

СканКромсатор v5.93 Пособие по программе

Неофициальная версия. Составил **monday2000** на основании объяснений автора программы, собранных на известных форумах (с неоценимой помощью **Arcand** и других), и на основании опыта использования программы. *Вследствие всего этого возможны ошибки и неточности.*

ВАЖНО: В Пособии даётся описание только таких элементов управления программы, про которые точно известно, для чего они и как они работают. В основном приводится описание наиболее необходимых и популярных возможностей программы. В отдельных случаях некоторые возможности программы **не описаны умышленно** - если они плохо работают, или же малопопулярны, и т.п.

Актуального и официального хелпа к программе СканКромсатор от автора программы (**bolega**) пока не существует и ожидается в будущей (6-ой) версии программы.

1 Основной интерфейс программы	3
1.1 Управляющие Tab-вкладки	3
1.1.1 Вкладка Pages	3
1.1.2 Вкладка Book	5
1.1.3 Вкладка Files	5
1.1.4 Вкладка Options	6
1.1.5 Вкладка Options 2	9
1.1.6 Вкладка Convert	10
1.2 Общие элементы	11
1.2.1 Элементы управления списка файлов	12
1.2.2 Виды выделений файлов в списке:	12
1.2.3 Групповое изменение опций	13
1.2.4 Управление списком файлов	13
1.2.5 Панель свойств загруженных файлов	14
1.2.6 Кнопки панели управления	14
1.2.7 Схема расположения резаков	15
1.2.8 Панель управления резаками	15
1.2.9 Управление резаками	16
1.2.10 Контекстное меню ПОЛЗУНКА любого резака	17
1.2.11 Пункты контекстного меню (главного окна)	17
1.3 Пункты меню	21
1.3.1 Пункт меню File	21
1.3.2 Пункт меню Edit	22
1.3.3 Пункт меню Result	22
2 Интерфейс редактора постобработки	23
2.1 Общие элементы	23
2.2 Пункты контекстного меню	25
2.2.1 Общее контекстное меню	25
2.2.2 Очистка сканов от грязи при помощи мыши	27
2.2.3 Окно поворота выделенного фрагмента	28
2.2.4 Контекстное подменю опций	29
3 Описание работы с программой	30
3.1 Окончательная обработка	31
3.2 Что такое Draft kromsate:	32
3.3 Что такое кнопка Process!	32
3.4 Как работают резаки	32
4 Version 5.93 Что нового	33

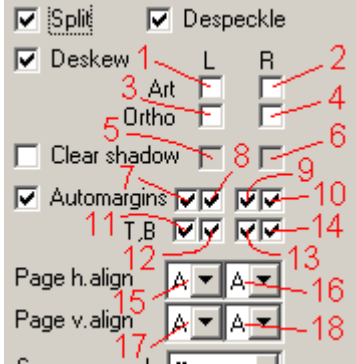
1 Основной интерфейс программы

1.1 Управляющие Tab-вкладки

1.1.1 Вкладка Pages

<div> <div>Pages Book Files</div> <div> <input type="checkbox"/> Split <input checked="" type="checkbox"/> Despeckle </div> <div> <input checked="" type="checkbox"/> Deskew <div> <div>L</div> <div>R</div> </div> <div> <div>Art</div> <div>Ortho</div> </div> </div> <div> <input type="checkbox"/> Clear shadow <div> <div></div> <div></div> </div> </div> <div> <input checked="" type="checkbox"/> Automargins <div> <div></div> <div></div> </div> <div> <div>T,B</div> <div></div> </div> </div> <div> <div>Page h.align</div> <div> <div>A</div> <div>A</div> </div> </div> <div> <div>Page v.align</div> <div> <div>A</div> <div>A</div> </div> </div> <div> <div>Spec. gaps</div> <div> <div>L</div> <div>R</div> </div> </div> <div> <input type="checkbox"/> Invert <input type="checkbox"/> Flip <div>special...</div> </div> </div>	Элемент	Описание
	Split	<p>Разделить страницу - Эта опция включает/выключает действие резака "Set internal margin or split line". Применяется для разрезания сдвоенных разворотов, или же для отрезания вертикальной полосы - ошмётка соседней страницы - в случае одиночного скана.</p>
	Deskew	<p>Автоматическое устранение перекоса. Рекомендуется всегда использовать этот режим при самой первой обработке сканов. Алгоритм - Hough transform с точностью 0.1.</p>
	Despeckle	<p>Автоматическое удаление соринки. Работает только на чёрно-белых сканах - а на серых и цветных не даёт никакого эффекта - это особенность самого алгоритма. Не рекомендуется к использованию на ч.б. сканах плохого качества - может удалить участки букв или грамматические точки.</p>
	Art	<p>Доп. режим работы Deskew - для страниц со сложными художественными рисунками, таблицами и сильно разреженных страниц. Применяйте в случаях, когда обычный Deskew неправильно определяет угол перекоса страницы.</p> <p>Если же и при включённом Art угол перекоса страницы определяется неправильно - тогда нужно заключить присутствующие на скане картинки или таблицы в Exclude-зону - чтобы исключить их из расчёта угла перекоса.</p> <p>Art имеет один недостаток - выполняется в два-три раза медленнее стандартного алгоритма. Поэтому не стоит применять его для всех страниц подряд.</p>
	Ortho	<p>Доп. режим работы Deskew - для страниц, содержимое которых расположено перпендикулярно, т.е. развёрнуто на 90 градусов. Допускается использовать одновременно с Art.</p>
	Automargins	<p>Распознающий режим резаков Эта опция должна быть всегда включена. Точное её назначение неизвестно.</p>

	Page h.align	Горизонтальное автоматическое выравнивание текста по площади скана. Может принимать значения: A (Auto) - в действительности это Лево; L (Left) - Лево; C (Center) - Центр; R (Right) - Право.
	Page v.align	Вертикальное автоматическое выравнивание текста по площади скана. Может принимать значения: A (Auto) - в действительности это Верх; T (Top) - Верх; C (Center) - Центр; B (Bottom) - Низ.

	<p>Если приглядеться, то видно, что на этом рисунке все пронумерованные элементы управления построены (автором программы) в 2 колонки - обозначенные буквами L и R. Это подопции левой (L) и правой (R) страниц. Понятия "левая" и "правая" страницы имеют смысл только при обработке сдвоенного разворота - когда одновременно с разрезанием обрабатываются образовавшиеся левая и правая половинки разворота. Активностью элементов в колонке R управляет галка Split - когда она включена, то все элементы в колонке R становятся активными - и наоборот. В случае одиночного разворота (т.е. когда ничего не нужно разрезать) этот одиночный разворот условно считается левой страницей. Итак, в колонке L находятся элементы с номерами 1, 3, 5, 7, 8, 11, 12, 15, 17; а колонке R - элементы с номерами 2, 4, 6, 9, 10, 13, 14, 16, 18.</p>
<p>Элемент № 7</p>	<p>Вкл./выкл режима распознавания контура текста у резака "Cut left margin".</p>
<p>Элемент № 8</p>	<p>Вкл./выкл режима распознавания контура текста у резака "Cut right margin".</p>
<p>Элемент № 11 (Т) (Top)</p>	<p>Вкл./выкл режима распознавания контура текста у резака "Cut top margin".</p>
<p>Элемент № 12 (В) (Bottom)</p>	<p>Вкл./выкл режима распознавания контура текста у резака "Cut bottom margin".</p>
<p>Обычно галки 7, 8, 11, 12 должны быть включены. Они управляют т.н. "режимом распознавания контура текста" у соответствующих резаков - верхнего, нижнего, левого и правого (при наведении мыши на ползунок каждого резака выдаётся контекстная подсказка с его названием). При включении этих галок соответствующие резаки принимают синий цвет, а при выключении - малиновый.</p> <p>При синем резаке соответствующая сторона (левая, правая, верхняя, нижняя) контура текста на скане определяется автоматически - при последующей обработке по кнопке Process!. При малиновом резаке эта сторона задаётся вручную - самим положением резака. Т.е. малиновый резак режет "там, где стоит", а синий - лишь примерно задаёт границу области распознавания контура текста, а программа потом сама при обработке определит истинную линию реза.</p> <p>Сначала существовали только синие резаки. Но из-за неизбежных погрешностей алгоритма распознавания пришлось ввести малиновые резаки - чтобы устанавливать линию реза просто вручную - в сложных случаях, когда синие резаки неправильно срабатывают - отрезают полезные элементы текста. Наиболее часто это отдельно стоящие номера страниц и т.п.</p>	

1.1.2 Вкладка Book

<div> <div>Pages Book Files</div> <div>Units Pixels</div> <div>H.Gap value 70</div> <div>V.Gap value 70</div> <div><input checked="" type="checkbox"/> vert.gap = hor.gap</div> <div>Spec. gap 0</div> <div>Page width Auto 0</div> <div>Page height Auto 0</div> <div><input checked="" type="checkbox"/> Use average width</div> <div><input type="checkbox"/> Merge pages after split</div> </div>	Элемент	Описание
	Units	Используйте всегда значение Pixels.
	H.Gap value	Размеры левого и правого полей, которые будут автоматически добавлены к контуру текста, определяемого программой (всё это происходит при обработке по кнопке Process!). Очень важный параметр, различный для разных dpi. Для 300 dpi следует задавать 70, для 400 dpi - 100, для 600 dpi - от 120 до 150.
	V.Gap value	Аналогично - здесь ширина верхнего/нижнего полей страницы.
	vert.gap = hor.gap	Вкл./выкл. режим равенства вертикальных и горизонтальных полей. Рекомендуется всегда включать.
	Page width	<p>Способ определения ширины страниц при обработке по кнопке Process!.</p> <ul style="list-style-type: none"> - Auto: Автоматическое определение средней ширины - самой обработкой по кнопке Process!. Рекомендованное значение. - Fixed: Обрезанная страница будет иметь заданные размеры, которые можно задавать вручную через предусмотренные справа поля. - None: Обрезанная страница не будет иметь полей - т.е. мы получим после обработки по кнопке Process! "голый текст" в чистом виде. <p>Примечание: После обработки по кнопке Process!, если было задано Auto, программа автоматически подставляет вычисленные средние размеры страниц в поля H.Gap value и V.Gap value, и тогда нужно обязательно сменить Auto на Fixed, чтобы при повторной переработке каких-то отдельных страниц (Process current file) их размер выдерживался равным вычисленному.</p>
	Page height	Аналогично - способ определения высоты страниц.

1.1.3 Вкладка Files

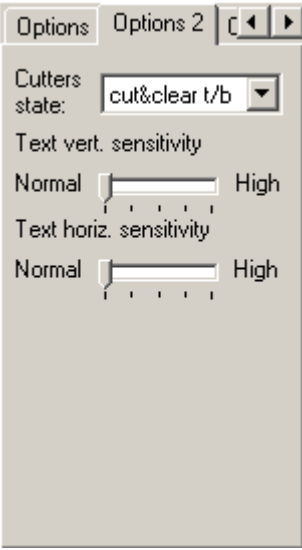
Элемент	Описание
Output dir:	Папка, в которую будут сохраняться результирующие файлы. Если папка не существует, программа попытается создать её.
Default	Использовать папку по умолчанию для результирующих файлов. Эта папка будет называться "out" и автоматически создастся внутри папки с исходными файлами.
Output prefix	Произвольный префикс, который будет добавляться к именам выходных файлов. Можно задать символ * - тогда на выходе будут сохраняться оригинальные имена файлов. Если при этом исходные файлы являются multi-page тифами, или просто разворотами, то исходное имя также будет сохранено, но с прибавлением соответствующего суффикса.
Start from/ step	Start from (1) - числовое значение в имени файла, с которого начнется нумерация выходных страниц. step (2) - шаг, используемый при нумерации выходных файлов.
Name length	Длина имени выходных файлов
Output format:	Формат выходных изображений.
DPI	Значение DPI выходных файлов.
Color	Режим цветности выходных файлов.
Input DPI	Входное dpi. Вводится только в том случае, если обрабатываемые файлы либо не содержат значение dpi, либо оно у них неправильное.

