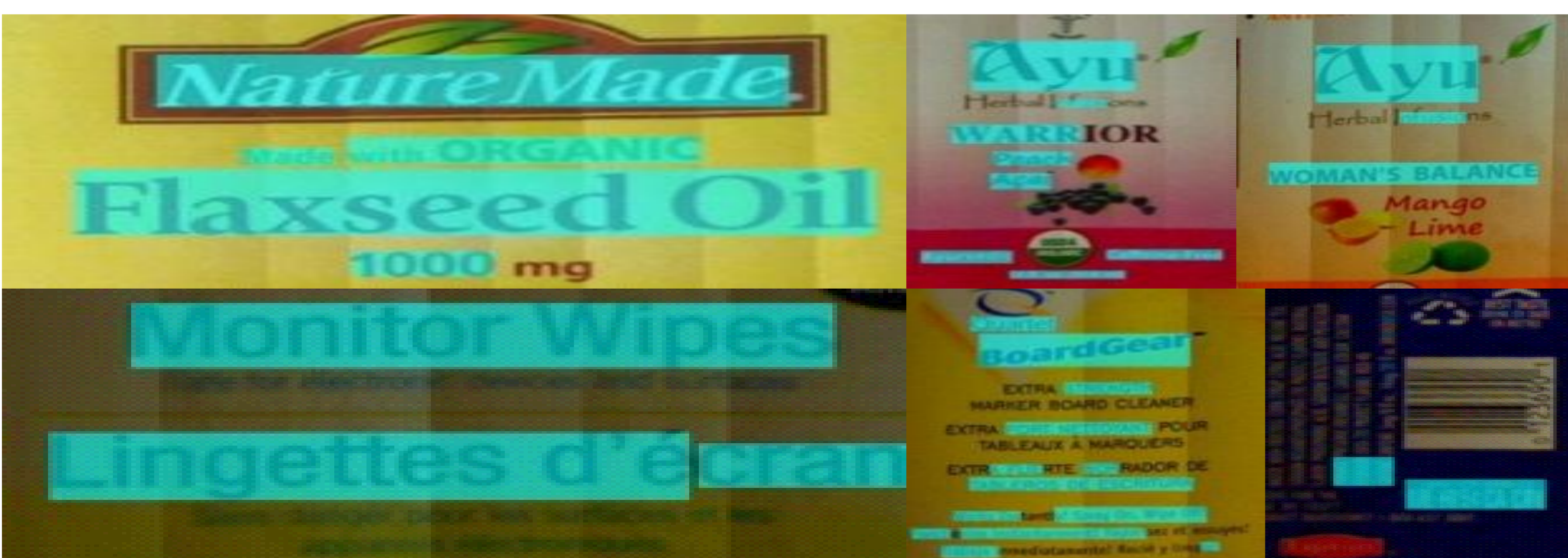
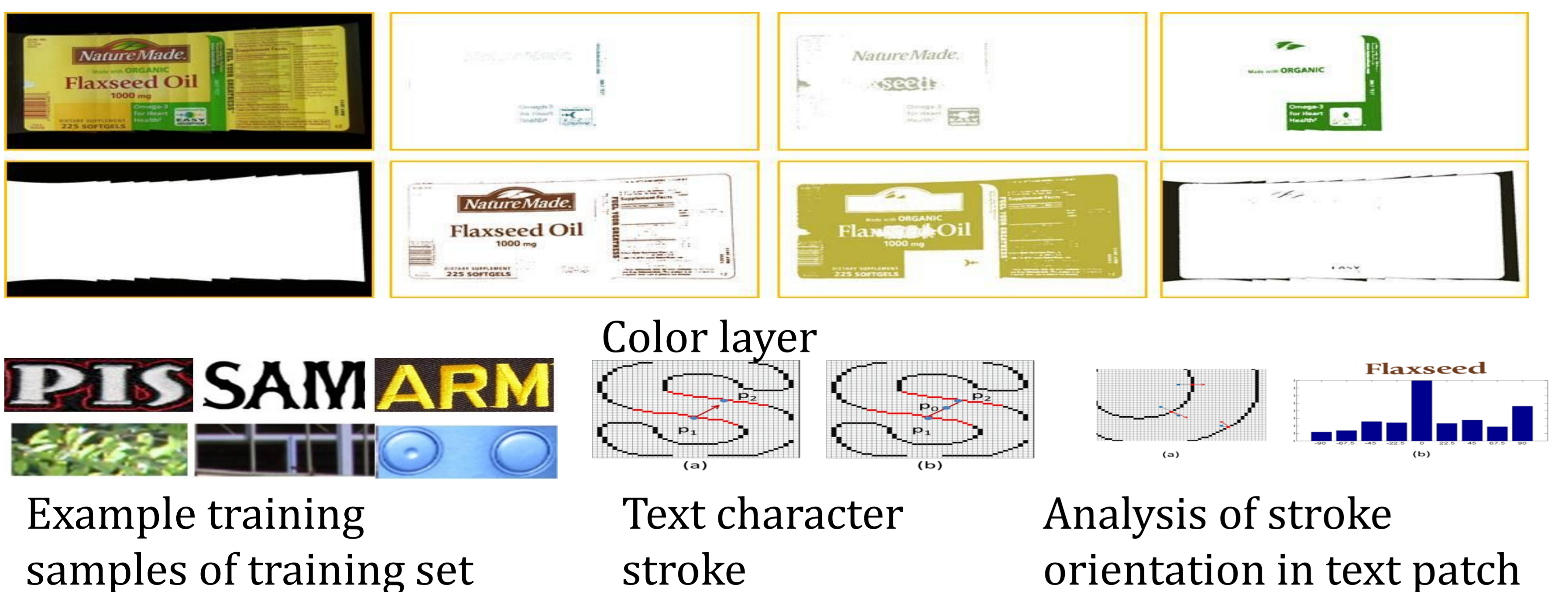
