

# ScanTailor Spectre

---

## ScanTailor Spectre

---



**Version 2.0a8** | macOS (Apple Silicon)

2025.01.13

Made with Claude Code 4.5

ScanTailor Spectre transforms raw scans into clean, publication-ready pages. Import a PDF or folder of images, process through an 8-stage workflow, and export a polished PDF.

## What's New in Spectre

- **PDF Import** - Open PDFs directly, no need to extract pages first
- **Finalize Stage** - New stage for color mode selection and output format
- **Export Stage** - New dedicated PDF export with quality presets
- **Redesigned UI** - Clean, focused interface
- **Apple Silicon Native**
- **Intelligent Color Detection** - Auto-detects B&W vs grayscale vs color pages
- **Batch Processing Dialogs** - Smart summaries after each stage to catch problems early

---

## Quick Start

### 1. Import Your Scans

From the startup screen or from the file menu:

- **Import PDF** - Extract pages from an existing PDF
- **Import Folder** - Load a folder of scanned images
- **Import Project** - Load a project you have previously saved.

Recently saved projects will be shown here.

## 2. Work Through the Filters

Each filter stage refines your pages. Work top to bottom:

### Stage Purpose

---

1 **Fix Orientation** Rotate pages right-side up

2 **Split Pages** Separate two-page spreads

3 **Deskew** Straighten tilted scans

4 **Select Content** Crop to content area

5 **Margins** Set page margins

6 **Finalize** Choose B&W, Grayscale, or Color

7 **Output** Apply image processing

8 **Export** Create final PDF

To run a stage, click on the little triangular play button next to the stage you are at. Stage 8 is special. You don't need to run it, you simply click on Export to PDF.

### 3. Batch Process

You may want to try doing everything automatically if your scan is clean. If you need to split spreads to pages, run stage 2. If not, or if you have run stage 2 already go to stage 6, Finalize, and run it and each of the next stages.

After batch processing certain stages, a summary dialog appears to help catch common problems:

- **Stage 2 (Split Pages)** - Shows pages that were/weren't split, lets you force changes
- **Stage 4 (Select Content)** - Shows pages with low content coverage that might need layout preserved
- **Stage 5 (Margins)** - Warns about unsplit spreads or outlier pages that cause margin issues

These dialogs let you jump directly to problem pages and take corrective action without hunting through hundreds of thumbnails.

