ., 1983, 188 p.
9. Bereznev V.A., Volkov A.Yu., Cherednichenko I.N. On the use of the Fourier transform in the handwriting recognition problem. *Voprosy modelirovaniya i analiza v zadachakh prinyatiya resheny: sb.* [Proc. Modeling and Analysis Issues in Decision-Making Problems]. Moscow, 2003, pp. 153–159 (in Russ.).
10. Bereznev V.A., Volkov A.Yu., Cherednichenko I.N. On the choice of parameters in the algorithm for recognizing separate handwriting text. *Voprosy modelirovaniya i analiza v zadachakh prinyatiya resheny: sb.* [Proc. Modeling and Analysis Issues in Decision-Making Problems]. Moscow, VTS RAN Publ., 2004, pp. 136–143 (in Russ.).

Примеры библиографического описания статьи

1. Сотников А.Н., Чередниченко И.Н. О задаче ретроконверсии неформатных документов в системах формирования и предоставления цифрового контента // Программные продукты и системы. 2017. Т. 30. № 4. С. 684–689. DOI: 10.15827/0236-235X.120.684-689.
2. Sotnikov A.N., Cherednichenko I.N. On the problem of non-format documents retro-conversion in digital content formation and provision systems. *Programmnye produkty i sistemy* [Software & Systems]. 2017, vol. 30, no. 4, pp. 684–689 (in Russ.). DOI: 10.15827/0236-235X.120.684-689.