<div>Files Options Opti</div>	Элемент	Описание
<div> <div>Deskew method</div> <div>Shear</div> <div>Antialias Threshold</div> <div>Auto</div> <div>AntiShadow Threshold</div> <div>Auto</div> <div>Resample filter</div> <div>Lanczos3</div> <div>Despeckle</div> <div>Normal</div> <div> <input type="checkbox"/> Sand despeckle <input type="checkbox"/> Strong con.speckles <input checked="" type="checkbox"/> LZW for non-bw TIFFs </div> </div>	Deskew method	<p>Выбор алгоритма поворота скана при Deskew. Рекомендуемый метод - Auto(shear).</p> <p>Fast: Стандартный поворот. Именно этот режим используется в большинстве граф. программ. Это lossless-поворот, т.е. буквы не утолщаются. Но он приводит к артефактам: сдвигу одной части буквы относительно другой и появлению эффекта "обратного italic" (для сканов ≤ 300 dpi). Кстати, в BookRestorer эти артефакты налицо. Fast хорош для сканов с dpi > 500 - тогда искажения практически незаметны и его можно спокойно применять - к тому же он намного быстрее всех остальных режимов поворота. Начиная с версии 5.9, Fast больше не нужен.</p> <p>Antialias (англ. "сглаживание, устранение контурных неровностей"): Для grey/color-сканов. Имеет интерполяцию для предотвращения артефактов (например, "ступеньки" в строках текста). Слегка неравномерно уширяет буквы. Для серых/цветных Antialias полностью идентичен Interpolate, когда-то он был придуман именно для b/w чтобы нейтрализовать артефакты.</p> <p>Interpolate: Для grey/color-сканов. Это Fast с билинейной интерполяцией - для предотвращения артефактов (например, "ступеньки" в строках текста). Хорош для gray и color, для b/w ничего особого не дает. На толщину повернутого текста влияет значение параметра convert to b/w threshold (закладка Convert). Если выбрать Interpolate и задать Convert = LowLight, то уширение практически исчезает, при том, что текст поворачивается так же хорошо (т.е. не деформируется ступеньками), как и при Antialias-повороте. Начиная с версии 5.9, метод Interpolate переписан на Ассемблере, поэтому работает быстрее, чем Fast.</p> <p>Shear (англ. "сдвиг (фрагмента изображения)"): Метод для b/w-сканов.</p> <p>Auto(shear): Метод для b/w-сканов. Shear с автоматической компенсацией артефактов. Это не-интерполяционный метод. Данный метод - самый рекомендуемый. В отличие от Antialias, без побочных эффектов. Начиная с этой версии, при этом режиме grey/color-сканы поворачиваются с помощью метода Interpolate (наилучшее качество), а b/w - как и раньше, с помощью Shear-метода.</p> <p>Вообщем, следует использовать только Auto(Shear), который в зависимости от цветности скана сам выбирает нужный, наиболее оптимальный метод, т.е. Interpolate для серых-цветных, Shear с автокомпенсацией артефактов - для b/w.</p>

	Antialias threshold	Порог сглаживания. Регулирует работу режима Antialias (см. выше). Чем этот порог меньше, тем меньше утолщение, но тем больше проявление искажения. Рекомендуемое значение - 150 - тогда буквы практически не утолщаются.
	Resample filter	Фильтр при изменении разрешения (т.н. ресемплинг). Используйте всегда значение Bicubic.
	Despeckle	<p>Выбор алгоритма поворота скана при Despeckle. Рекомендуемый метод - Normal.</p> <p>Примечание: любой из алгоритмов Despeckle предназначен только для ч.б. сканов - на grey/color-сканах он не производит никакого эффекта.</p> <p>ВАЖНО: Иногда Despeckle вообще не следует применять - на бледных или низкокачественных сканах - чтобы не удалить кусочки букв или грамматические точки.</p> <p>Fine: Это "геометрический" алгоритм. Удаляются все спеклы, меньшие определённого размера (размер выбирается Кромсатором самостоятельно), и даже спеклы чуть большего размера, но имеющую определённую форму, явно указывающую на "мусорное" происхождение.</p> <p>Распознаются буквы i и j и метод не позволяет "зачищать" их точки. Особенно это актуально, когда эти символы используются как индексы в математических формулах. Такой защиты, насколько известно, нет ни в одной другой программе. Метод показал свою 100% эффективность и надежность. Это самый "чистящий" метод.</p> <p>Normal: Это классический простейший алгоритм Despeckle (фильтр по среднему с апертурой 3x3 и с пропуском белых точек).</p> <p>Fine+Normal: Интеллектуальный метод очистки. Например, он не вычищает точки над i и j.</p> <p>Safe: Метод для бледных сканов, с разорванными буквами. В этом случае в процессе обработки области, занимаемые буквами, чиститься не будут. Суть его - в предварительной кластеризации текста, т.е. выделении букв, и удалении мусора только в области, не принадлежащей букве. Метод не идеальный, но буквы страдают гораздо меньше.</p>
	Sand despeckle	Эта опция хоть как-то соединяет точки "песочно-образных" букв (т.е. состоящих из разрозненных точек).

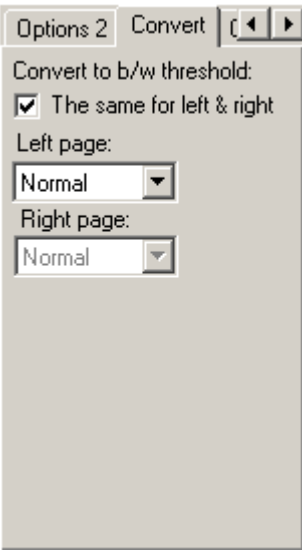
		<p>Это strong connected pixels (4-connected pixels). В противоположность 8-connected pixels. Т.е. если эту опцию включить, то пиксели, касающиеся друг друга только по диагонали, будут интерпретироваться как два разных, т. е. изолированных.</p> <p>Эта опция влияет на чистку. По умолчанию выключена, что означает: если пиксел касается другого только по диагонали, то они все равно считаются принадлежащими одному спеклу, и размер его получается больше, и вероятность его удаления понизится. Поэтому лучше не включать этот флаг.</p>
--	--	--

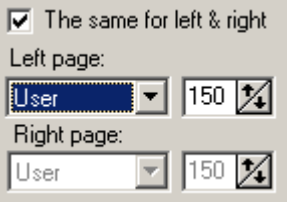
1.1.5 Вкладка Options 2

<div>Options Options 2</div>		Элемент	Описание
	Cutters state:		<p>Управление режимом очистки резаков. Значение по умолчанию - cut&clear t/b, но рекомендуемое значение - only cut. Влияет также на вид галок панели управления резаками:</p> <div data-bbox="659 837 1286 904"> <input type="checkbox"/> Xl=0 <input type="checkbox"/> Xr=601 <input type="checkbox"/> Yt=4 <input type="checkbox"/> Wl=601 <input type="checkbox"/> Wr= <input checked="" type="checkbox"/> Xi1= <input type="checkbox"/> Xi2=-16 <input type="checkbox"/> Yb=422 <input type="checkbox"/> Hl=418 <input type="checkbox"/> Hr= </div> <p>cut&clear t/b: выключён/включён и засерен/включён (только резаки "верх-низ").</p> <p>only cut: только включён/выключён.</p> <p>cut&clear all: выключён/включён и засерен/включен (все резаки).</p> <p>Засерен означает, что область за резаком будет не отрезаться при кромсании, а только очищаться. При включённом Automargins это даёт в итоге один и тот же результат. Если же юзер взял весь контроль за размерами страниц на выходе на себя (Automargins = off), то это имеет смысл.</p>
		Text vert. sensivity	<p>Вертикальная регулировка точности работы алгоритма, распознающего контур текста (т.е. алгоритм по кнопке Process!) Изменять этот параметр НЕ рекомендуется никогда.</p> <p>При автоматическом определении программой контура текста могут возникать неточности. Так, могут быть проигнорированы одинокие, далеко отстоящие от основного текста символы ("одиночные выступающие символы" (термин bolega) - ОБ-символы), напр. номера страниц, состоящие из одной цифры. Или последние 1-2 символа выступающей строки текста. Это известная и пока не решённая проблема. В результате эти символы могут быть потеряны (если их не перекроет задаваемое поле H.Gap Value или V.Gap Value).</p> <p>Рекомендуемый способ регулировки: после проработки Draft kromsate переключить в Automargins соответствующий резак с синего на малиновый и выставить его вручную сразу за ОБ-символом, который некорректно отрезается алгоритмом по кнопке Process!.</p>

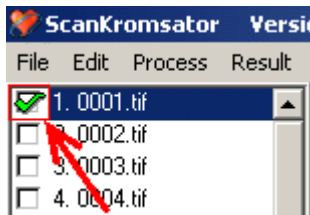




	Text hor- riz. sen- sivity	То же самое, что и строкою выше, только по горизонтали.
--	---	---

1.1.6 Вкладка Convert

	Элемент	Описание
	Convert to b/w threshold:	<p>Порог конвертации в ч./б. Рекомендуемое значение - Normal. Влияет на процедуру конвертации grey->b/w Параметр определяет порог grey-цвета, ниже которого точка становится черной, а выше - белой. Диапазон 0-255. Для удобства ряд ходовых порогов имеют имена (см. ниже). Это странично-ориентированный параметр, т.е. для каждой страницы его можно задавать различным.</p> <p>Влияет на результат, как это ни странно, при обработке цветных сканов (без перевода в b/w). Можно его поварьировать (сделать более dark, если скан бледный, и light, если наоборот).</p>
	The same for left & right	Одно и то же для левого и правого (Вкл./Выкл.) Имеет смысл только при работе со сдвоенными разворотами.