---

### The Workflow Explained

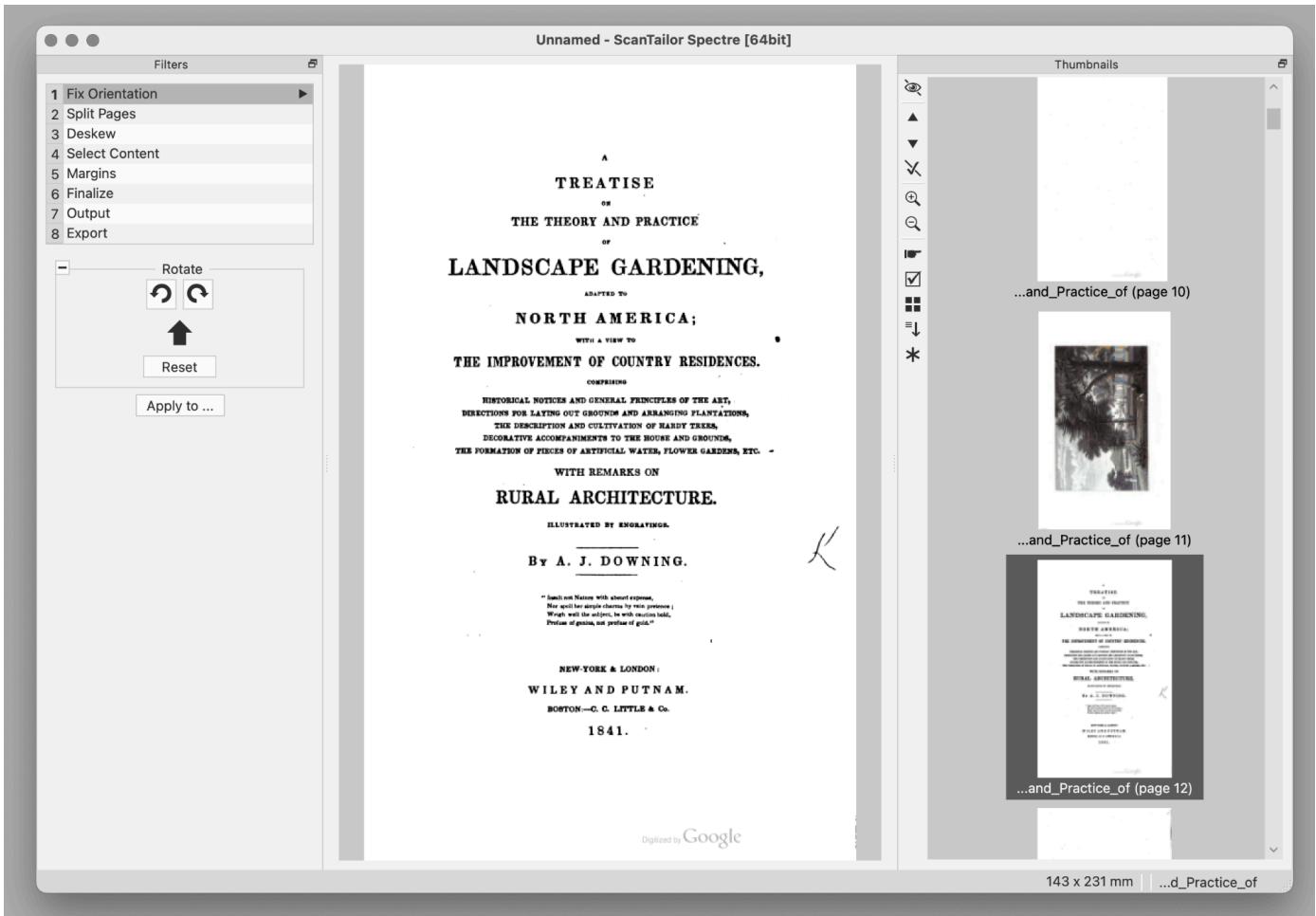
Each stage is automatic and potentially requires minimal input, but some scans can be more difficult to process than others. For example, a decent scan with only text pages might require virtually no input on your part. A 500 page book with color pages, some of which are skewed, many of which have color casts, and so on may require substantial intervention.

Although ScanTailor Spectre is designed to be as automated as possible, it can make mistakes. It's always best to check your output.

After you run a stage, the parameters that stage set are stored in your project on a per-page basis. Let's say you set a page to color once. If you rerun the automatic color detection, it will not alter that

setting even if it thinks it is grayscale.

This is beta software. Save your project frequently.



## Stage 1: Fix Orientation

Corrects pages that are upside-down or rotated. Auto-detection handles most cases; use the rotate buttons for manual override. It is rare that you will have to even run this separately, I always start at Stage 2 or later. This is not the same as Stage 3, Deskew.

## Stage 2: Split Pages

Separates book spreads (two pages per scan) into individual pages. The split line is auto-detected but can be adjusted manually.

After batch processing, a summary dialog appears with two toggle buttons:

- **Not Split (X)** - Pages kept as single pages (may need to be split)
- **Split (Y)** - Pages that were split (may have been split incorrectly)

Double-click any page to jump to it for inspection. Use the action buttons to force changes:

- In "Not Split" view: **Force Two-Page** makes pages split
- In "Split" view: **Force Single Page** makes pages unsplit

