ScanScribe: Perceptually Supported Diagram Image Editing

Eric Saund and James Mahoney

Palo Alto Research Center
Palo Alto, CA 94304, USA
{saund,mahoney}@parc.com

Abstract. We have implemented a prototype document image editor that incorporates principles and algorithms for Perceptual Organization in order to facilitate selection and manipulation of visually salient image objects. Called *ScanScribe*, the program serves two purposes. First, it offers an illustration of the advantageous use of Gestalt principles of segmentation and grouping of text and line-art found in diagrams. Second, it is a practical tool for modifying existing diagrams and composing new diagrams from mixed source material. ScanScribe's user interface design and foundational representations are designed to scale to support recognition of domain objects found by structural matching through subgraph correspondence or other techniques.

1 Introduction

Over the past decade many investigators have identified the goal of building "intelligent" computer programs for human work with sketches and diagrams. The companion poster to this one discusses this motivation and associated experimental artifacts. This poster presents our efforts to build a practical image editing tool embodying some of the perceptual abilities of human users.

2 Artifact: The ScanScribe Image Editor

A fundamental operation of image and document editing programs is selection. Once the user has selected certain desired markings, they may proceed to move, copy, delete, rotate, scale, or otherwise modify them. Our research has led to the implementation of a number of programs for perceptually-supported selection of diagrammatic material. The most mature of these, called *ScanScribe*, is available for download and will be offered as a demonstration to accompany this poster [4]. ScanScribe implements basic forms of foreground/background separation, object segmentation, and Gestalt grouping. Grouping principles currently implemented are smooth continuation of line-work, the detection of perceptual (but not necessarily geometrically) closed paths in line-art [3], and grouping of text elements by spatial proximity and alignment. See Figure 1.

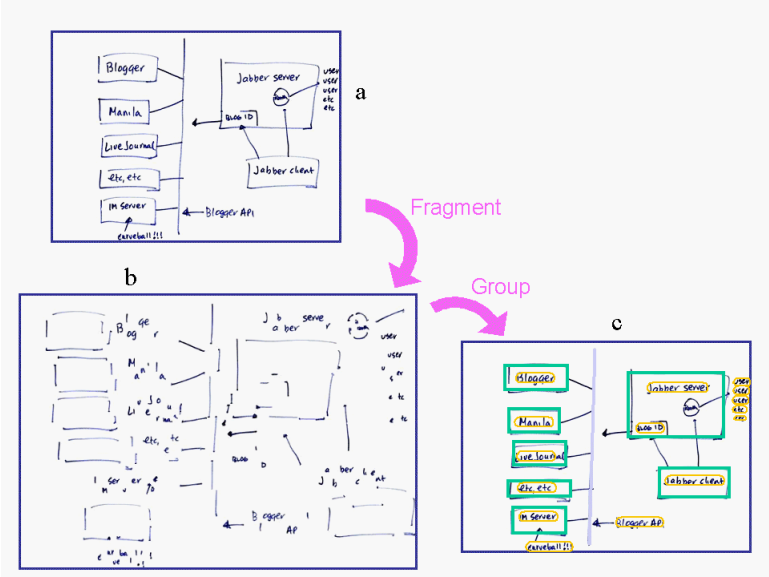


Fig. 1. Stages of perceptual organization. a. Input image. b. Image fragments corresponding to individual characters, blobs of touching characters, and simple curves. c. Grouping rules attempt to identify groups of fragments exhibiting smooth continuation, perceptually closed contour paths, and spatial proximity and alignment.

Using ScanScribe, it is easy to modify existing diagrammatic material starting from scanned images. Figure 2 illustrates. Here, a practitioner of Category Theory wishes to illustrate through a diagram that commutativity may be applied to derive one formal structure from another. After importing the image of Figure 2a into the program from a tiff file, ScanScribe performs a number of steps that facilitate transformation of the original into the desired figure. First, foreground markings are distinguished from from white or light background; the background is then rendered transparent. Next, foreground markings are decomposed into elemental fragments corresponding to characters and simple curves. Finally, groups are formed of proximal text and aligning curve fragments. Sophisticated user interface techniques are used to offer the user a combination of selection by enclosure and selection by point-and-click. In just a few minutes Figure 2c is obtained. As part of the process, the practitioner wished to include an arcing directed arrow between two nodes. They sketched the arrow freehand on scratch paper, then used a digital camera to import the scratch paper image into ScanScribe. ScanScribe's foreground/background separation algorithms eliminated lighting artifacts, and the arrow was easily sized and rotated into position. In other words, the program enables the free and fluid mixture of diverse image material, aspiring to a WYPIWYG (What You Perceive is What You Get) style of smart interface.

3 Prospect: Recognition of Visual Language Structure

To take smart diagram manipulation to the next level will involve recognition of more diagram-specific and domain-specific aspects of structure, above and beyond Gestalt principles. For example, Figure 2a is recognizable as an instance of a generic class, the Node/Link diagram. As such, one might expect that operations such as moving nodes should cause their associated links to follow in order to preserve syntactic connectivity. To recognize shapes and configurations in diagrams such as these, our group has adopted the well-known approach of subgraph matching. We have found that perceptual organization techniques offer benefits with regard to some of the difficulties these methods have traditionally been found to encounter [2, 1]

While the capability to recognize and respect spatial structure at the level of visual language constructs is in some respects readily available to today's structured graphics editors, the ability to manipulate in this way, say, *any* node-link diagram obtained from *any* source remains a significant research goal. To achieve it will raise the image itself to the status of common currency between perceptually-enabled machines and human visual cognition.