	<p>Left page:</p>	<p>Значение порога конвертации для левой страницы. Подбирается экспериментально в зависимости от степени бледности скана. Неправильно выбранный порог приводит либо к потере деталей, либо к излишнему замусорению скана.</p> <p>Normal, LowDark, MiddleDark, LowLight, MiddleLight, HighLight: обычные ходовые пороги.</p> <p>HighDark: В отличие от остальных, не является пороговым. Он адаптивный, порог высчитывает автоматически, исходя из энтропии изображения.</p> <p>Smooth: Исключение. Применяется спец.алгоритм со сглаживанием (анализирует энтропию), никак не связанный с порогом конвертации в b/w. Может сглаживать, "округлять" (слегка ужирнять) буквы (в т.ч. очень бледные части букв) и даже увеличивать их связность (убирать разрывы в буквах).</p> <p>Smooth очень хорошо работает на чрезвычайно блёклых сканах или сканах поганого качества (например, с очень низким dpi).</p> <p>Недостаток - очень чувствителен к фону (шумам), поэтому если скан сильно грязный (с неравномерным фоном), то smooth вносит очень много мусора, который в итоге так сливается с текстом, так что никакой Despeckle потом не помогает, поэтому лучше использовать его совместно с Histogram Thresholding для подавления грязи (фон слегка забледняется с помощью Histogram, и поэтому не вносит мусора при конвертации).</p> <p>Серые полутоновые рисунки при Smooth следует включать в Exclude region. После Smooth текст слегка жирнеет. Чтобы этого не происходило, можно, к примеру, на закладке Quality -> Gray enhance... -> Contrast сдвинуть ползунок Brightnes (яркость) на позицию 20 влево.</p> <p>User: Значение порога задаётся пользователем. При выборе User появляются окошки выбора значения порога:</p> 
	<p>Right page:</p>	<p>Значение порога конвертации для правой страницы</p> <p>Аналогично - для правой страницы.</p>

1.2.1 Элементы управления списка файлов

Элемент	Описание
	<p>Жирная зелёная галочка (другое название - "marked"-файл). Её смысл:</p> <ol style="list-style-type: none"> 1. Пометка на обработку (по кнопке Process!) данного файла - это сейчас основной смысл. Только те файлы, напротив которых стоит эта галка, будут обработаны после нажатия кнопки Process!. 2. Запоминание (для последующей обработки по кнопке Process!) текущих изменённых индивидуальных настроек файла - уже устаревший смысл (из-за кнопки "Auto-accept options is on/off"). <p>Когда-то после изменения опций (например, подвинули резак) необходимо было "перещёлкнуть" эту галочку - т.е. снять её и вновь поставить - иначе изменённые опции не запоминались, но в последних версиях Кромсатора появилась специальная (для этой проблемы) кнопка "Auto-accept options is on/off"  - когда она включена, "перещёлкивать" галочки (после изменения опций) уже не надо. Эта кнопка теперь включена по умолчанию.</p>
	Щёлкнуть мышкой на любой (обычно последний) файл в списке, нажать эту кнопку - автоматически проставятся жирные зелёные галочки на всех файлах в списке, предшествующих ему.
	Щёлкнуть мышкой на любой файл в списке, нажать эту кнопку - на нём проставится жирная зелёная галочка.
	Щёлкнуть мышкой на любой (обычно (пред)последний) файл в списке, нажать эту кнопку - автоматически проставятся жирные зелёные галочки на всех alternate ("чередующихся", т.е. одинаковой чётности) файлах в списке, предшествующих ему. Т.е. если стоим на чётном файле - пометятся все предыдущие ему чётные файлы, если стоим на нечётном - то пометятся все нечётные.

1.2.2 Виды выделения файлов в списке:

Название	Описание	Способ выделения	Групповое изменение
Marked	С жирной зелёной галочкой	Левая кнопка мыши	Клавиша Ctrl
Selected	Выделенные красным цветом.	Клавиша Пробел (или Insert) (при активном списке файлов). Чтобы выделить группу, нужно пробелом отметить первый файл в группе, затем стать на последний файл и нажать Shift + Пробел .	Клавиша Shift
Bold selected	Выделенные чёрным жирным.	Клавиша Ctrl + Пробел .	

Из таблицы видно, что, чтобы поменять опцию у группы файлов, то нужно (при изменении опции):

- Для marked-файлов держать нажатым **Ctrl**.

- Для selected-файлов держать нажатым **Shift**.

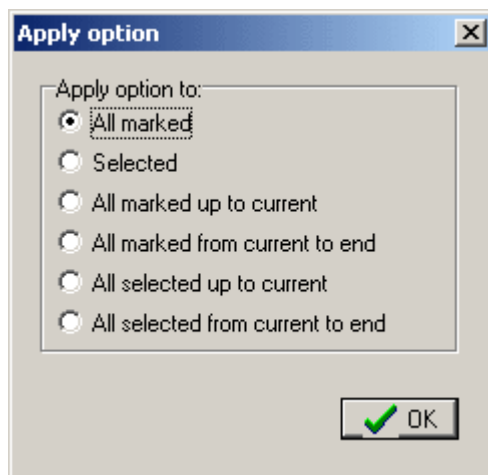
Выделение **Marked** (как это указано чуть выше) используется, чтобы пометить файлы, которые будут обрабатываться обработкой по кнопке **Process!**

Выделение **Selected** - это пользовательское выделение, используемое в случаях, когда нужно выделить группу файлов в списке и проделать над ней некую операцию.

Выделение **Bold selected** используется обработкой **Draft kromsate** - для того, чтобы обозначить те файлы, где хотя бы один из резаков (после **Draft kromsate**) установился **малиновым**, а не **синим**.

1.2.3 Групповое изменение опций

Любую опцию можно установить/сбросить/изменить у любой группы файлов. Для этого нужно при изменении опции держать нажатым **Alt**. Кромсатор предложит на выбор, к какой группе применить изменение (в любом случае, выдав сначала такое предупреждение):



1.2.4 Управление списком файлов

1.2.4.1 Перемещение файлов:

а. **Единственный файл** - нажимаем на него левой кнопкой мыши и перетаскиваем на другую позицию в списке.

б. **Группа файлов:**

- Команда **Move selected**: перемещает все выделенные (selected) файлы (выделение может быть не обязательно сплошным, а, например, через один) в определённую позицию в списке файлов. Например, можно выделить все чётные страницы и переместить их в конец списка, где они расположатся уже последовательно.

- Команда **Shift selected**: все выделенные файлы сдвигаются на заданное количество строк вверх или вниз, сохраняя при этом расстояние между друг другом.

- Команда **Paste interliev** ("вставлять файлы через один"): вставляет в список (скопированные в буфер файлы) через один, т.е. на место только чётных или только нечётных.

Если изменение порядка (как для единственного файла, так и для группы) выполняется уже после обработки по кнопке **Process!** (т.е. когда уже получены выходные файлы), то программа будет автоматически переименовывать и выходные файлы так, чтобы сквозная нумерация на выходе не нарушалась.

1.2.4.2 Навигация по списку файлов:

ВАЖНО: Для перемещения по списку файлов (в главном окне, а также в окне просмотра результатов), можно использовать "горячие клавиши": "q" или "[" (предыдущий) и "]" или "w" (следующий).

В списке исходных файлов нажатие стрелок, Home, End при нажатом Shift перемещают курсор

на следующий selected, предыдущий selected, первый selected, последний selected соответственно.

1.2.4.3 Многостраничные TIF-файлы:

Если открывать много TIF-файлов скопом, программа считает их одностраничными.

Если к первому открытому одностраничному TIF-файлу добавить многостраничные, она их тоже считает одностраничными.

Если первый многостраничный - то все нормально, т.е. программа распознаёт корректно все последующие добавленные файлы, какие бы (одно- или много- страничные) они ни были.

Для того, чтобы открыть несколько (не один) TIF-файлов как многостраничные, нужно при выборе команды меню (или при щелчке на кнопке открытия файлов) держать нажатым Shift. Такое явное указание было введено из-за чрезмерной длительности автоматического определения "многостраничности" открываемых TIF-файлов.

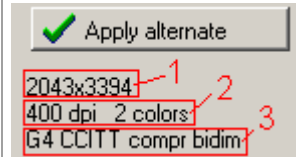
1.2.4.4 Разное:

Начиная с 4-й версии, Кромсатор поддерживает полуавтоматическое добавление файлов в задание. Например, вы сканируете книгу, и файлы попадают в папку MyScan. Вы загружаете сколько есть файлов в задание. В дальнейшем, если в этой папке будут появляться от сканера новые изображения, можно просто нажимать в кромсаторе "Н" (раньше для этого нужно было вызывать пункт меню add files) и они добавятся в конец задания (при условии, что сканирующая программа нумерует их).

В уже существующее задание можно добавлять новые файлы (с возможностью пересортировки, если необходимо), при этом допускаются одинаковые имена файлов (при условии, что папки у них все-таки разные).

Теперь в задании можно использовать любую смесь форматов файлов. (Ранее к multi-page-тифам, например, нельзя было добавить bmp или jpg).

1.2.5 Панель свойств загруженных файлов



	Элемент	Описание
	1.	<горизонтальный размер изображения в пикселях> x <вертикальный размер изображения в пикселях>
	2.	<разрешение изображения в DPI> <количество цветов изображения>
	3.	<Вид сжатия изображения>











При запуске Кромсатора эта панель ВСЕГДА не видна. Чтобы её увидеть, надо "растянуть вверх" панель списка файлов.

1.2.6 Кнопки панели управления

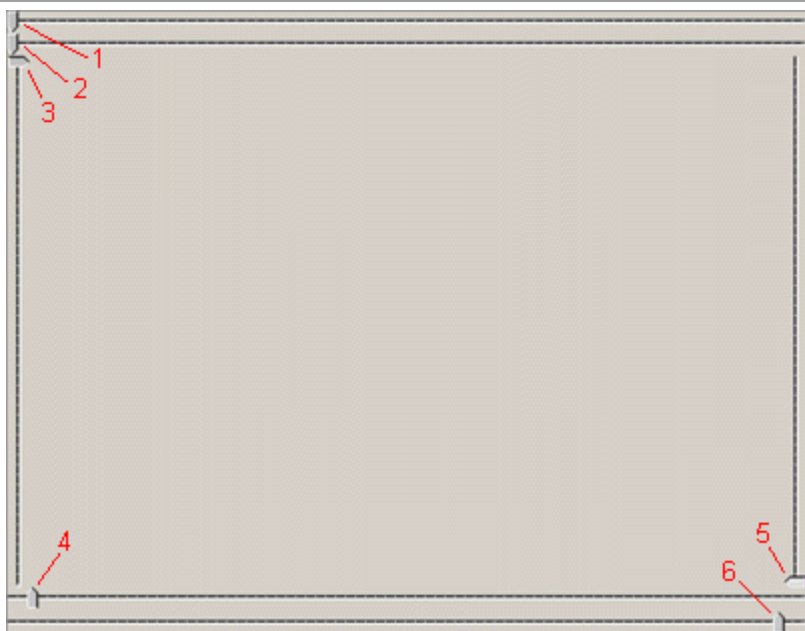


- Общий вид основной панели управления.

Элемент	Описание
	Load image files - загрузить графические файлы.
	Save current task - сохранить текущее задание.

	Fit image - растянуть изображение по ширине окна.
	Draft kromsate - т.е. запустить процедуру автоматической предварительной расстановки резакoв - по всем загруженным сырым сканам.
	Запустить обработку marked-файлов (<i>т.е. помеченных жирными зелёными галочками</i>).
	Process current file - обработать текущий файл.
	Process left page - обработать текущую левую страницу. Имеет смысл только при обработке сдвоенных разворотов.
	Process right page - аналогично для текущей правой страницы.
 Выкл.)  (Вкл.)	Auto-accept changings option - Автопринятие изменения опций Включена по умолчанию. Рекомендуется иметь эту кнопку всегда включённой . Выключать её имеет смысл только при использовании кнопки Preview на вкладке Quality - т.к. при смене dpi кромсатор не корректирует положения резакoв.
	Image zoom - масштаб изображения.
	Rotate image - поворот изображения. Можно держать Ctrl или Alt при изменении значения в этом комбобоксе. Ctrl - повернет для всех, Alt - спросит, для каких.