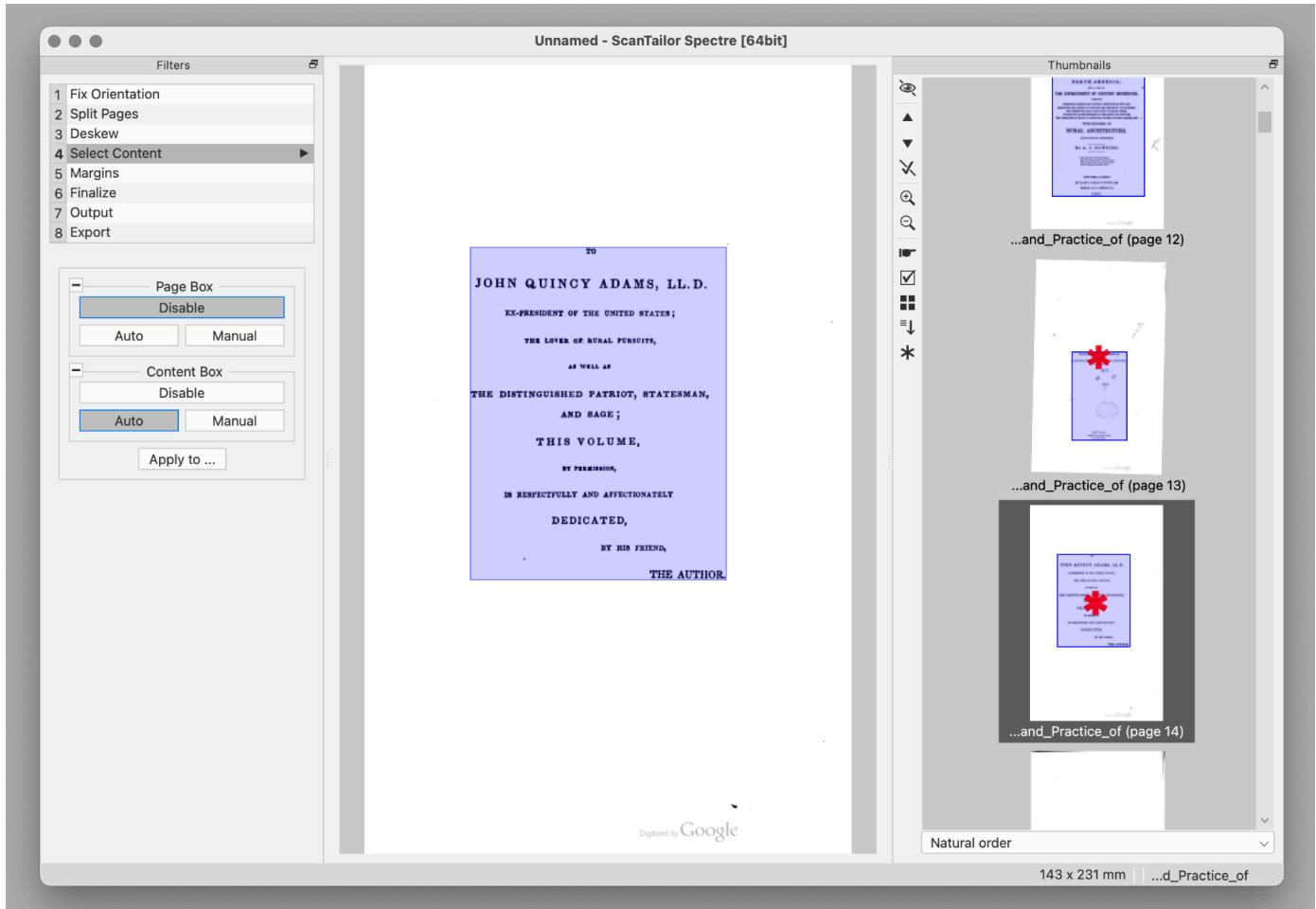
## Visual indicators in the list:

- Gray strikethrough = page was visited/jumped to
- Dark green strikethrough = action was taken (force split or unsplit)

## Stage 3: Deskew

Straightens pages that were slightly tilted during scanning. Even 1-2° of tilt is noticeable in output.

**Tip:** Sort by "decreasing deviation" to review the most-skewed pages first. The sort order is located at the bottom of the thumbnail panel on the right.



## Stage 4: Select Content

This stage has two separate but related functions: defining the **Page Box** (the physical page boundaries) and the **Content Box** (the area containing actual content like text and images).

### Page Box

The Page Box defines the boundaries of the physical page within the scan. This is useful when:

- The scanner captured more than just the page (scanner bed edges, background)
- You want to define a consistent page size across all scans
- The automatic page detection didn't work correctly

## **Page Box options:**

- **Disable** - Don't detect page boundaries; use the entire image
- **Auto** - Automatically detect the page edges (looks for the transition from scanner background to paper)
- **Manual** - Draw the page boundary yourself by dragging corners and edges

When Auto mode is selected, **Fine Tune Page Corners** adjusts corner positions by looking for black edges, useful for books where corners may be shadowed or bent.

In Manual mode, you can enter exact **Width** and **Height** values to set a specific page size.

## **Content Box**

The Content Box defines what part of the page contains actual content. Everything outside this box becomes white margin in the final output.

## **Content Box options:**

- **Disable** - Don't detect content; use the entire page
- **Auto** - Automatically find text and images on the page
- **Manual** - Draw the content area yourself

This excludes:

- Scanner bed edges
- Book margins you want to remove
- Fingers or page holders

Auto detection is not always perfect, especially with images or decorative elements. Sort by "Order by completeness" to review pages that may need manual adjustment.

