

НОВЫЕ МЕТОДЫ И ТЕХНОЛОГИИ ОБРАБОТКИ ИСТОРИЧЕСКИХ ИСТОЧНИКОВ

NEW METHODS AND TECHNOLOGIES OF HISTORICAL
SOURCES PROCESSING

«VESTIGIUM» — КОМПЛЕКС ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ АНАЛИЗА НЕТЕКСТОВОЙ ИНФОРМАЦИИ РУКОПИСНЫХ ПАМЯТНИКОВ¹

“VESTIGIUM”, A SOFTWARE FOR EXAMINATION
OF NON-TEXTUAL DATA IN HANDWRITTEN DOCUMENTS

Исаев Борис Львович,

научный сотрудник Лаборатории
кодикологических исследований
и научно-технической экспертизы документа
Отдела рукописей Российской национальной
библиотеки.

E-mail: boris_isaev@mail.ru

Boris L. Isaev

Ляховицкий Евгений Александрович,

кандидат исторических наук, заведующий
Лабораторией кодикологических
исследований и научно-технической
экспертизы документа Отдела рукописей
Российской национальной библиотеки.

E-mail: stoglav@bk.ru

Evgenij A. Lyakhovitskiy

Цыпкин Денис Олегович,

кандидат исторических наук, доцент,
заведующий кафедрой истории
западноевропейской и русской культуры
Института истории Санкт-Петербургского
государственного университета.

E-mail: d.tsypkin@spbu.ru

Denis O. Tsypkin

¹ Исследование выполнено при поддержке гранта РФНФ № 13-01-12010.

Чиркова Александра Викторовна,
кандидат исторических наук, старший
научный сотрудник Научно-исторического
архива Санкт-Петербургского института
истории РАН.
E-mail: chirok-alexandr@list.ru

Alexandra V. Chirkova

Программный комплекс «Vestigium» предназначен для исследования памятников письменности посредством анализа их цифровых изображений. В состав комплекса входят следующие инструменты: 1) «измерения» предназначены для тонкой подстройки кривых, позволяя со значительно большей точностью и удобством измерять контуры со сложными очертаниями, а также углы между объектами на изображении; 2) «плотностная структура» позволяет визуализировать распределение плотностных значений путем выделения на самом изображении; 3) «цвет» дает возможность описывать цветовые характеристики через реконструкцию длин волн из данных цифрового изображения путем геометрического соотнесения координат цветового пространства; 4) «профиль» предназначен для реконструкции заточки примененного при исполнении документа письменного прибора, а также для оценки рельефа поверхности писчего материала; 5) «пропорции» предназначены для анализа пропорциональных соотношений различных объектов на изображении путем наложения на него мерной модульной сетки. Отдельная утилита *Vestigium PhotoPlate* предназначена для быстрого сбора образцов исследуемых объектов (восковых, пальцевых следов, элементов орнамента и пр.) из цифровых снимков документов.

Ключевые слова: палеография, кодикология, почерковедение, оптико-электронная экспертиза, модульный анализ письма.

“Vestigium” is the software for manuscript study that works by means of analysing digital images. The software has the following tools: 1) *Measurement* is used for fine-tuning curves. It can measure complicated external shapes and angles between the objects in the image with greater precision; 2) *Density structure* visualises the distribution of density values by means of sorting image pixels according to their density values; 3) *Colour* is particularly efficient if used to define colour through wavelength reconstruction of the image data; 4) *Profile* evaluates the frequency and differential of density values within the digital image. It helps to reconstruct the way of sharpening the pen or any other writing instrument. It would evaluate the surface of the writing material; 5) *Proportions* analyses proportional correlation of different shapes within the image by dimensional modular grid application. The separate utility *Vestigium PhotoPlate* would rapidly take the specimens of objects (vestiges of wax, fingerprints, ornamental elements, etc.) from the digital images.

Key words: paleography, codicology, handwriting studies, optic and electronic examination, module analysis of writing.

Задачами историко-документной экспертизы являются максимально точная датировка и локализация по месту производства исторического документа и его элементов; установление его подлинности; идентификация создателей документа в целом и его элементов; реконструкция процесса производства и бытования документа. Все эти задачи требуют анализа нетекстовой информации памятника. Одним из основных условий для анализа является объективный формализованный учет этой информации. Эра цифровых копий рукописных памятников представляет богатые возможности в этом отношении, но чтобы реализовать эти возможности, необходимы специализиро-

ванные программные инструменты. К сожалению, в настоящее время наблюдается явный дефицит последних. Ученый, специализирующийся на исследовании нетекстовой информации письменных памятников, должен пользоваться универсальными программными инструментами, ориентированными на дизайнеров и фотографов, такими как *Adobe Photoshop*, или на широкий круг исследовательских задач (прежде всего естественно-научных), как *ImageJ*.

Важным шагом на пути преодоления этой ситуации стала разработка итальянскими специалистами инструмента *Graphoskop* — плагина для *ImageJ*, добавляющего в данный программный комплекс спе-

специализированные функции линейных измерений¹. При наличии в этом плагине весьма интересных и полезных функций, среди которых следует особенно выделить возможность оценивать «вес письма» путем анализа соотношения «темных» и «светлых» пикселей в изображении, *Graphoskop* ориентирован на решение достаточно узкого спектра задач, находящихся в основном в сфере линейных измерений.

Программный комплекс *Vestigium*², разработанный в ходе совместного проекта СПбИ РАН и РНБ, представляет собой первый самостоятельный специализированный комплекс программного обеспечения, позволяющий решать широкий спектр исследовательских задач в области исследования рукописных памятников³.