1.2.7 Схема расположения резакoв

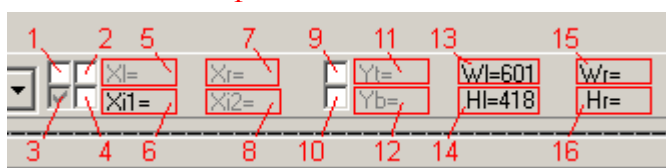
 <i>(Изображение уменьшено)</i>	Элемент	Описание
	1. - (Set internal margin)	?
	2. - (Set internal margin or split line)	Резак для разрезания сдвоенных разворотов
	3. - (Cut top margin)	Верхний резак
	4. - (Cut left margin)	Левый резак
	5. - (Cut bottom margin)	Нижний резак
	6. - (Cut right margin)	Правый резак

1.2.8 Панель управления резакoв



- Общий вид панели управления резаками.

При запуске Кромсатора эта панель ВСЕГДА не видна полностью. Чтобы её увидеть, надо "за-
узить" панель списка файлов.



Элемент	Описание
1. (Cut left margin)	Вкл./выкл. левый резак
2. (Cut right margin)	Вкл./выкл. правый резак
3.	? (Возможно, это ранее было управление резаком для разрезания сдвоенных разворотов, который теперь не отключается, а просто его действенность включает-ся/выключается галкой Split)
4. (Cut internal2 margin)	Вкл./выкл. резак Set internal margin (?)
5. (left cutter position)	Координата левого резака
6. (first internal cutter position)	Координата резака для разрезания сдвоенных разворотов
7. (right cutter position)	Координата правого резака
8. (second internal cutter position)	Координата резака Set internal margin (?)
9. (Cut/clear top margin)	Вкл./выкл. верхний резак
10. (Cut/clear bottom margin)	Вкл./выкл. нижний резак
11. (top cutter position)	Координата верхнего резака
12. (bottom cutter position)	Координата нижнего резака
13. (left page width)	Расстояние между левым и правым резаком (при выкл. Split) или между левым резаком и резаком для разрезания сдвоенных разворотов (при вкл. Split)
14. (left page height)	Расстояние между верхним и нижним резаком
15. (right page width)	Расстояние между резаком для разрезания сдвоенных разворотов и правым резаком (при вкл. Split)
16. (right page height)	Расстояние между верхним и нижним резаком

1.2.9 Управление резаками

Чтобы двигать сразу два резака синхронно (например, верхний и нижний), нужно держать на-
жатым **Ctrl**.

Резаки можно наклонять (чтобы лучше отсекал грязь и полосы на перекошенных сканах). Для этого нужно двигать ползунок резака правой клавишей мыши, либо левой, но с нажатым **Shift**. Двой-
ной щелчок на ползунке выпрямляет резак.

Если нужно убрать тень и мусор от разворота, но при этом не разрезать разворот на две части,
то нужно убрать опцию **Split**, но оставить оба внутренних резака. В этом случае область между ними

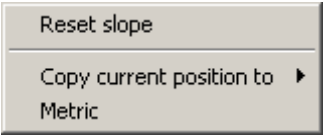
будет просто очищаться.

Резаки синхронизируются с картинкой при изменении масштаба, повороте и отражении.

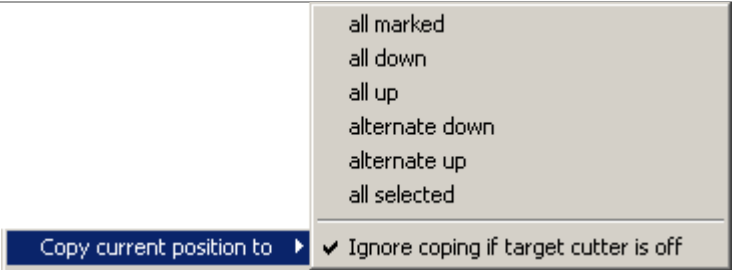
Можно на панели управления резаками в контекстном меню элемента, в котором выводятся координаты резаков ($X_{i1}=...$, $X_{i2}=...$, и т.п.) выбрать команду **Reset cutter position** - соответствующий резак подойдет к соответствующему краю страницы и станет виден.

Отключение резаков скопом не предусмотрено (чревато). Но выход есть - нужно вывести резаки за пределы картинки.

1.2.10 Контекстное меню ПОЛЗУНКА любого резака



Элемент	Описание
Reset slope	Убрать наклон резака (т.е. "выпрямить" его).
Copy current position to >	Распространить позицию данного резака на ... <i>Подробнее - см. ниже таблицу</i>



Copy current position to	Распространить позицию данного резака на ...
Пункт меню	Назначение
all marked	все файлы, отмеченные жирной зелёной галочкой.
all down	все нижележащие файлы
all up	все вышележащие файлы
alternate down	все нижележащие файлы той же чётности
all selected	все выбранные красным файлы

1.2.11 Пункты контекстного меню (главного окна)

Редактировать теперь можно и исходное изображение (в отличие от старых версий), хотя операций и меньше, чем для выходного. Поддерживается редактирование для многостраничных тифов.

Исходный файл можно чистить (Mouse-Up mode -> Clear selection) и сохранять (Auto save).

Для исходного изображения имеется операция быстрого удаления любой тонкой сплошной длинной чёрной линии (хорошо удалять подчёркивания ручкой). Пока только для исходного b/w-скана. Нажав Ctrl + Shift, щёлкните по любой черной точке контура. Но учтите, что если линия касается буквы, то и буква удалится.

При редактировании изображений (как исходного, так и выходного) поддерживаются не только прямоугольные выделения, но и эллиптические и произвольной формы, нарисованные мышкой.

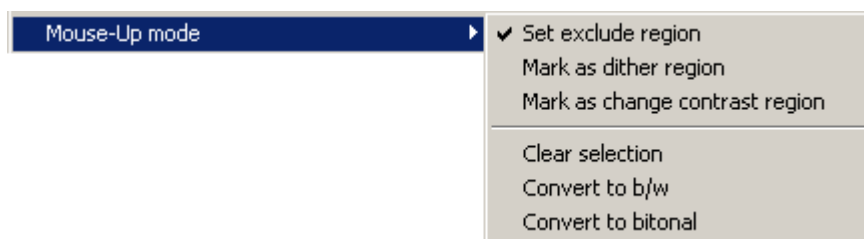
Zoom rectangle	
Show selection	
Crop selection	
Invert selection	
Deselect	
Save image	
Undo	Ctrl+Z
Exclude region	
Exclude and Mark as dither region	
Mark as bitonal region	
Mark as change contrast region	
Mark as floodfill region	
Delete all regions	
Delete selected region	
✓ Ruler	
Units	►
No Zoom Filter	
● Linear Zoom Filter	
Mouse-Up mode	►
Selection form	►
✓ Auto save	
✓ Do not show save dialog	
Convert to bitonal	Ctrl+Y
Convert to b/w	Ctrl+U
Convert to grey	

Пункт меню	Назначение
Zoom rectangle	Увеличить выделенный фрагмент во весь экран
Show selection	По этой команде создаётся небольшое окошко (их можно создавать сколько угодно), в котором будет отображаться выделенная область. В нём можно таким образом "хранить" разные варианты обработки и визуально сравнивать их.
Crop selection	Вырезать выделенный фрагмент в отдельное новое изображение
Invert selection	<p>Инвертировать выделение - поменять местами выделенную и невыделенную область - т.е. выделенная область становится невыделенной, а пространство скана за пределами этой области становится выделенным.</p> <p>Так можно, например, выделить на изображении области вне рисунков: выделяем картинки (множественное выделение поддерживается, держать shift), и выбираем команду invert selection - тогда выделение инвертируется, т.е. выделяется все, кроме картинок.</p>
Deselect	Снять выделение
Save image	Сохранить изображение
Undo	Откат изменения. Глубина отката - неограниченна

Exclude region	<p>Зона исключения. Допускается до 20 зон на страницу. Исключает действие (внутри зоны) следующих операций:</p> <p>Despeckle - для защиты полутоновых или мелкозернистых рисунков и изображений. Это первоочередная задача Exclude region.</p> <p>Deskew (Auto Shear, Interpolate, Antialias) - зона не будет участвовать в определении skew-угла (например, именно изображения приводят к неправильной детекции угла наклона). Зона будет поворачиваться без использования интерполяции, чтобы не внести дополнительный муар.</p> <p>Выравнивание освещенности - если на скане имеется полутоновые иллюстрации, то при они могут сильно пострадать.</p> <p>Все опции Gray image enhance, за исключением Histogram.</p> <p>Метод удаления фона Convert -> Convert to b/w threshold: -> Smooth.</p>
Exclude and Mark as dither region	Данная зона при обработке будет переводиться в b/w по специальному алгоритму, предназначенному для правильной передачи полутоновых или цветных изображений.
Mark as bitonal region	Bitonal-конвертирование зон. Просто спасение для точной передачи многотонных иллюстраций.
Mark as change contrast region	Зона изменения контраста. Значение контраста берется из окошка Contrast, закладка Contrast, при условии, что применение изменения контраста глобально (ко всему изображению) отключено. Т.е. если на закладке Contrast стоит галка на Enabled, то contrast-zones игнорируются.
Mark as floodfill region	FloodFill-зоны работают только для b/w. Заливается белым цветом чёрная область, начиная от центра зоны. Если окажется, что в центре зоны - белый пиксел, то кромсатор будет искать черный пиксел в пределах зоны, начиная от центра и двигаясь по спирали. Если в пределах зоны не будет чёрных пикселей, то ничего выполняться не будет. Заливка чёрной области выполняется применительно ко всей области, т.е. не ограничивается размером зоны. Т.е. зона нужна только для того, чтобы задать своим центром начальную координату, с которой будет выполняться заливка.
Delete all regions	Удалить все выделенные участки
Delete selected region	Удалить выделенный участок

✓ Ruler	<p>С его помощью можно мерить не только расстояния, но и одновременно угол наклона.</p> <p>Запись измеренного угла: включить режим измерения углов (меню Ruler), щёлкнуть на первой точке линии измерения и, не отпуская кнопку мышки, подвести её курсор к другой точке этой же линии. За мышкой будет следовать "резиновая" линия, рядом с которой будет выводиться угол наклона линии. Если затем отпустить кнопку мыши и при этом держать нажатым Ctrl (или Shift), то измеренный угол занесётся в поле Fixed skew angle (вкладка Pages кнопка special...) с тем же знаком, что и знак, выдаваемый Ruler. На этот жёстко заданный угол кромсатор и будет поворачивать скан при обработке. Кстати, все углы в кромсаторе трактуются одинаково, т.е. + означает поворот против часовой стрелки.</p>
Units >	Единицы измерения
No Zoom Filter	Выключение фильтра отображения изображения. Это удобно при работе с цветными сканами - чтобы меньше тормозило.
• Linear Zoom Filter	Использовать линейный фильтр отображения изображения.
Mouse-Up mode >	Выбор смысла прямоугольного выделения мышью <i>см. далее подробнее</i>
Selection form >	Выбор формы инструмента выделения фрагмента <i>см. далее подробнее</i>
✓ AutoSave	Автосохранение текущего файла при переходе к следующему. Удобно при редактировании каждой страницы.
✓ Do not show save dialog	?
Convert to bitonal Ctrl+Y	<p>Преобразовать в битональное изображение. То есть применить dithering (или растривание) к серому изображению - т.к. его прямое преобразование его в ч.б. испортит его. Dithering (дизеринг) - это имитация серого (8-битного) изображения мелкими чёрно-белыми точками, построенными на скане так, что dithering-изображение издалека очень похоже на исходное серое - однако, является уже 1-битным - т.е. чёрно-белым.</p> <p>Алгоритм работы с Convert to bitonal: страницы с худ.картинками сканируем в greу, преобразуем иллюстрации (т.е. отдельные фрагменты страницы) в bitonal, нажимаем Preview (при этом т.к. выходной формат задан b/w, получаем исходный файл уже как b/w), сохраняем его и в итоге на входе программа уже имеет b/w-файл, но с иллюстрациями, почти не попорченными прямым конвертированием в b/w. Надо еще чтобы use error dithering был включен, там же, в меню.</p>
Convert to b/w Ctrl+U	Преобразовать в ч.б. При этом порог конвертации берется из текущего значения Convert to b/w threshold на закладке Convert). При этом формат файла не изменяется, просто заданный кусок преобразуется в b/w и возвращается на место, т.е. файл как был, так и остаётся 8-битным (серым) - а преобразованная картинка построена теперь уже только лишь

	из чёрного и белого цветов.
Convert to grey	Конвертирует выделенную область на цветном изображении в grey. Команда не является опцией обработки.