## **Content Coverage Summary**

After batch processing, a summary dialog appears showing pages grouped by content coverage:

- **Low Coverage** - Pages where detected content covers less than the threshold percentage of the page area. These may have intentional layout (chapter titles, short paragraphs) that shouldn't be centered.
- **Normal Coverage** - Pages where content fills most of the page area.

Use the **Coverage threshold** slider (20%-80%) to adjust what counts as "low coverage."

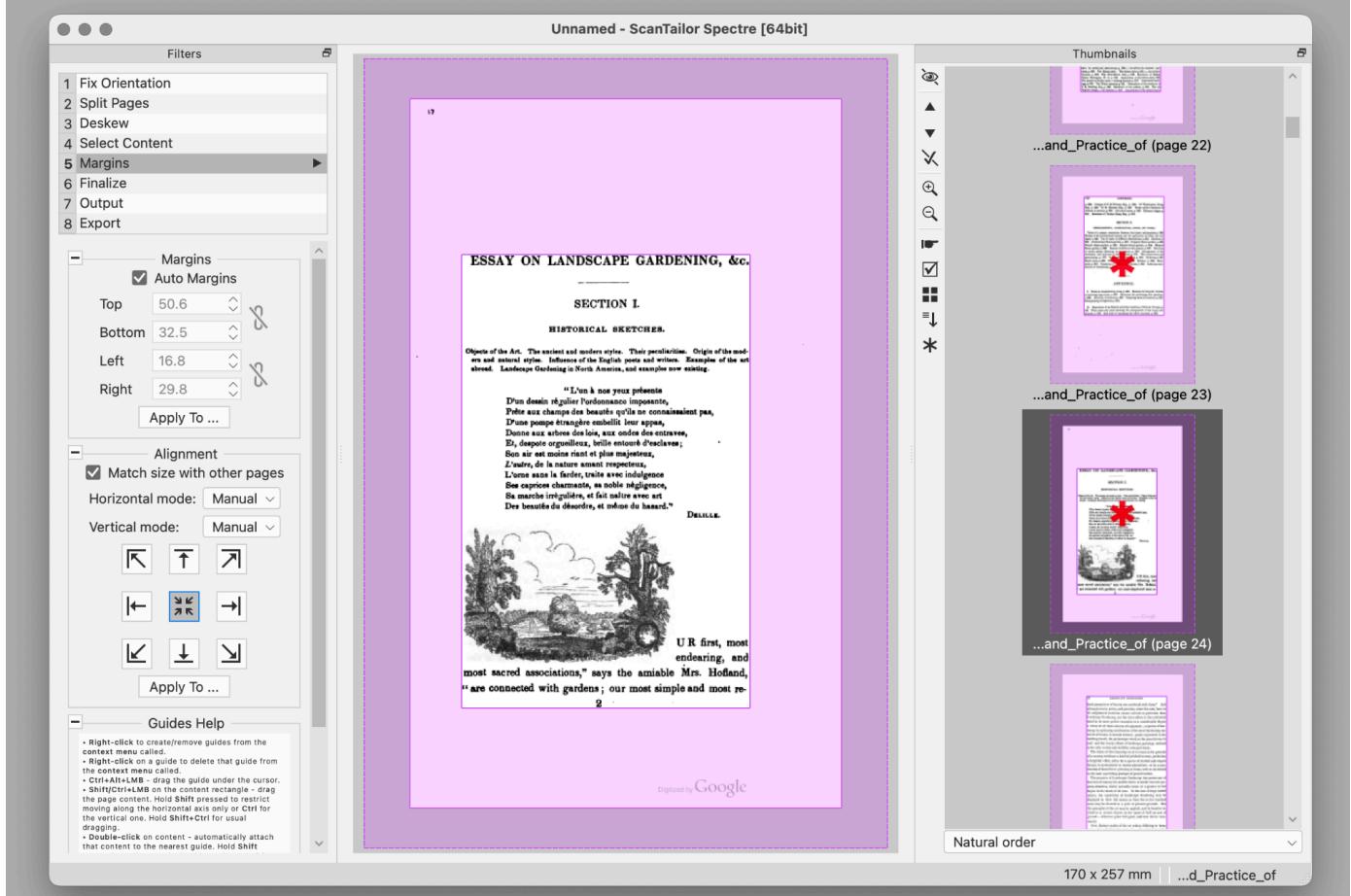
## **Actions for low coverage pages:**

- **Jump to Selected** - Navigate to a page for manual inspection

- **Preserve Layout (Selected/All)** - Disable content detection for these pages, keeping their original layout intact

## Visual indicators:

- Gray strikethrough = page was visited/jumped to
- Dark green strikethrough = action was taken (preserve layout)



## Stage 5: Margins

Sets white space around content in the final output and page size.