Программный комплекс включает в себя несколько аналитических инструментов.

«Измерения». По сравнению с аналогичными инструментами других программных комплексов данный инструмент обладает возможностью тонкой подстройки кривых, позволяя со значительно большей точностью и удобством измерять контуры со сложными очертаниями (особенно с закруглениями). Принцип работы данного инструмента аналогичен инструменту «paths» в графическом редакторе *Adobe Photoshop*, где не предусмотрена возможность использовать этот инструмент для измерений. Данный инструмент комплекса *Vestigium* может найти себе применение в любых исследованиях, где необходимо измерение объектов на цифровом изображении.

Инструмент «Измерения» имеет два режима работы: «курвиметр» и «угломер». «Курвиметр» — режим для линейного измерения двухмерных объектов на цифровом изображении. Выполнение измерений обеспечивается заложенным в программную оболочку механизмом задания масштаба для цифрового изображения (его предлагается задать для каждого открываемого файла). Масштаб можно задать двумя способами: «фотография» и «сканер». При выборе способа «фотография» пользователь должен установить масштаб, воспользовавшись тест-объектом (например, линейкой), введенным в кадр при фотосъемке. При выборе способа определения масштаба «сканер», масштаб задается исходя из разрешения сканирования в точках на дюйм (dpi). Собственно процесс измерения состоит в следующих процедурах: 1) пользователь наносит на изображение контур, который необходимо измерить, отмечая на изображении точки, которые соединяются линией; 2) после этого пользователь имеет возможность включить механизм редактирования контура по якорным точкам (режим «искривление») (см. рис. 1).

Режим «угломер» предполагает измерение углов на цифровом изображении. Пользователь дол-

жен обозначить требуемый угол на изображении, отметив точку, определенную как вершину угла, и еще две точки, необходимые для построения сторон угла. Интерфейс инструмента предполагает возможность инверсии построенного угла.

«Плотностная структура». В отличие от инструментов анализа плотностей других программных комплексов, которые способны визуализировать распределение плотностей только в виде гистограммы, данный инструмент позволяет визуализировать распределение плотностных значений путем выделений на самом изображении. Пиксели с различными значениями плотности могут быть объединены в группы и показаны единой заливкой на изображении. Возможно выделение любого необходимого числа групп, выделяемых на изображении заливкой разного цвета (см. рис. 2). Интерфейс инструмента позволяет отображать выделенные группы и фон в любых сочетаниях исходя из ситуации.

Данный инструмент предназначен прежде всего для почерковедческих исследований (выявление нажимных характеристик), а также для оценки равномерности просвета бумажных листов (которая является одной из важнейших интегральных технологических характеристик бумаги), но может быть востребован в любых исследованиях, где необходимо зонирование изображения по оптическим плотностям. Инструмент также может применяться для обработки изображений (в частности, для создания масок для сложных выделений в графических редакторах).

«Цвет». Данный инструмент значительно отличается от инструментов анализа цветовых характеристик других программных комплексов. Цветовые характеристики описываются не координатами в цветовых пространствах, а значениями основных физических характеристик — длины волны и насыщенности, которые реконструируются из данных цифрового изображения путем геометрического соотношения координат цветового пространства⁴. Это позволяет в определенной степени представить физические характеристики исследуемого объекта без применения специального оборудования (спектрофотометров и пр.). Значения в цветовых пространствах RGB, XYZ и значения яркости также могут быть выведены в качестве дополнительной информации. Инструмент дает возможность измерения цвета как с усреднением для выделенной области, так и путем детального описания состава изображения по цветовым характеристикам составляющих его точек. Данный инструмент будет востребован в любых исследованиях, где в зоне интереса находятся цветовые характеристики объекта.

В дизайн «Цвет» инструмента заложена возможность корректировки баланса белого анализируемого изображения, которая позволяет сопоставлять изображения, снятые с разными установками баланса белого или при разном освещении. Всем измерениям предшествует соответствующая изображению настройка источника света. Для этого следует нажать на информационное окно, где выведена текущая установка источника света. Появится окно «Настройка источника света», где следует выбрать установку баланса белого из пяти возможных:

«Белый» — при выборе данного источника света предполагается, что баланс цвета в цифровом изображении соответствует цветовой температуре 5000 К;

«А» — 2856 К воспроизводит условия искусственного освещения электрическими лампами накаливания;

«В» — 4874 К воспроизводит условия прямого солнечного освещения;

«С» — 6774 К воспроизводит условия освещения рассеянным дневным светом;

«Специальный» позволяет установить баланс белого для измерения, воспользовавшись введенным в кадр нейтральным тест-объектом.

Инструмент «Цвет» имеет два режима работы: «Колориметр» и «Цветоделение». Режим «Колориметр» предполагает измерение усредненной цветовой характеристики выделенной на изображении области (при этом выделенная область будет залита соответствующим цветом). Результаты измерения — вербальное определение цвета (например, «фиолетовый»), реконструированная длина волны и насыщенность цвета — выводятся на главном окне. Для вербального определения цветовой характеристики в *Vestigium* используется следующее разделение длин волн:

- синий: 436 нм — 480 нм
доп. 570 нм — доп. 567 нм
- голубой: 481 нм — 506 нм
- зеленый: 507 нм — 564 нм
- желтый: 565 нм — 581 нм
- оранжевый: 582 нм — 600 нм
- красный: 601 нм — 700 нм
доп. 500 нм — доп. 493 нм
- фиолетовый: доп. 506 нм — доп. 501 нм

Интерфейс инструмента предусматривает возможность вывода описания данного цвета в пространствах RGB и XYZ, а также значения яркости (см. рис. 3).