Mouse-Up mode	<i>Выбор смысла прямоугольного выделения мышью</i>
Пункт меню	Назначение
✓ Set exclude region	Установить зону исключения
Mark as dither region	Пометить как dithering-зону
Mark as change contrast region	Пометить как зону изменения контраста
Clear selection	Очистить выделение. Это очень удобный режим для ручной чистки от мусора. Выделяем грязь прямоугольником мышью, отпускаем левую кнопку мыши - и весь выделенный прямоугольник очищается.
Convert to b/w	Конвертировать в чёрно-белое
Convert to bitonal	Конвертировать в dithering

1.3 Пункты меню

ПРИМЕЧАНИЕ: В этом разделе будут рассмотрены только наиболее важные и интересные пункты меню. О смысле и значении остальных пунктов меню (которые мы не будем рассматривать) легко догадаться, исходя из прочих разделов данного Пособия.

1.3.1 Пункт меню File

В программе существует такое понятие, как "Задание". Задание - это информация о загруженных в программу файлах, расставленных по ним резакам, выставленным опциям и т.д. Задание можно сохранять и открывать заново. Использование задания позволяет прерывать и снова возобновлять обработку сканов

Open task	Открыть задание
Save task	Сохранить задание
Close task	Закрыть задание
Task options settings	Открыть окно выбора пакета опций

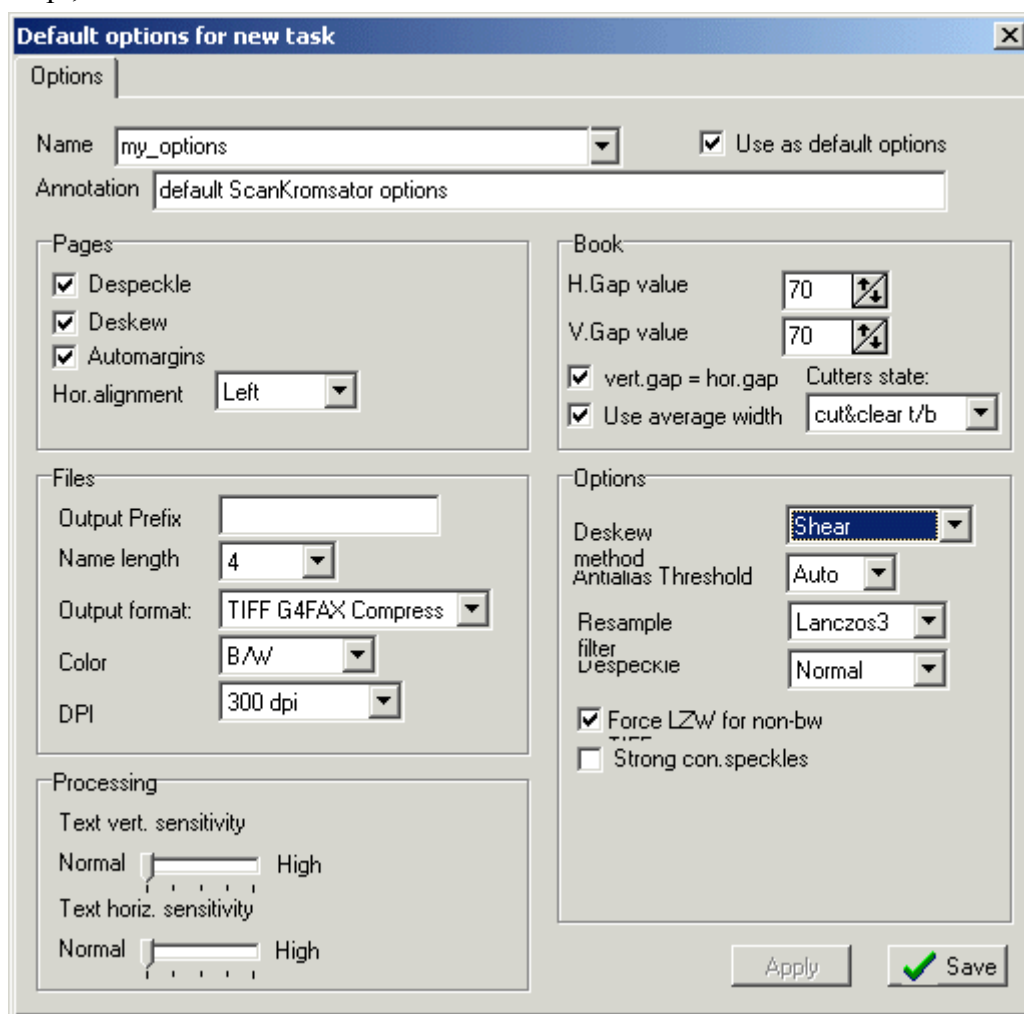
1.3.1.1 Подпункт Task option settings

Опции программы, изменённые пользователем, можно сохранять в виде набора, и потом использовать повторно.

Для этого в окне File->Task options settings задайте нужный набор опций, присвойте этому на-

бору любое имя и сохраните его (кнопка Save). Можно создать сколько угодно таких наборов default-опций - со своими именами. Один из таких наборов будет набором по-умолчанию (поставите в нём галку на Use as default options) - он будет автоматически загружаться при открытии Кромсатора.

Перед созданием нового задания можно загружать любой набор - тогда эти опции восстановятся. После этого можно загружать новые изображения. Например, можно сделать отдельные наборы для 300dpi, 600dpi, и т.д.



1.3.2 Пункт меню Edit

Clear all options	Очистить все опции. Эта команда сбрасывает все опции программы в нулевые значения (причём, они не совпадают со значениями по-умолчанию). Очень удобно для нестандартных обработок.
--------------------------	---

1.3.3 Пункт меню Result

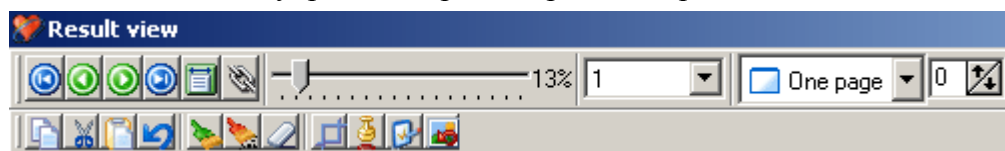
Result	
Пункт меню	Назначение
View all	<p>Открывает редактор постобработки, если в Кромсаторе имеются файлы, прошедшие обработку, то есть выполняются условия:</p> <ol style="list-style-type: none"> 1. В списке файлов есть хотя бы один marked-файл. 2. Существует папка для выходных файлов. 3. В этой папке есть хотя бы один файл, соответствующий какому-либо marked-файлу (он и загрузится в редактор постобработки). <p>Информацию о файлах, прошедших обработку, можно сохранить в файл задания (меню File -> Save task) и затем загрузить её обратно в про-</p>

	грамму (меню File -> Open task). После этого можно опять открыть редактор постобработки (выбрав View all) и продолжить там обработку выходных файлов.
Show source files	Возможность в окне просмотра результата редактировать исходные файлы.

2 Интерфейс редактора постобработки







2.1 Общие элементы

Общий вид панели управления редактора постобработки:



Обе панельки инструментов - плавающие, их можно пристыковать к бокам или превратить в окошки и закрыть (Right click - Hide) - чтобы увеличить "эффективный" размер экрана. Для того, чтобы опять включить их, надо выбрать в контекстном меню Options... -> Show navigate command's panel и/или Show edit command's panel.

Элемент	Описание
- First page	Перейти к первой странице
- Previous page	Перейти к предыдущей странице
- Next page	Перейти к следующей странице
- Last page	Перейти к последней странице
- Fit	Сделать масштаб "по ширине страницы"
- Sync original scan with current image	Если включить опцию "Sync", то при навигации по выходным файлам будет производиться автоматическая смена изображений и в окне с исходными файлами.
- Image zoom	Регулировка масштаба
- Page number	Переход к странице с нужным номером
- Panels layout	Режим отображения страниц
- Shift image increment value	- Можно двигать изображение клавишами "Alt + стрелка". Здесь задаётся шаг смещения.
- Copy	Копировать
- Cut	Вырезать
- Paste	Вклеить
- Undo	Отменить
- Auto-clear mode	Режим автоочистки

 Auto-despeckle mode	Режим автоудаления соринок
 Eraser	Ластик
 Crop	Вырезающий прямоугольник
 Show stamp	Показать штамп
 Clear options	Показать диалог с опциями очистки
 Rotate selection	Показать диалог вращения выделенного фрагмента

Чтобы пометить файл (как selected, т.е. красным) в списке файлов основного окна, находясь при этом в окне редактора постобработки **Result view**, нужно нажать F2 или F2 и F3 (для левого и правого файла, в случае, если в окне просматриваются сразу два выходных файла). Это бывает необходимо, чтобы отметить неправильно обработанный файл. Пометив таким образом все (или один) такие файлы, можно вернуться в главное окно Кромсатора, подправить помеченные файлы и дать команду process selected.

Во окне **Result view** нажатие Backspace перемещает на предыдущую просмотренную страницу (например, смотрели 10-ую, потом 15-ую - нажали Backspace и вернулись к 10-ой, а не к 14-ой). Ctrl + Backspace - на страницу, с которой началось просматривание в окне **Result view**.

Форму выделения можно запомнить в буфер обмена (Ctrl-Insert), потом восстановить - на этом же, или на другом файле (Shift-Insert).

Может ли Кромсатор делать despeckle для белых точек на чёрном фоне? Да, но только в окне редактора постобработки **Result view** (т.е. в основном окне программы не может).