- **Top/Bottom/Left/Right** - Individual margin sizes
- **Alignment** - Where content sits within the page

## Page Size Warning

After batch processing, if the app detects potential problems it will show a warning dialog. This catches two common issues:

## Unsplit Two-Page Spreads:

If you have book spreads that weren't split in Stage 2, they will cause all other pages to have large margins (because the output tries to make all pages the same size). The dialog shows which pages appear to be unsplit spreads based on their dimensions.

## Outlier Pages:

Pages that are significantly larger or smaller than the median page size. These outliers can cause unexpected margins on other pages.

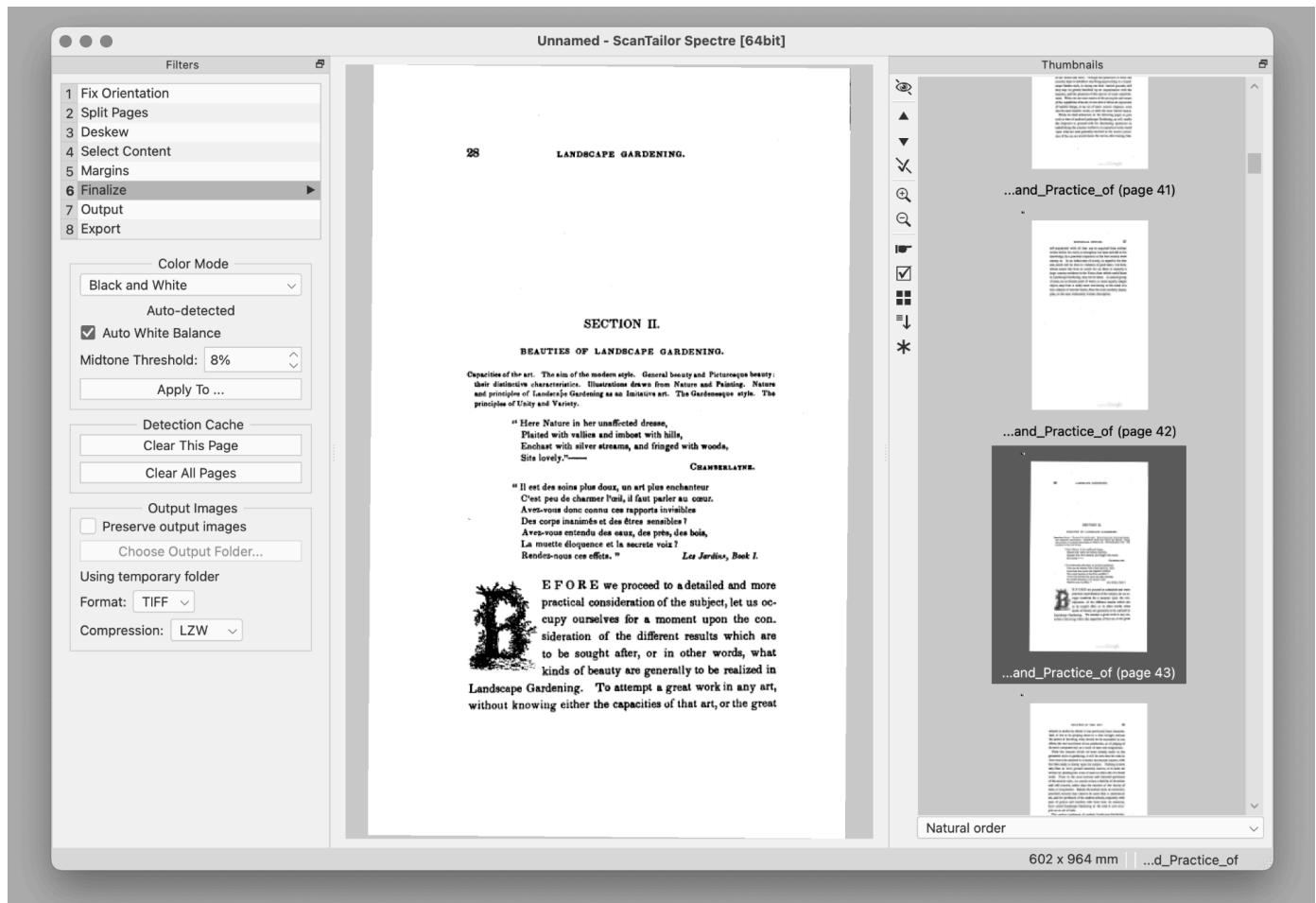
## Actions:

- **Jump to Selected** - Navigate to a page to inspect it
- **Split to Pages** - Go back to Stage 2 and split the selected spread into two pages

Double-click any page in the list to jump directly to it.