Режим работы «Цветоделение» предполагает подразделение точек, входящих в изображение или выделенную на нем область исходя из их цветowych характеристик. Результаты анализа выводятся в виде списка, который имеет два варианта отображения: «волны» (открывается по умолчанию) и «по

группам». При первом варианте список выводится в трех колонках: «волна» (т. е. реконструированная длина волны), группа (согласно вербальным определениям цветowych характеристик), количество пикселей с данными значениями. При втором варианте список выводится в двух колонках: группы и количество (для данной группы) (см. рис. 4а). При выделении одной или нескольких строк в списке соответствующие точки будут выделены на изображении цветной заливкой (см. рис. 4б). Цвет заливки данной точки определяется ее принадлежностью к той или иной группе. Невыделенные точки можно скрыть (см. рис. 4в).

«Профиль». Данный инструмент предназначен для реконструкции заточки примененного при исполнении документа письменного прибора, а также для оценки рельефа поверхности писчего материала. Принцип работы данного инструмента — расчет частоты и степени перепадов между различными плотностными значениями в составе цифрового изображения. Предусмотрено три режима работы инструмента:

- а) «горизонтальный» — измерение производится по поперечному срезу изображения (см. рис. 5а);
- б) «вертикальный» — измерение производится по вертикальному срезу (см. рис. 5б);
- в) «произвольный» — измерение производится по отрезку произвольной длины и расположения (см. рис. 5в).

Во всех трех режимах при измерении отображается диаграмма, описывающая перепад плотностей в рамках измеряемого отрезка. Предусмотрена возможность вывести значение плотности конкретной точки, а также рассчитать расстояние между двумя точками. Основные результаты измерения выводятся в виде «Коэффициента изрезанности» и «Коэффициента неравномерности».

«Коэффициент изрезанности» — по аналогии с понятием, использующимся в топографии для описания береговой линии водоемов, — описывает отклонение линии поверхности от идеально ровной. «Коэффициент неравномерности» описывает степень выраженности рельефа (через соотношение крайних плотностных значений). Для облегчения статистической обработки данных выводится также интегральный показатель, описывающий рельеф измеряемой поверхности — отношение коэффициентов изрезанности и неравномерности.

«Пропорции». Данный инструмент, не имеющий аналогов, предназначен для анализа пропорциональных соотношений различных объектов на изображении путем наложения на него мерной модульной сетки. Данный инструмент разработан

для исследования письма, его пропорционального строя, а также анализа пропорциональных отношений различных элементов композиции листа кодекса (поля текста, полей, заставки и пр.).

Инструмент «пропорции» имеет два режима работы: «квадратная сетка» и «прямоугольная сетка». Режим «Квадратная сетка» подразумевает отрисовку сетки на изображении по отмерянному пользователем отрезку — «шагу сетки» (см. рис. 6). В режиме «Прямоугольная сетка» пользователь отмеряет на изображении два отрезка — по оси X и по оси Y, исходя из которых происходит построение сетки с неравносторонними ячейками. Интерфейс инструмента предполагает два варианта нанесения сетки на изображение — с автоматически назначенной точкой отсчета шагов сетки (в этом случае точкой отсчета будет крайняя правая нижняя точка), или с точкой, назначенной пользователем вручную (кликом на изображении). Имеет

ся возможность изменения цвета сетки. Сохранить результаты работы с инструментом можно в форме снимка рабочего окна программы или копии анализируемого изображения со всей нанесенной на него в момент сохранения разметкой (доступно сохранение в форматах tif и jpeg).

«*Vestigium PhotoPlate*». Отдельная утилита в составе комплекса программного обеспечения *Vestigium*, предназначенная для быстрого сбора образцов исследуемых объектов (восковых, пальцевых следов, элементов орнамента и пр.) из цифровых снимков документов. Позволяет объединять выбранные образцы в форме «фототаблиц» (см. рис. 7а, б), которые можно затем сохранить как изображение в формате tif или jpeg.

В дальнейшем предполагается развитие программного комплекса в сторону усиления возможностей каталогизации, автоматизированной поточной обработки изображений, поисковых функций.