Для листания страниц используются следующие "горячие клавиши": вперед - w и], назад - q и [. Листает также нажатие средней кнопки мыши. В последнем случае направление листания можно задать, выбрав соответствующую опцию в контекстном меню Options... -> Middle button click -> (Forward, Backward, None).

Ctrl-Shift-щелчок мыши на b/w работает как floodfill, на grey/color - как magic wand.

Hotkey: в случае, если zoom таков, что изображение не вмещается целиком в окно просмотра, "A" или ":" делают видимым верхнюю часть изображения, а "S" или "'''' - нижнюю, т.е. быстрый скролл вверх-вниз.

В версии 5.6A почти все операции редактирования доступны во всех режимах просмотра (одиночная страница/две страницы/сравнение с оригиналом). Активное окно подсвечивается.

Некоторые возможности редактора постобработки:

Ластик, стирающий одним из трех цветов: белым, чёрным, или средним по заданной области (это для цветных и серых).

На страницу можно поместить штамп (с поддержкой прозрачности), двигать его мышкой, изменять размеры, и затем слить с изображением.

Magic-wand, усреднение цвета и т.д.

Возможность применить despeckle не ко всему тифу, а к выделенному участку. Причем имеется и white-despeckle, когда убираются не черные точки, а наоборот - белые на чёрном фоне.

При редактировании изображений (как исходного, так и выходного) поддерживаются не только прямоугольные выделения, но и эллиптические и произвольной формы, нарисованные мышкой.

В выходном изображении поддерживаются Copy, Cut, Paste, Paste с прозрачностью.

В clipboard можно Copy не только из СканКромсатора, но и из любых графических программ.

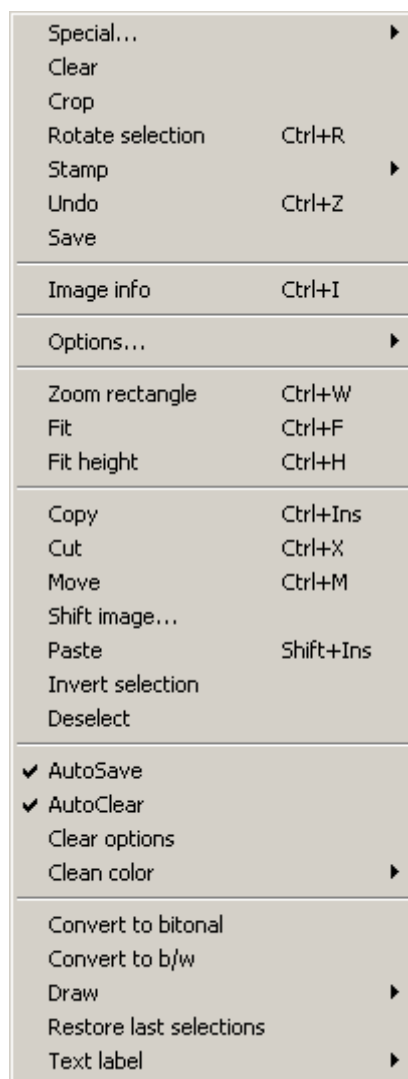
После Paste фрагмент можно двигать, растягивать и только потом merge.

Поворот выделенной области.

Масштаб клавиатурой через (+,-).

2.2 Пункты контекстного меню

2.2.1 Общее контекстное меню



Пункт меню	Назначение
Special... >	Специальные возможности
Clear	Очистить выделение
Crop	Вырезать выделение в отдельный файл
Rotate selection Ctrl+R	Поворот выделения. Rotate selection нельзя пользоваться в режиме показа Two pages. Просто в этом случае не всегда понятно, к какому из двух изображений применять команду. Сейчас в принципе это можно определить по тому, где вызвано контекстное меню, но когда будет общее меню и плавающая панель инструментов - тогда проблема.
Stamp >	Работа со штампом
Undo Ctrl+Z	Безлимитный
Save	Сохранить
Image info Ctrl+I	Информация об изображении

Options... >	Подменю с опциями
Zoom rectangle Ctrl+W	Увеличить выделенный прямоугольник во всё окно
Fit Ctrl+F	Сделать масштаб по ширине окна
Fit height Ctrl+H	Сделать масштаб по высоте скана окна
Copy Ctrl+Ins	В clipboard можно Copy не только из кромсатора, но и из любых графических программ
Cut Ctrl+X	Вырезать
Move Ctrl+M	Двигать
Shift image	Сместить изображение
Paste Shift+Ins	После Paste фрагмент можно двигать, растягивать и только потом merge. Выделяем, Copy. Paste - участок появится поверху страницы и он ещё не слит с ней, т.е. его можно таскать или изменять размеры. Если же ещё нажать клавишу T (разницу почувствуете сразу же после нажатия), то последующий merge будет выполнен с прозрачностью, т.е. белый цвет будет прозрачным. Для того, чтобы окончательно слить участок с изображением, выбираете команду merge в контекстном меню или щелкаете в любом месте за пределами участка. В новой версии добавлена ещё возможность сделать прямоугольнички по краям участка (grips) однопиксельными, а то при копировании малых участков эти грипы все заслоняли.
Invert selection	Обратить выделение
Deselect	Убрать выделение
✓ AutoSave	Автоматическое сохранение данного файла при переходе к последующему.
✓ AutoClear	После взвода AutoClear выделения мышкой будут чиститься автоматически, в момент отпускания клавиши мыши.
Clear options	Опции очистки
Clean color >	Цвет, которым будет закрашиваться очищаемая область
Convert to bitonal	Конвертировать в dithering
Convert to b/w	Конвертировать в чёрно-белое
Draw >	Вызов подменю рисования

Restore last selections	Начиная с версии 5.05, в Кромсаторе есть одна хитрая возможность - предположим, что вы почистили (в режиме autoclear) изображение и вдруг обнаружили, что по каким-то причинам страницу нужно пере-обработать (иногда такое бывает: обнаруживаешь, что часть страницы была блеклая и despeckle эту часть повредил). Тогда заключаете эту часть страницы в exclude-зону, которую despeckle трогать не будет, переобрабатываете страницу, и чтобы не чистить страницу по новой, выбираете команду restore last selections и нажимаете Delete. В результате команда отсеlectирует все зоны, которые были зачищены при предыдущей обработке, и по нажатию del очистит их. Если же при ручной чистке использовалось despeckle-removing (с помощью Ctrl-Mouse Up), то такие зоны тоже передеспеклятся (в таком же режиме) автоматом, а на странице выделятся желтой подсветкой, аналогично и для случая, если имело место despeckle белых пятен на чёрном фоне (Ctl-Shift-MouseUp), только подсветка будет зеленой.
Text label >	Работа с текстовой пометкой

2.2.2 Очистка сканов от грязи при помощи мыши

Любой из возможных алгоритмов Despeckle не в состоянии полностью и автоматически очистить скан от соринок, грязи, мусора, и т.п. Поэтому иногда следует вручную почистить сканы после Despeckle (а иногда и вместо Despeckle - если сканы бледные, плохо пропечатанные - и Despeckle может их попортить).

Настройка:

включаем в контекстном меню **AutoSave** (чтобы не надоедал сообщениями о необходимости сохранять изменённый файл).

включаем в контекстном меню **AutoClear** (теперь после выделения в момент отпускания мышки будет производиться чистка/деспеклирование, какие - см. дальше).

в окне **Clear options** (см. чуть ниже) задаем **Speckle size (fine)** = 20.

Чистка:

а. Простая очистка:

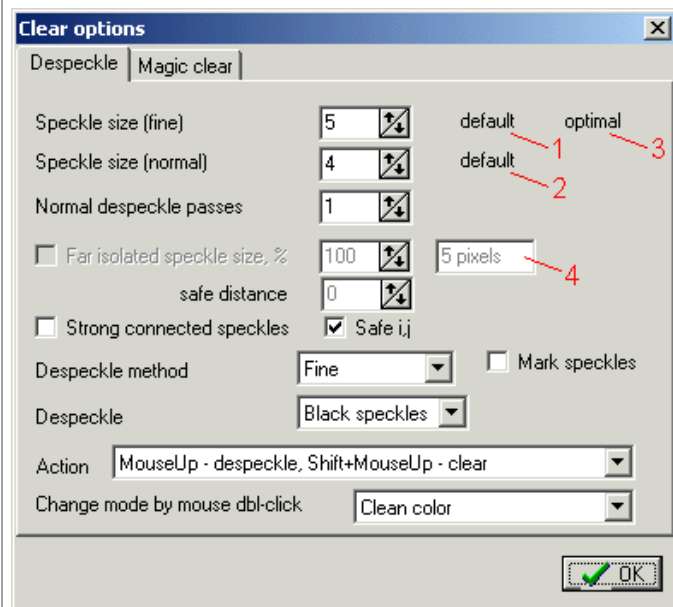
Выделяем фрагмент, в момент конца выделения, когда кнопка мышки отпускается, выделенный фрагмент полностью очищается. Цвет, которым заливается, зависит от опции Clean color (white/black). Цвет можно также быстро изменить просто двойным щелчком мышки на изображении.

б. Despeckle-очистка: (рекомендуемая)

Если при этом держать нажатым Shift, то будет не очистка, а despeckle (обращаю внимание, что despeckle применяется не ко всему изображению, а только к выделенной области). Следует увеличить (см. ниже окно **Clear options**) допустимый Speckle size (рекомендуемое значение - **Speckle size (fine)** = 20), и тогда чистка будет намного лучше, чем просто Despeckle при обработке (по кнопке Process!), но нужно следить, чтобы выделение **не захватывало знаков препинания**. Режим хорош тем, что не нужно точно позиционировать выделение на мусор, можно захватывать и текст, он в большинстве случаев не пострадает. Таким образом, Despeckle-очистка гораздо менее трудоёмка, чем простая очистка, только самые большие кусочки мусора надо вычищать всё-таки простой очисткой (т.е. через Shift при Despeckle-очистке).

Если при этом держать нажатым и Shift, и ещё Ctrl, то despeckle будет работать не для чёрного мусора, а для белого, т.е. закрашивать мелкие белые пятна на чёрном фоне. Если после этого остаются

ся пятна, которые больше, чем заданный Speckle size, то делаем двойной щелчок (т.е. переключаемся на заливку чёрным), выделяем пятно, отпускаем мышку - пятно закрасилось. Снова двойной щелчок, чтобы вернуться к белой заливке. При despeckle белых пятен нужно быть осторожным с текстом - внутренние замкнутые полости ряда букв часто имеют малые размеры и могут при этом легко закраситься чёрным!



Можно сделать немного по-другому: с самого начала в окне **Clear options** установить **Change mode by mouse dbl-click = Clear/despeckle** (чтобы двойным кликом быстро переключаться от чистки полной до чистки по типу despeckle).