## Visual indicators:

- Gray strikethrough = page was visited/jumped to
- Dark green strikethrough = action was taken



## Stage 6: Finalize

**New in Spectre.** This stage determines how each page will be processed:

- **Color Mode:** Black & White, Grayscale, or Color
- **Output Format:** TIFF or PNG
- **Output Location:** Where processed files are saved

The app auto-detects the appropriate color mode. Obviously, Color pages require the most storage space, Grayscale requires less, and Black and White the least of all.

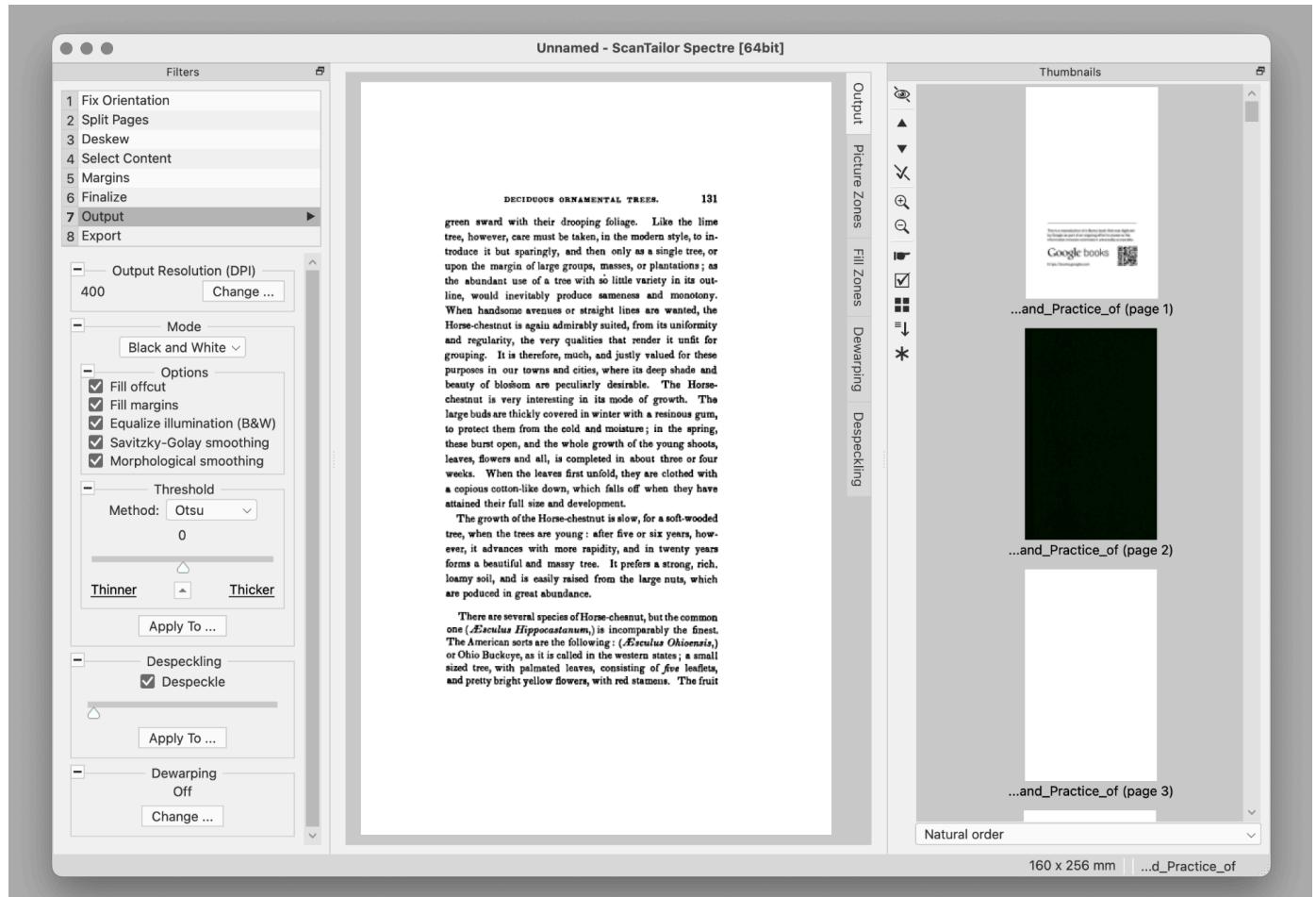
The **Midtone Threshold** slider adjusts detection sensitivity, useful for smaller images (lower = more pages will be inspected and detection will be slower). You can override the determination after inspection.