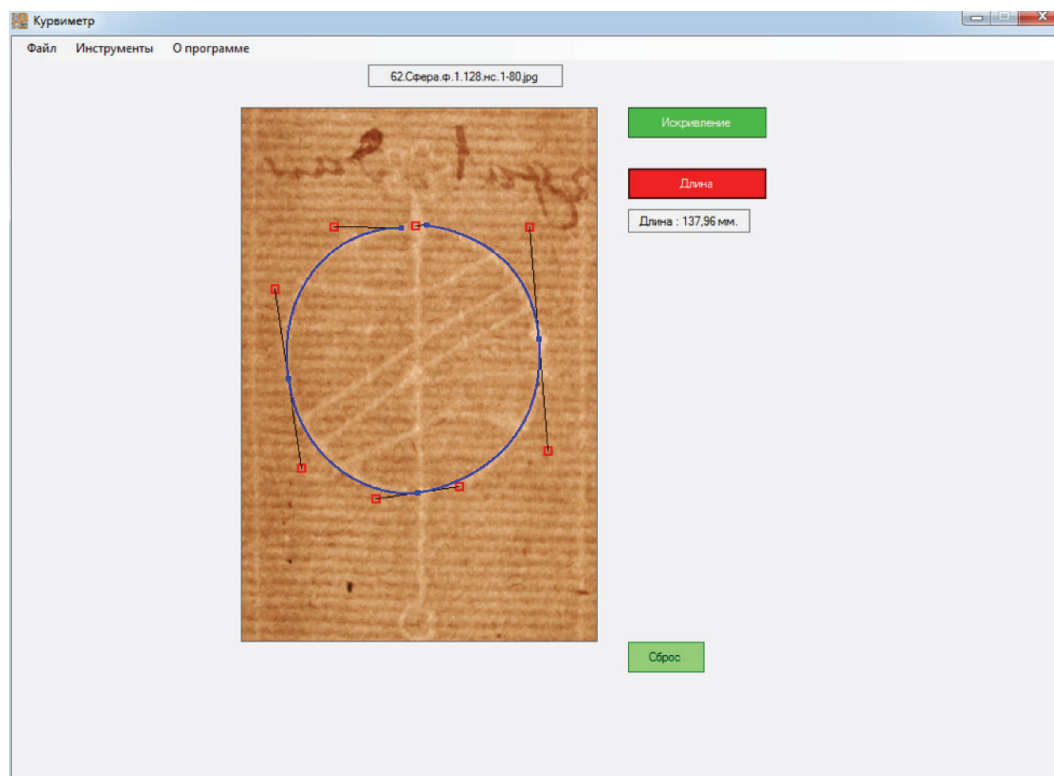


Рис. 1. Пример работы инструмента «Измерение» в режиме «Курвиметр»: измерение длины отпечатка проволоки, образующего окружность водяного знака «сфера» (ОР РНБ. Собр. П. П. Дубровского. Авт. 62. Л. 128)

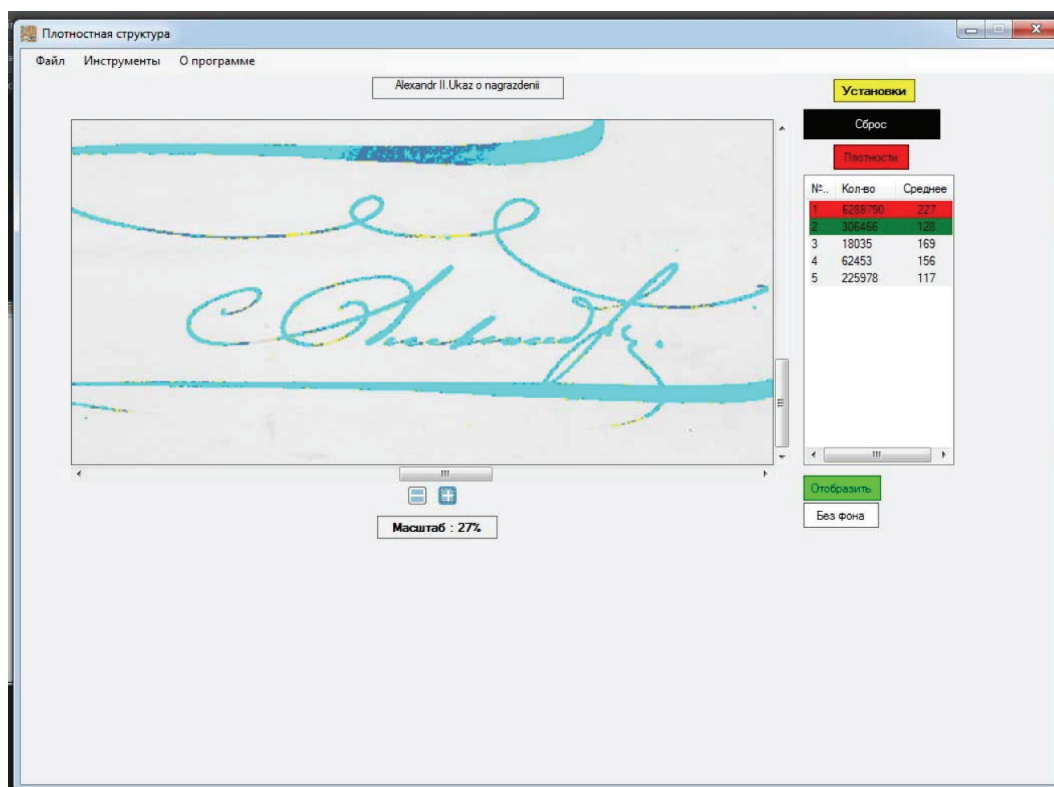


Рис. 2. Пример работы инструмента «Плотностная структура»: анализ нажима в подписи императора Александра II (ОР РНБ. Ф. 341. Д. 347. Л. 1об.)