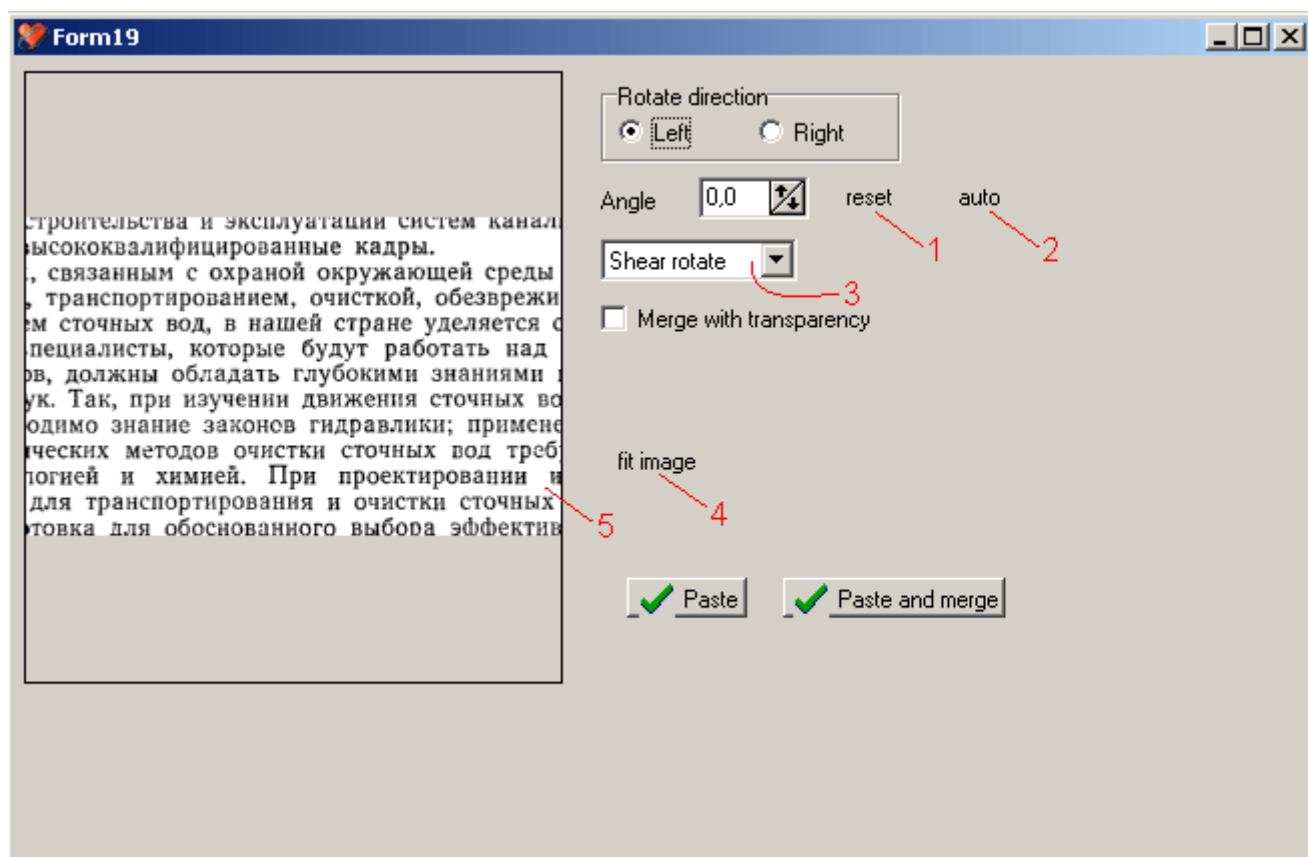
Вообще поведение клавиш (Shift, Ctrl) можно изменять - в окне **Clear Options**. Т.е. можно задать, чтобы при простом отпускании мыши был despeckle, а с Shift - чистка. Кому как нравится. По умолчанию используется **Clear Options -> Despeckle method -> Fine**. Можно дополнительно указать использовать **Fine+Normal**. Но размеры допустимых спеклов для **Normal** (они задаются отдельно, для **Fine** - одни (поле **Speckle size (fine)**), для **Normal** - другие (поле **Speckle size (normal)**) лучше не увеличивать. Т.к. **Normal** не просто чистит мусор, но и убирает лохматинки с букв.

Элемент	Описание
Speckle size (fine)	Максимальный размер соринки, которые будут удаляться выделением мышью
Action	MouseUp - clear, Shift+MouseUp - despeckle MouseUp - despeckle, Shift+MouseUp - clear MouseUp - clear, CapsLock=On - MouseUp - despeckle Выбор режима очистки мышью

Там ещё есть несколько операций с magic color. Это для цветных сканов (или для цветных иллюстраций). Штука сильная. Позволяет "выравнивать" цвет, делать как бы цветовой despeckle и clean и т.п. Как это работает, объяснять долго, это будет описано в будущей доке.

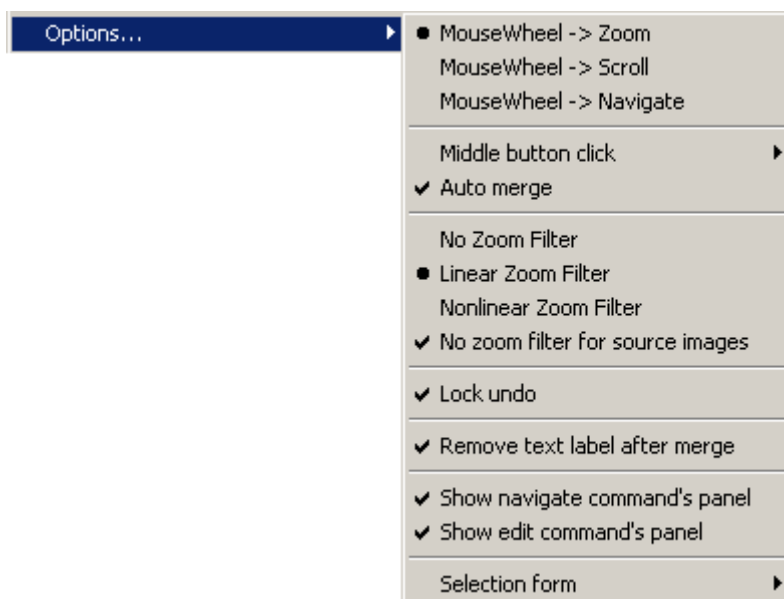
В Clear options сохраняется список цветов, использованных при magic clear (только в пределах сессии).

2.2.3 Окно поворота выделенного фрагмента




Rotate selection. (Поворот выделения). Работает для выделенного участка (можно всю страницу). Тут можно задать угол поворота, либо рассчитать его автоматом (кнопка **auto** 2), потом нажать **Paste** (положит обратно, но не сольёт, т.е. можно будет подвигать мышкой) или **Paste and merge** (положит и сольёт сразу). Сброс угла - кнопка 1, выбор направления поворота - в группе **Rotate direction**; можно выбрать алгоритм поворота в списке 3. Можно включить режим прозрачного наложения **Merge with transparency**. Есть окошко предварительного просмотра 5, где можно щёлкать мышью на картинке, и она будет увеличиваться (для возврата к исходному масштабу нажмите кнопку **fit image** 4), а ещё можно "схватить" и "тащить" картинку мышью в окне предварительного просмотра.

2.2.4 Контекстное подменю опций



Options...	
Пункт меню	Назначение
• Mouse Wheel -> Zoom	Включить режим, когда вращение колёсика мыши изменяет масштаб
Mouse Wheel -> Scroll	Включить режим, когда вращение колёсика мыши прокручивает скан
MouseWheel -> Navigate	Включить режим, когда вращение колёсика мыши листает сканы
Middle button click >	Выбор смысла работы средней кнопки мыши
✓ Auto merge	?
No Zoom Filter	Отключить фильтр отображения сканов
• Linear Zoom Filter	Включить линейный фильтр отображения сканов Одна страница 600dpi grey или color занимает в несжатом виде (а именно так windows работает с битмапами) от 50 до 100 МБ - именно из-за этого Кромсатор долго отображает такие сканы при включенном фильтре, ведь фильтр на каждую точку анализирует 9 соседних, т.е. умножьте объём ещё на 10. Нагрузка на кэш процессора колоссальная.
Nonlinear Zoom Filter	Включить нелинейный фильтр отображения сканов
✓ No zoom filter for source images	Если эта опция активна, в режиме "сравнение с оригиналом" левое изображение, т.е. оригинал, отображается без фильтра.
✓ Lock undo	Возможность блокировки undo-менеджера.
✓ Remove text label after merge	?
✓ Show navigate command's panel	Показать плавающую кнопочную панель - Панель навигации
✓ Show edit command's panel	Показать плавающую кнопочную панель - Панель редактирования
Selection form >	Выбор формы выбирающего инструмента - прямоугольник или полигон

3 Описание работы с программой

При самом первом сеансе работы с Кромсатором убедитесь, что кнопка с галкой **Auto-accept changings option is on/off** (в верхней панели) включена и галка на ней имеет зелёный цвет. Кнопка должна выглядеть вот так: . Это означает, что все изменения опций, которые Вы сделаете, будут фиксироваться автоматом. Иначе, как только Вы сместите фокус с файла, все выставленные Вами опции (например, позиции резаков) уйдут.


Коротко алгоритм работы со СканКромсатором:

(Это лишь концептуальное описание - для общего понимания сути работы с программой)

Предварительная обработка

Загружаем сырые сканы и разрезаем **вручную** (для полного визуального контроля над процессом) сдвоенные развороты или же, если это одиночные сканы, отрезаем у каждого скана ошмётки

соседней страницы в виде вертикальной полосы.


Это делается так: Выбираем в меню **Edit -> Clear all options**. Затем кликаем мышью последний файл в списке загруженных и нажимаем кнопку **Apply up to current**  - от этого на всех файлах в списке проставятся жирные зелёные галочки. Удерживая **Ctrl**, выставляем на вкладке **Pages** галку **Split** (разрезать сканы) и галку **Deskew** (чтобы применить Deskew ко всем получаемым в процессе разрезания сканам). Позиционируем резак **Set internal margin or split line** над линией смыкания 2-х соседних страниц на самом первом скане. Нажимаем правой кнопкой мыши на ползунок этого резака, и в появившемся контекстном меню выбираем пункт **Copy current position to -> all down**. Тем самым, данная позиция резака распространится на все нижележащие сканы. Теперь нужно пройти дальше по всем сканам - при помощи клавиатурных клавиш "q" и "w" (или "[" и "]") и вручную подкорректировать позицию разрезающего резака - чтобы она проходила над линией смыкания 2-х соседних страниц на всех сканах. При необходимости повторяйте **Copy current position to -> all down**.

Выставляем прочие нужные опции (**Output dir** и т.п.) и нажимаем кнопку **Process!**. Ждём, пока все сканы не разрежутся.

Выходим из программы.

3.1 Окончательная обработка

Загружаем в программу полуобработанные после 1-го этапа сканы. Нажимаем кнопку **Draft**

kromsate . В открывшемся окне нажимаем ОК. Запустится процедура автоматической предварительной расстановки резаков - по всем загруженным сканам.

После окончания работы **Draft kromsate** проходим вручную по всем сканам и корректируем неправильно выставившееся резаки. Принцип в том, что резаки не должны отсекал ничего полезного, но в то же время они должны **по максимуму** отсекал всю грязь от текста. Некоторые резаки после **Draft kromsate** будут вообще отключены - тогда их можно включить (если нужно) на панели управления резаками - вот её изображение:



На разряженных страницы, где мало текста (титulyные страницы, оглавления, страницы с выступающими одиночными строками и т.д.), **Draft kromsate** автоматически переключит некоторые резаки с **синего** на **малиновый** - чтобы подчеркнуть место неуверенного распознавания контура текста. В списке файлов такой скан будет выделен чёрным жирным цветом. Зачастую следует вручную подкорректировать **малиновые** резаки (после **Draft kromsate**) и переключить их обратно на **синий** режим - для простоты.