### White Balance (Color and Grayscale modes):

- **Force white balance** - Automatically corrects color casts from aged paper or poor lighting
- **Pick Paper Color** - Click this, then click on an area that should be white/neutral in the preview. The app will correct the entire page based on that sample.

**Clear All Pages** resets all pages to unprocessed state, useful if you want to re-run automatic detection with different settings.

Previous versions of Scantailor exported to images, the workflow in this one is images or PDF to PDF. Images will be discarded when you close the project unless you choose to preserve them at this point by ticking the box to "Preserve Output Images."



## Stage 7: Output

Applies image processing to generate the output files. Options vary by color mode.

**Output Resolution:** Sets the DPI of the output image. 400 DPI is the Library of Congress recommendation for most documents. Higher = sharper but larger files.

## Black & White Mode Options

- **Binarization Threshold** - Controls the cutoff between black and white. Lower values = more black, higher = more white. Adjust if text appears too thin or too bold.
- **Normalize Illumination** - Evens out lighting variations across the page. Helps with scans that have shadows or uneven exposure.
- **Morphological Smoothing** - Smooths jagged edges on text and lines.
- **Morphological Opening/Closing** - Advanced cleanup options for removing small artifacts.

## Grayscale & Color Mode Options

- **Normalize Illumination** - Evens out lighting variations.
- **Sharpen/Blur filters** - Enhance or soften the image.

## Common Options (All Modes)

- **Dewarping** - Flattens curved pages from book spines. Set to "Auto" for automatic detection or "Manual" to adjust control points yourself. Most useful for thick books where pages curve near the spine.
- **Despeckle** - Removes small dots and noise:
  - *Cautious* - Minimal cleanup, preserves detail
  - *Normal* - Balanced approach
  - *Aggressive* - Heavy cleanup, may remove fine details
- **Equalize Illumination** - Additional lighting correction.
- **White Margins** - Ensures margins are pure white.

## Zones

- **Picture Zones** - Draw rectangles around photographs or illustrations. These areas are processed differently to preserve detail and gradients instead of being binarized.
- **Fill Zones** - Draw areas to fill with a solid color (usually white). Useful for removing stamps, stains, or unwanted marks.