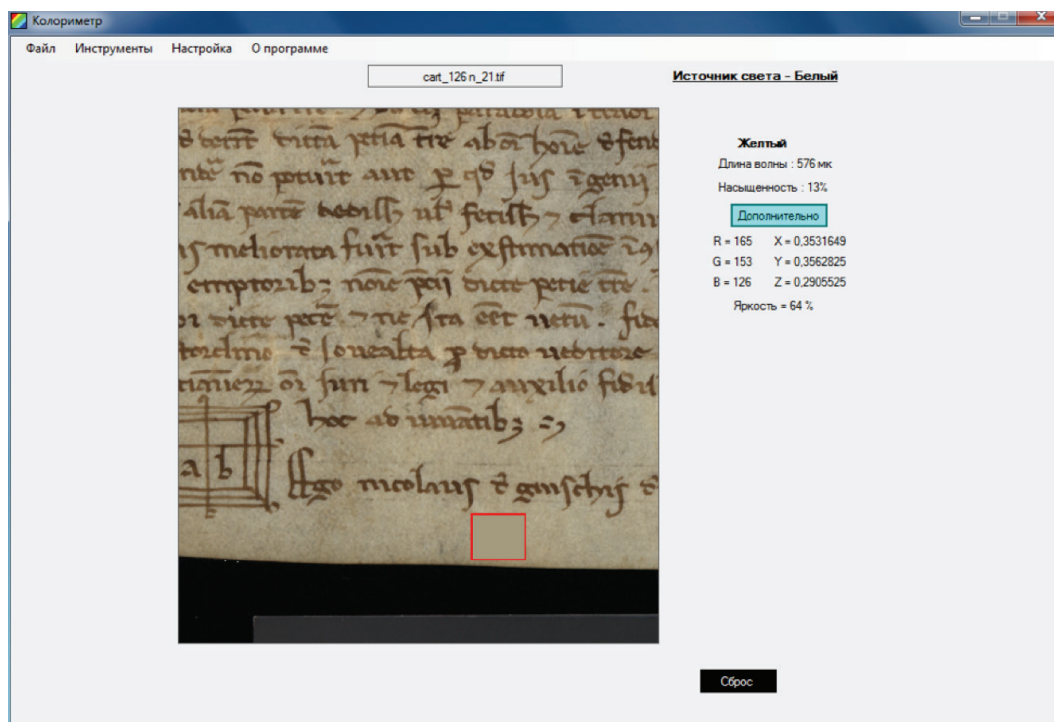
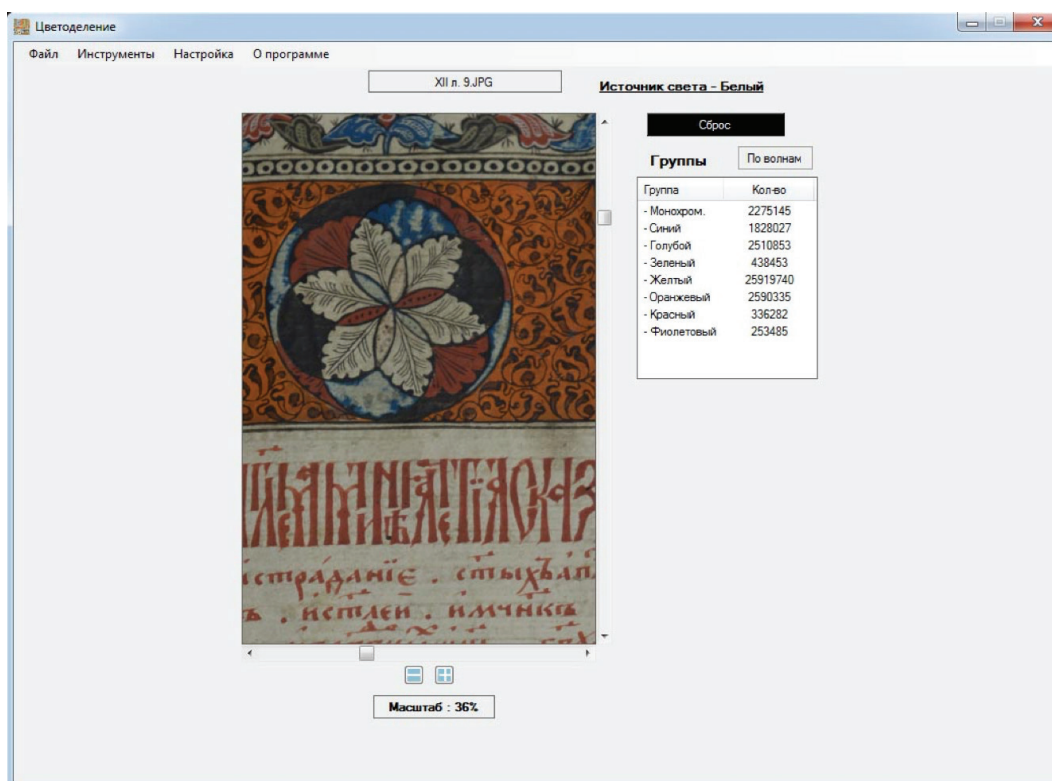
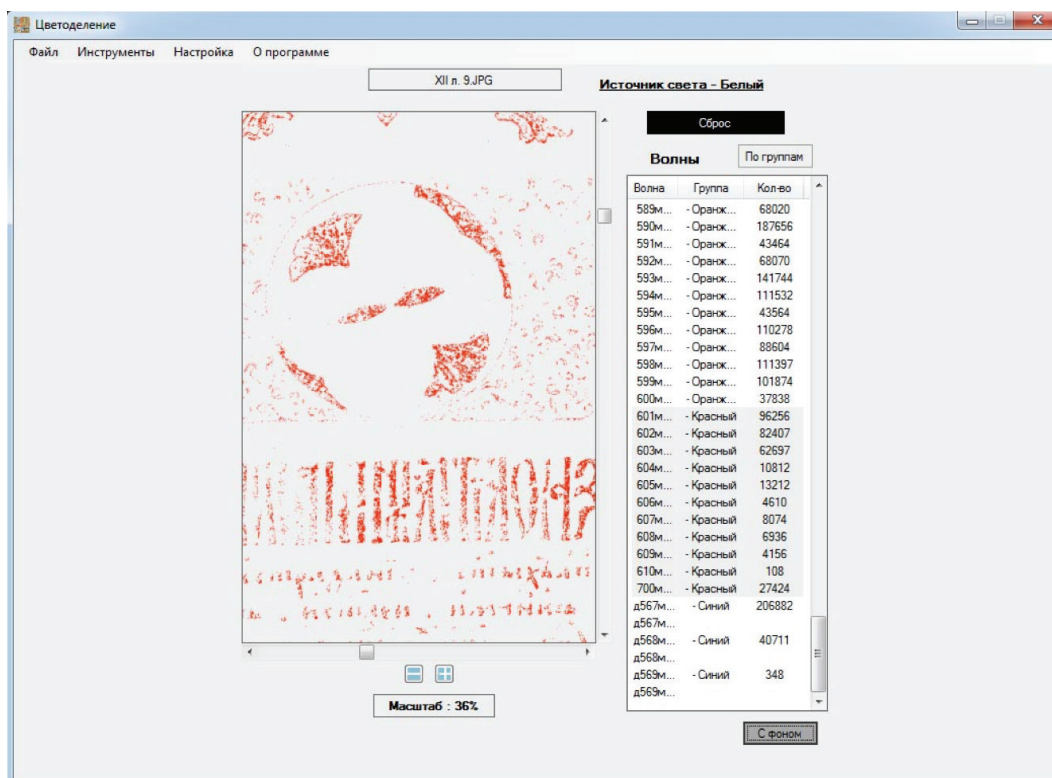