References

1. Mahoney, J., and Fromherz, M.;[2002]; "Handling Ambiguity in constraint-based recognition of stick figure sketches," *Proc. Document Recognition and Retrieval IX*, San Jose, California.
2. Messmer, B.;[1995]; "Efficient graph matching algorithms for preprocessed model graphs," PhD thesis. Bern Univ., Switzerland.
3. Saund, E.; [2003]; "Finding Perceptually Closed Paths in Sketches and Drawings," *IEEE TPAMI*, V. 25, No. 4., 475-491.
4. Saund, E., Fleet, D., Larner, D., and Mahoney, J.;[2003]; "Perceptually-Supported Image Editing of Text and Graphics," *Proc. ACM UIST*, Vancouver, BC, 183-192.

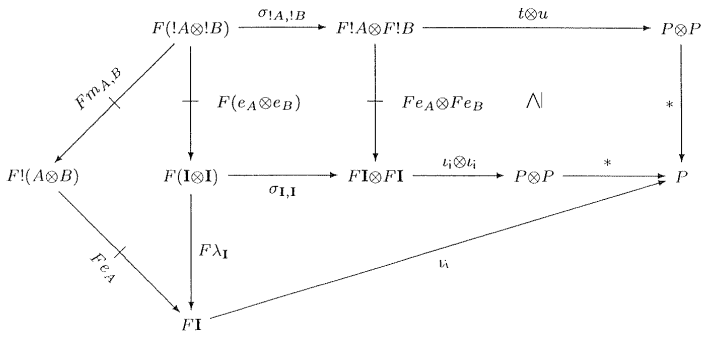
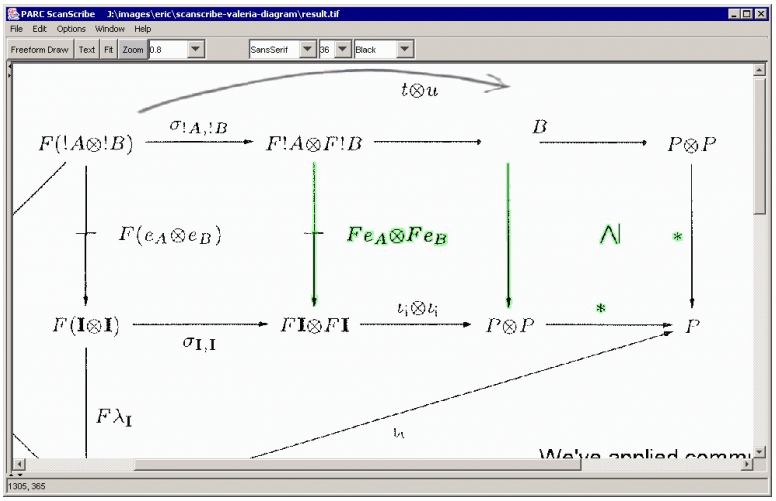
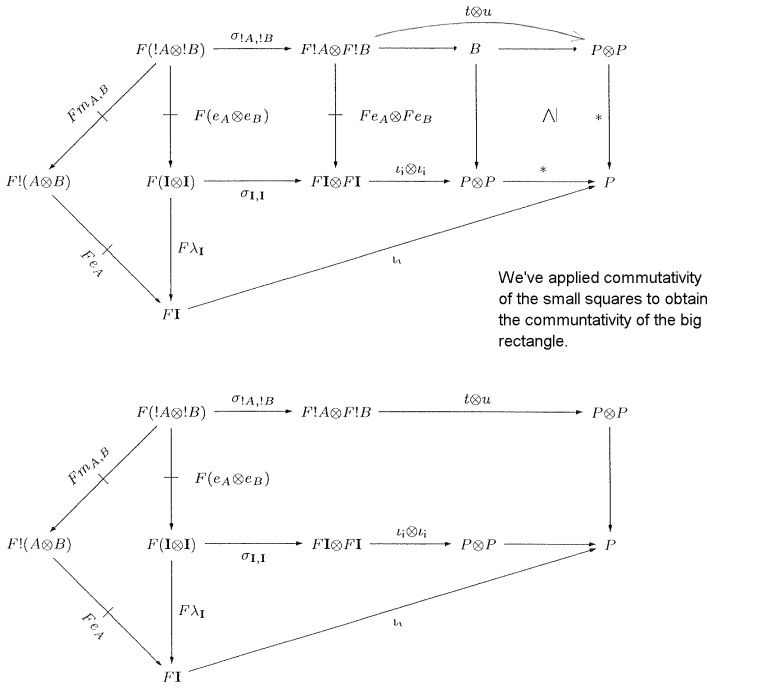


Fig. 2. a. Diagram scanned from paper. (Full caption is on following page).



b



c

Fig. 2. Image-based editing of a Category Theory diagram using the ScanScribe image editor. a. Diagram scanned from paper. b. Screen snapshot of material selected by point-and-click (clearly visible in color as a green halo behind several image objects). c. Modified figure derived from a, illustrating a point about a derivation in this domain. Note the added arc-arrow which was imported from a digital camera photo of an arrow drawn freehand on scratch paper, then scaled and positioned within the diagram.