Если нужно - можно на данном этапе сохранить задание на диск (меню **File -> Save task**) и возобновить работу позже.

Выставляем на вкладках программы желаемые опции обработки (**Despeckle** и т.п., можно повторно **Deskew** для надёжности).

Нажимаем кнопку **Process!** и ждём. Начинается автоматическая финальная обработка сканов (включая обрезку).

После обработки (по кнопке **Process!**), переключаемся на вкладку **Book**. Программа автоматически подставила вычисленные средние размеры страниц в поля **H.Gap value** и **V.Gap value**. Теперь нужно обязательно сменить там **Auto** на **Fixed**, чтобы при повторной переделке каких-то отдельных страниц (**Process current file**) их размер выдерживался равным среднему вычисленному.

Если нужно - можно на данном этапе сохранить задание на диск (меню **File -> Save task**) и возобновить работу позже.

Просматриваем полученные сканы в редакторе постобработки (меню **Result -> View all**). Если мы замечаем, что где-то неправильно отрезался отдельно стоящий номер страницы, или же кусочек строки - возвращаемся в главное окно программы, становимся на нужный скан, и корректируем нужный резак. Как правило, нужно при этом переключить этот резак с **синего** на **малиновый** - путём сброса соответствующей галки в группе **Automargins** на вкладке **Pages**. Откорректировав резаки, нажимаем кнопку  **Process current file** (обработать текущий файл) - и данный скан перенарежется по-новому.

Вот и всё - таков вкратце алгоритм работы с программой.

3.2 Что такое Draft kromsate:

Draft kromsate - это автоматическая расстановка резаков самой программой, т.е. автоматизация самой нудной и рутинной работы. **Draft kromsate** - это т.н. **алгоритм автоматической сегментации**, т.е. определение логических блоков, занимаемых на скане полезным текстом и грязью. Задача **Draft kromsate** - только расставить резаки на сканах (приблизительно), он не вычисляет размеров финальных страниц.

3.3 Что такое кнопка Process!

Основная обработка по кнопке **Process!** - это автоматический алгоритм распознавания контура текста на скане. В процессе этой обработки программа автоматически анализирует область внутри резаков, находит контур текста, и обрезает по нему. Попутно к скану применяются **Despeckle**, **Deskew** и т.п. Получается т.н. "голый текст", без полей. Далее рассчитываются средние ширина и высота голых текстов (и заносятся в поля **Page width** и **Page height** на вкладке **Book**), создаётся пустой шаблон с этими размерами, каждый голый текст копируется на копию этого шаблона (с относительным выравниванием, задаваемым в полях ввода **Page h.align** и **Page v.align** на вкладке **Pages**), а затем к этой копии добавляются поля заданного размера (размеры берутся из полей ввода **H.Gap value** и **V.Gap value** на вкладке **Book**). Так и получаются конечные обработанные сканы.

Кроме того, кнопка **Process!** служит для запуска более простых обработок, задание для которых составляет пользователь, выставив свои опции в программе.

3.4 Как работают резаки

Резаки имеют 2 состояния - **синий** режим и **малиновый** режим. Резаки - это всего лишь подсказка для работы обработки по кнопке **Process!**. Важно понимать, что **синий** резак - не режет, он лишь очерчивает зону, внутри которой обработка по кнопке **Process!** будет **полностью автоматически** искать контур текста - и по нему реально и отрежет. В случае **синего** резака линия реального разреза будет где-то внутри (зоны, очерченной **синими** резаками) и она будет параллельно данному **синему** резаку (но есть исключение - **синие** Split-резаки **Set internal margin** и **Set internal margin or split line** не меняют никогда свой цвет и всегда режут скан по тому месту, над которым находятся - т.к. при разрезании сканов пополам никакое зональное распознавание не требуется).

ВАЖНО: Между **синим** резаком и соответствующей стороной контура текста (по которой и будет реально отрезано) может находиться произвольное количество белого пространства - это никак не влияет на правильную работу обработки по кнопке **Process!**. Важно лишь, чтобы там было как можно меньше грязи - иначе она будет принята **Process!**ом за полезный текст, что приведёт к ненужному увеличению средних ширины и высоты голых текстов.

Иногда **Draft kromsate** автоматически отключает некоторые **синие** резаки - это означает, что

Draft kromsate счёл, что с данной стороны текста никакой грязи нет и резак как подсказка не нужен. Однако, это не всегда так - в этом случае нужно вручную включить нужный **синий** резак (установкой соответствующей галки на панели управления резаками) и вручную выставить его.

Однако, полностью автоматическое (алгоритмическое) определение контура текста невозможно в принципе. Поэтому было введено понятие "**малиновый** резак". **Малиновый** резак - это "настоящий" резак, т.е. он режет скан именно по тому месту, над которым находится (только потом ещё к скану тоже поля навешиваются). **Малиновый** резак - это средство ручной корректировки работы обработки по кнопке **Process!**.

Наиболее типичное применение **малинового** резака - это борьба с некорректным отрезанием ОВ-символов ("одиночные выступающие символы") - отстоящих далеко от основного текста номеров страниц, параграфов, кусочков строк и т.п. - которые программе трудно опознать как смысловые элементы - для этого потребовалось бы проводить полноценный OCR (т.е. распознавание текста).

4 Version 5.93 Что нового

- Унифицированы диалоги импорта pdf и djvu. Добавлена возможность импорта произвольного набора страниц. Например 2-10,15-20,66-. Если после последнего тире не стоит номер, то он считается равным последнему номеру страницы в файле
- Улучшен импорт pdf
- Добавлен фильтр Smart blur (окошко Gray enhance). Предусмотрено три вида силы размытия. Самый безопасный – Low. Для фильтра задается: Radius (чем больше, тем сильнее размытие), Threshold – контролирует защиту от размытия контуров букв, Passes – количество проходов. Задается также место применения фильтра: перед чисткой фона, после чистки фона, перед фильтром sharpen из закладки Quality.

Фильтр следует применять с осторожностью при наличии букв с тонкими перемычками.

- В опции picture-зон добавлен фильтр Smart blur
- При создании pdf можно использовать jpg2000-сжатие. СК использует хоть и demo, но полностью работоспособную версию кодека Kakadu. Кодек можно скачать по адресу

<http://www.kakadusoftware.com/Executables/Win32.zip>

Из архива потребуются только два файла:

kdu_compress.exe – поместить в любое место на диске, путь к файлу прописать в настройках СК: File->Options->Apps->JPG2000 codec

kdu_v61R.dll – поместить в папку windows\system32.

Чтобы СК кодировал серые/цветные сканы в pdf формате JPG2000 (JPX в терминологии Adobe), на закладке PDF необходимо включить соответствующую опцию. Опция влияет и на способ кодирования зон, однако его можно при необходимости изменить для всех или избранных зон. Все внешние зоны (т.е. используемые в задании для создания pdf) имеют собственный переключатель, задающий индивидуальный формат кодирования зоны:

default (формат наследуется из глобальной опции на закладке PDF), JPG или JPG2000 (кодирование с потерями), TIF (кодирование без потерь). Степень сжатия для JPG и JPG2000 задается независимо друг от друга.

- Новая команда в меню Service: JPG2000 batch converter. Выполняет пакетное конвертирование графических файлов в формат JP2 с использованием кодека Kakadu.

- Улучшен draft kromsate: уменьшено количество ошибок при расстановке резаков (отрезания номеров страниц, «спотыкания» на линиях таблиц и крупных иллюстраций, а также на мусоре в нижней или верхней части скана). Автоматическое определение вертикального выравнивания

страницы (при наличии достаточных условий для такого определения).

- Добавлен новый метод бинаризации Auto. Работает хорошо в большинстве случаев. При Auto после обработки СК отображает вычисленное им значение порога. При необходимости можно переключиться на метод Custom, и подкорректировать порог, используя вычисленное значение при Auto как базис.

- Новый метод выравнивания освещенности Soft. Метод мягкий, не портит иллюстраций, но не такой чистящий, как Normal. Рекомендуется использовать метод Normal со значением Strength=5 для страниц без иллюстраций и метод Soft для страниц с картинками. Метод Normal прекрасно удаляет локальные неровности, в том числе и избавляет от содержимого, просвечивающего с обратной стороны листа. Метод Soft такими свойствами не обладает.

Убран метод Aggressive. В ближайшей версии также будут убраны методы Combo и Smart. При открытии заданий, в которых использовались упомянутые методы, они будут автоматически заменены на Normal, т.к. их действие практически идентично.

- В File->Options->Processing новая опция Crop page content at finalization stage. При включенной опции на стадии финализации контур страниц обрезается по текущим размерам с небольшим запасом. Запас задается в поле safe gap (в относительных единицах). **КРАЙНЕ НЕ РЕКОМЕНДУЕТСЯ** включать опцию Crop page!!! Это может привести хоть и к незначительным, но обрезаниям части букв!

- Новый метод despeckle – Safe50. Это вариация метода safe. Такой же безопасный как и safe, но кластеры букв чистит несколько лучше. Рекомендуется для плохо-пропечатанных сканов, а также при наличии букв с тонкими перемычками и разрывами всегда использовать метод Safe50. Метод можно задать как глобально для всей книги (на закладке опций Options), так и выборочно для отдельных страниц (Special на закладке опций Page).

- Все типы зон можно вращать мышкой. Для этого необходимо нажать Shift и подвести указатель мышки к любой вершине зоны. Как только форма указателя сменится на закругленную стрелку, нужно нажать клавишу мышки и движением производить вращение. Вращение происходит вокруг точки – геометрического центра зоны. Вращать можно как прямоугольные, так и непрямоугольные зоны, как в обычном режиме, так и в режиме редактирования вершин.

Таким образом, при отпущенном Shift, при указателе вблизи вершины, мышка изменяет положение вершины зоны (размер зоны), при нажатом Shift – поворачивает зону.

- Улучшен skew. Теперь он справляется и со сложными случаями

- При отключенной опции Split внутренний резак Int1 скрывается (ранее он отображался всегда)

- Добавлена поддержка перетаскивания файлов в окно SK. Можно перетаскивать графические файлы, файл задания skt (только один), а также pdf и djvu. Перетаскивание файлов сразу нескольких форматов не допускается, напр., если в перетаскиваемом списке помимо изображений есть pdf- или djvu-файлы, то он будет проигнорирован.

- Если SK в процессе работы находится свернутым в tray, то его иконка будет менять свой вид: а) по окончании обработки, б) если требуется вмешательство пользователя (напр., при ошибке). Окно СК сворачивается в tray, если в Options включена опция Minimize app to tray. Если опция включена, окно сворачивается в панель задач. В этом случае по окончании обработки, либо в случае ошибки, заголовок окна будет мигать.

- Опция драфта Safe Top/Bottom теперь сохраняется в ini-файл.

- Ликвидирована зависимость от dll. Теперь они больше не нужны

- Закладка опций Convert переименована в Binarization

- В профили добавлены опции создания PDF
- Для JPG и JPG2000 файлов в основном окне СК в информационной панели отображается используемая цветовая схема (RGB/GRAY/CMYK/YcbCr/YCbCrK)