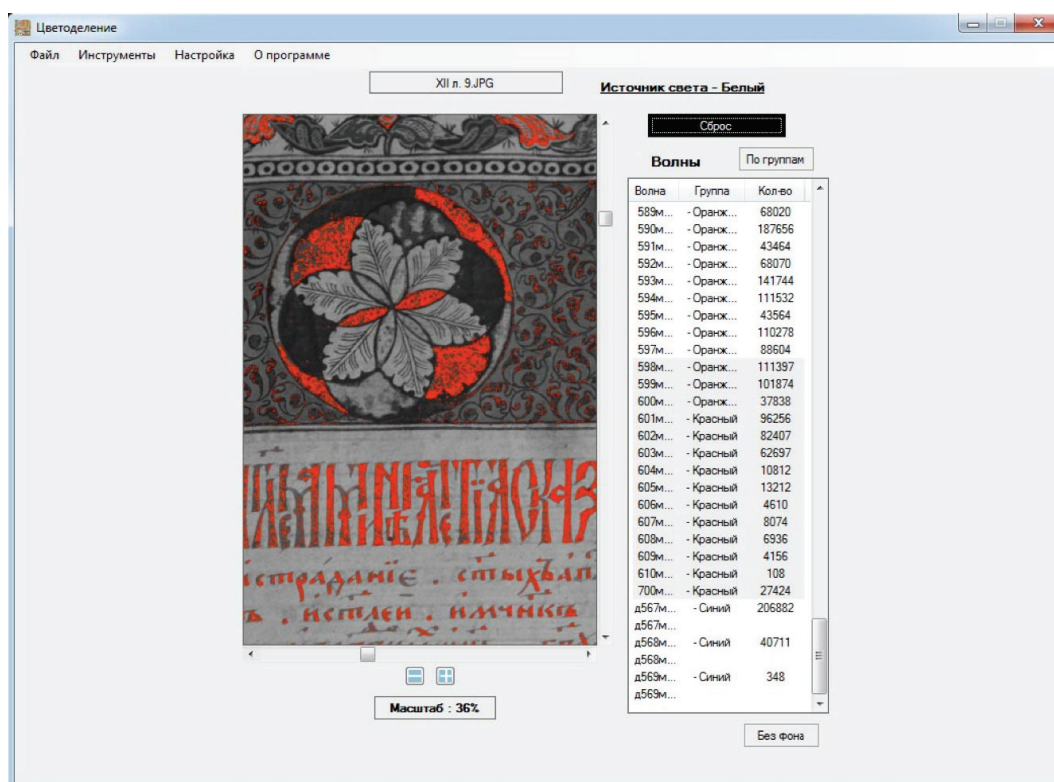
Рис. 3. Пример работы инструмента «Цвет» в режиме «Колориметр»: определение цветовых характеристик пергамена (использован нотариальный акт 1244 г.: Западноевропейская секция НИА СПб ИИ РАН. Ф. 4: Акты Кремоны. Карт. 126. № 21)



а)



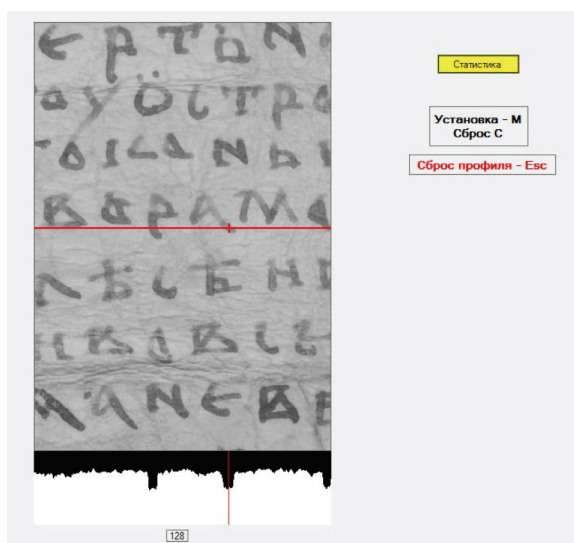
б)



