PPTLens: Create Digital Objects with Sketch Images

Changcheng Xiao*
Shanghai Jiao Tong University
Shanghai, China
xchangcheng@gmail.com

Changhu Wang†
Microsoft Research
Beijing, China
chw@microsoft.com

Liqing Zhang
Shanghai Jiao Tong University
Shanghai, China
zhang-lq@cs.sjtu.edu.cn

ABSTRACT

In this work, we introduce the PPTLens system to convert sketch images captured by smart phones to digital flowcharts in PowerPoint. Different from existing sketch recognition system, which is based on hand-drawn strokes, PPTLens enables users to use sketch images as inputs directly. It's more challenging since strokes extracted from sketch images might not only be very messy, but also without temporal information of the drawings. To implement the 'Image to Object' (I2O) scenario, we propose a novel sketch image recognition framework, including an effective stroke extraction strategy and a novel offline sketch parsing algorithm. By enabling sketch images as inputs, our system makes flowchart/diagram production much more convenient and easier.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing; H.5.2 [**User Interfaces**]: User-centered Design

Keywords

Offline Sketch Recognition; Sketched Flowchart Recognition

1. INTRODUCTION

Drawing flowcharts and diagrams on a whiteboard or tablet is an effective way for users to convey and express information. Although with the development of touch-screen devices, increasing research is conducted on hand-drawn flowchart or diagram recognition, most work targets at interpreting online drawings with temporal information (online sketch recognition), with little effort on understanding sketch images captured by smart phones (offline sketch parsing). However, it will make flowchart/diagram production much more convenient and easier if we can use sketch images as inputs, especially with the popularity of smart phones.

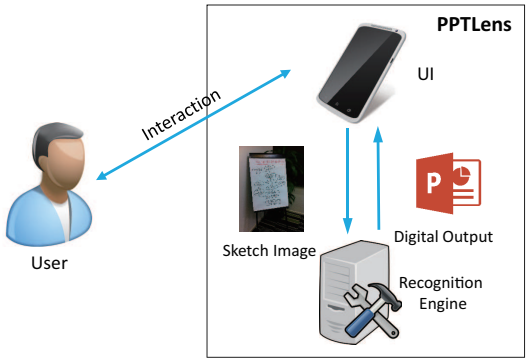


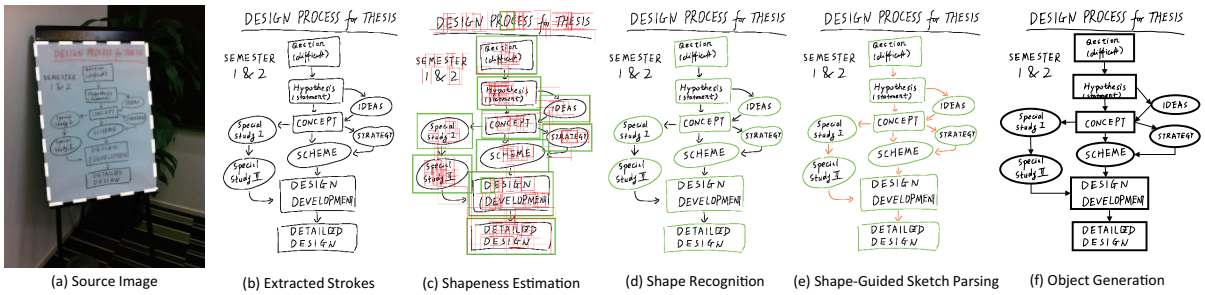
Figure 1: Illustration of the PPTLens system. Users can take a photo of the sketched flowchart, which can be automatically converted to a digital flowchart in PowerPoint with shape symbols and arrows replaced by form digital objects.

In spite of the increasing interest in building system that can automatically recognize sketch images, there is still a huge gap between sketch images and digital objects. This gap mainly comes from the following two aspects: **1) the challenges of extracting strokes from sketch images**. Different from online sketch recognition [5], in which strokes can be obtained directly from the user interface, it's very challenging to extract strokes from sketch images. On the one hand, sketch images may have uneven drawing quality and large variations of luminous intensity, resulting in low stroke extraction recall. On the other hand, the missing of drawing orders results in high error rate for the stroke construction. **2) the challenges of offline sketch parsing**. To recognize the symbols of a sketch, a typical approach in the literature is to first generate a number of candidate stroke groups, followed by the recognition of each group [3, 4]. Theoretically, for N strokes, there will be 2^N different stroke groups to recognize. To avoid the exponentially growing recognition cost, several constraints are leveraged, such as temporal constraint [1] and spatial constraint [4]. However, the strokes extracted from sketch images might be very messy, resulting in much more groups. Moreover, the unavailable of the temporal constraint makes it harder to reduce the search space.