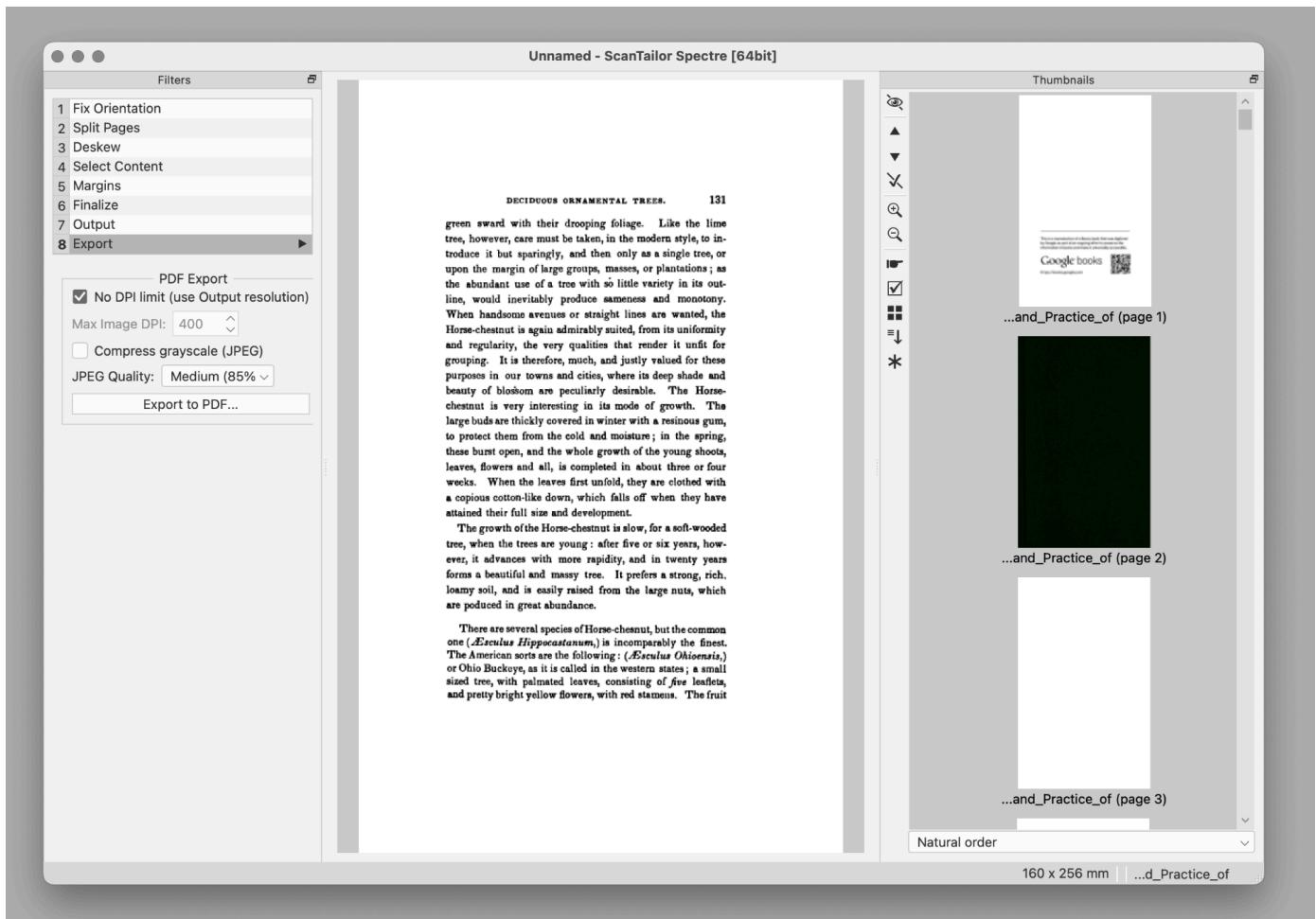
To draw a zone, select the zone tool, then click and drag on the preview image.

## Apply To...

Once you've tuned settings for one page, use **Apply To...** to copy those settings to other pages. You can apply to:

- All pages
- Selected pages only
- Pages with the same color mode

This is essential for efficiently processing large documents.



## Stage 8: Export

**New in Spectre.** Creates the final PDF.

**Max DPI:** Limits output resolution for grayscale and color pages.

400 DPI is the Library of Congress recommendation for most documents.

If your scan is lower quality or to be read on screen only, you may find that lowering the resolution as low as 72 DPI gives acceptable results and a much lower final PDF size. At 72 DPI, however, it would be very difficult to read black and white pages, which is why they are controlled by the master output resolution in Stage 7.

## Keyboard Shortcuts

### Key Action

**Page Up/Down** Previous/next page

**Home / End** First/last page

**Cmd+S** Save project

## Tips

## Scanning

- **300-400 DPI** for text documents
- **600 DPI** for fine detail or small text
- Scan in **color** even for B&W content - better conversion results

## Processing

- Work through stages **in order** - each depends on the previous
- Configure **one page well**, then batch apply to similar pages
- Use **page ordering** options to find problem pages
- You can right click on the thumbnails to remove pages from the project
- **Save frequently** - all settings are preserved in the project file

## Color Mode Selection

- **B&W (1-bit)**: Smallest files, pure black and white, ideal for text-only
- **Grayscale**: Better for pages with photos or illustrations
- **Color**: Only when color information matters

## Typical File Sizes (100-page book)

- B&W PDF: 5-15 MB
- Grayscale PDF: 20-50 MB
- Color PDF: 50-150 MB

---

## Project Files

Projects are saved as **.ScanTailor** files containing:

- References to source images (not copies)
- All settings for every stage
- Processing state

**Important:** Keep source images in place - projects reference them by path.

---

## Credits

ScanTailor Spectre is based on:

- [ScanTailor Experimental](#) by zvezdochiot
- [ScanTailor Advanced](#) by 4lex4
- [ScanTailor](#) by Joseph Artsimovich

## License

GPL-3.0 - See [LICENSE](#) for details.