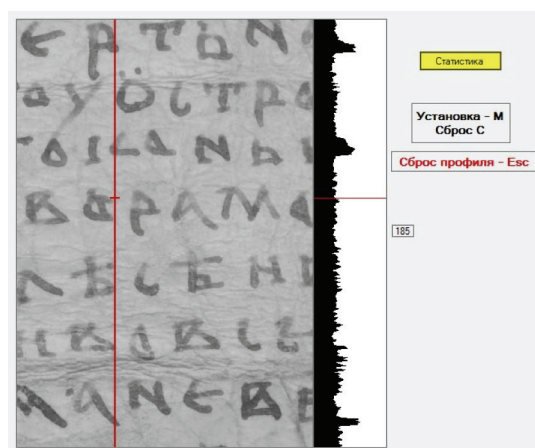
в)

Рис. 4. Пример работы инструмента «Цвет»
в режиме «Цветоделение» (использован фрагмент рукописной книги:
ОР РНБ. Библиотека Новгородского Софийского собора. № 1318
(октябрьский том Софийского комплекта Великих Миней Четьих). Л. 9):

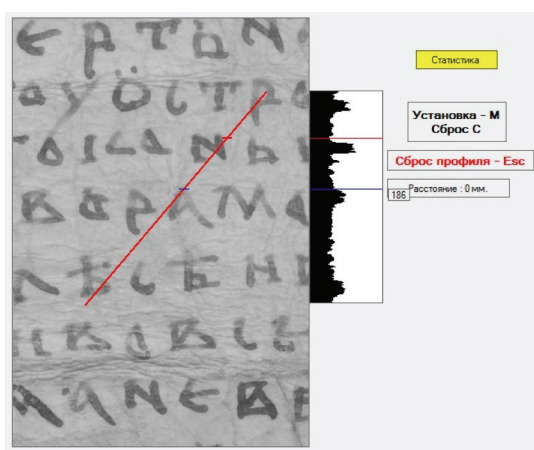
а) отображение данных о цветовых характеристиках точек изображения по группам; б) отображение данных о цветовых характеристиках точек изображения по реконструированным длинам волн; цветовая заливка точек, выделенных в списке; в) отображение выделенных в списке точек без фона



а)



б)



в)

Рис. 5. Пример работы инструмента «Профиль»
 (использован фрагмент пергаменной грамоты: ОР РНБ.
 Особое собрание актов и грамот. Оп. 1. Д. 11. Л. 1): а) режим «горизонтальный»;
 б) режим «вертикальный»; в) режим «произвольный»