To bridge the gap between sketch images and digital objects, in this work, we first propose an effective stroke ex-

*Changcheng Xiao performed this work while being an intern at Microsoft Research Asia.
†Corresponding author.

MM'15, October 26–30, 2015, Brisbane, Australia.
ACM 978-1-4503-3459-4/15/10.
DOI: http://dx.doi.org/10.1145/2733373.2807974 .



(a) Source Image (b) Extracted Strokes (c) Shapeness Estimation (d) Shape Recognition (e) Shape-Guided Sketch Parsing (f) Object Generation

Figure 2: Illustration of the sketch image recognition framework. Given a hand-drawn flowchart image, we first detect the region with sketched contents (a), and extract strokes from the detected region (b). Then a large number of stroke groups are generated as candidates (red&green bounding boxes), followed by the shapeness estimation algorithm to select a few 'good' candidates (green bounding boxes) in (c). The selected 'good' candidates are then fed into a robust recognizer to predict their specific shape types in (d), followed by the shape-guided parsing to recognize all other parts (e). The final sketch parsing results are converted to a digital flowchart in PowerPoint that can be further edited (f).

traction strategy, which is based on the observation that sketches hold almost fixed stroke width. The main process of our stroke extraction strategy is: 1) detect the borders of the whiteboard/paper, then crop and rectify the image [6], 2) binarize the image by leveraging the Stroke Width Transform [2] and MRF optimization, 3) extract the skeleton of the binary image and construct strokes. To reduce the search space of sketch recognition efficiently, we propose the idea of shapeness estimation, measuring how likely a stroke group could be a *good* shape. Owing to the shapeness estimation process, our system can efficiently remove most of the *bad* shape groups while still preserve high recall for the *good* shape groups.

Based on the above technologies, we present a novel sketch image recognition framework (Fig. 2) and build the PPTLens system to recognize camera-captured sketch images and convert them to digital flowcharts in PowerPoint.

2. SYSTEM OVERVIEW

In this section, we introduce the main functions of the PPTLens system from the following two aspects:

System Input

The input of the PPTLens system should be an image, containing a flowchart drawn on a whiteboard or a paper. A typical flowchart, as in Fig. 2, mainly contains three categories of objects, i.e., shape symbols, non-shape symbols and noises. The five basic shape symbols are process (rectangle), decision (diamond), terminator (ellipse), data (parallelogram) and connection (circle). The two non-shape symbols are arrow and text. In this work, we mainly recognize the five shape symbols and arrows while leaving the text recognition for future work.

System Output

The output of the PPTLens system is a PowerPoint file with three slides. The first slide contains all the strokes extracted from the sketch image. The second slide contains the unrecognized strokes and the formal digital objects replacing recognized strokes. The third slide contains the original input image.

Fig. 3 shows two example sketch images and their corresponding results of the PPTLens system.

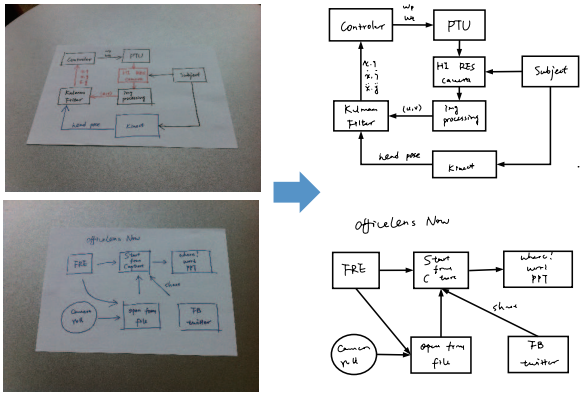


Figure 3: Example sketch images and corresponding results.

3. ACKNOWLEDGMENTS

The work of first and third authors was supported by the National Natural Science Foundation of China (Grant No. 61272251) and National Key Basic Research Program of China (Grant No. 2015CB856004).

4. REFERENCES

[1] C. Carton, A. Lemaitre, and B. Couasnon. Fusion of statistical and structural information for flowchart recognition. In *ICDAR*, 2013.

[2] B. Epshtein, E. Ofek, and Y. Wexler. Detecting text in natural scenes with stroke width transform. *CVPR*, 2010.

[3] L. B. Kara and T. F. Stahovich. Hierarchical parsing and recognition of hand-sketched diagrams. *UIST*, 2004.

[4] E. J. Peterson, T. F. Stahovich, E. Doi, and C. Alvarado. Grouping Strokes into Shapes in Hand-Drawn Diagrams. *Artificial Intelligence*, 2010.

[5] J. Wu, C. Wang, L. Zhang, and Y. Rui. Sketch Recognition with Natural Correction and Editing. AAAI, 2014.

[6] Z. Zhang and L. W. He. Whiteboard scanning and image enhancement. *Digital Signal Processing*, 2007.