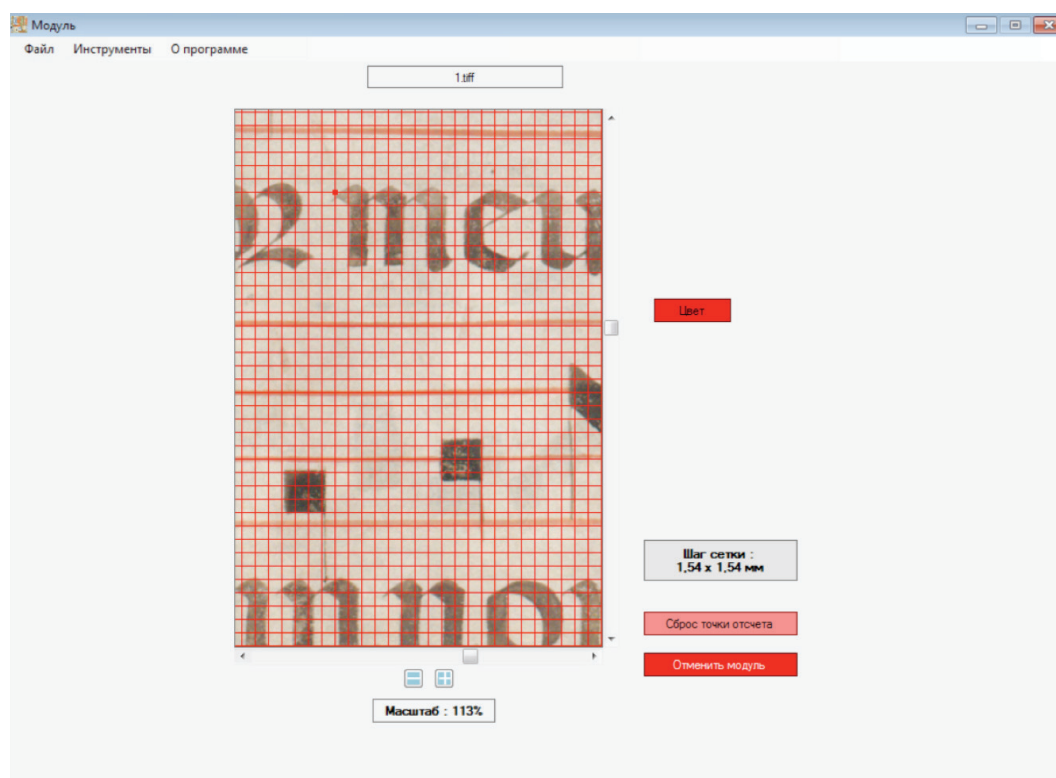
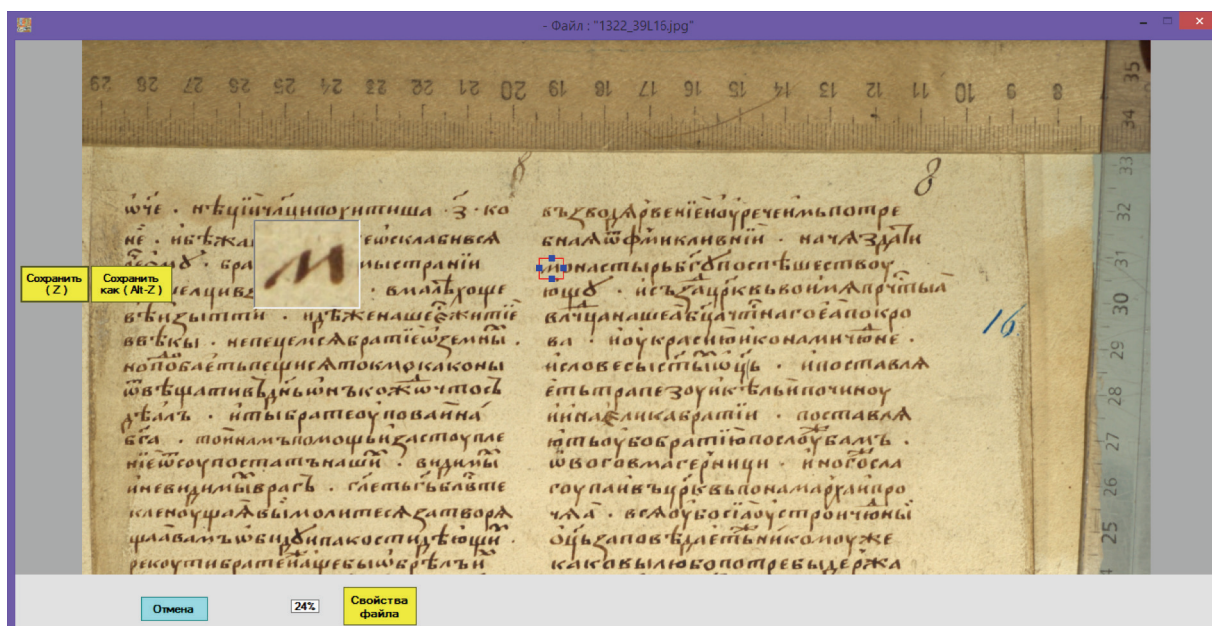
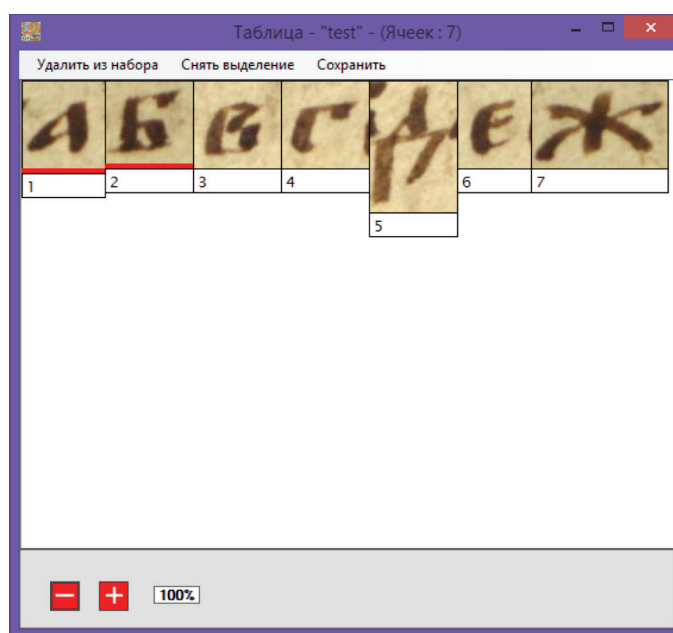


Рис. 6. Пример работы инструмента «Модуль» в режиме «квадратная сетка»: анализ пропорций письма (использован фрагмент болонского антифонария второй четверти XIV в.: Западноевропейская секция НИА СПбИИ РАН. Кол. 46 «Рукописная книга». Карт. 562. № 1. Л. 102)



а)



б)

Рис. 7. Пример работы утилиты «Vestigium PhotoPlate» (использован фрагмент рукописной книги: ОР РНБ. Библиотека Новгородского Софийского собора. № 1322 (июньский том Софийского комплекта Великих Миней Четьих): а) процесс сбора образцов; б) таблица с собранными образцами

ПРИМЕЧАНИЯ

- ¹ Graphoscope [Electronic Resource]. URL: <http://www.palaeographia.org/graphoskop/index.htm>. См.: Gurrado M. «Graphoskop», uno strumento informatico per l'analisi paleografica quantitativa // Kodikologie und Paläographie im digitalen Zeitalter / Codicology and Palaeography in the Digital Age / Hrsg. von M. Rehbein, P. Sahle, T. Schaßan. Norderstedt, 2009. P. 251–259.
- ² Vestigium (лат.) — след. Объекты исследования, для которых разрабатывалось ПО, представляют собой оставленные человеком следы в процессе создания рукописных памятников и их использования в прошлом. Для заставки был выбран фрагмент инициала из хоровой книги XIII в. (Западно-европейская секция НИА СПбИИ РАН. Кол. 47 «Листы пергаменных кодексов». Карт. 622. № 13).
- ³ См. подробное описание комплекса программного обеспечения и находящийся в открытом доступе инсталлятор на сайте СПбИИ РАН: URL: <http://www.spbiiran.nw.ru/vestigium/>
- ⁴ При разработке колориметрической составляющей программного комплекса *Vestigium* разработчики опирались на работу: Гуревич М. М. Цвет и его измерение. М. ; Л., 1950. Для тех цветовых объектов, положение которых выходит за границы используемого в программе цветового тела, применяются условные (дополнительные) значения длин волн (обозначаются как «доп.»).