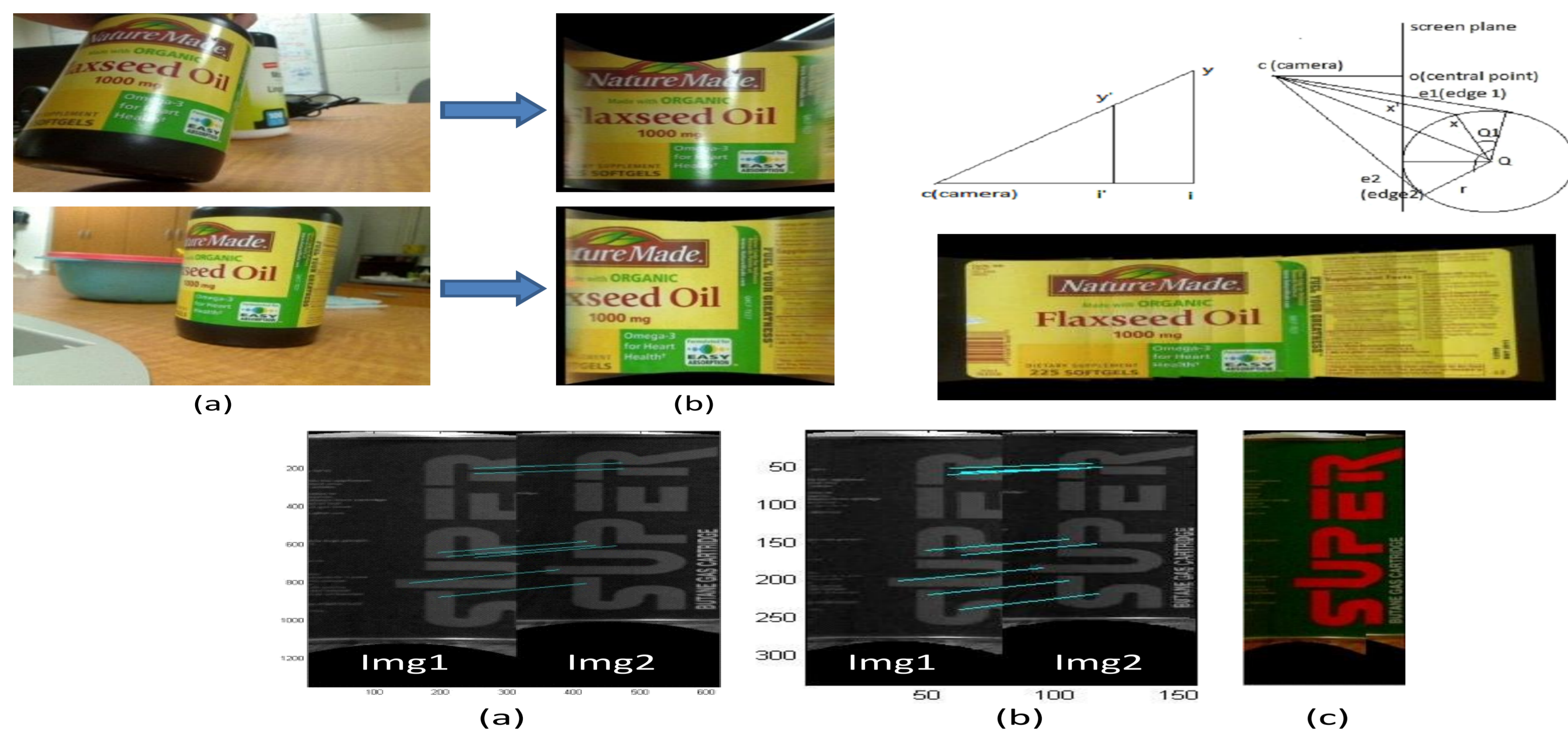
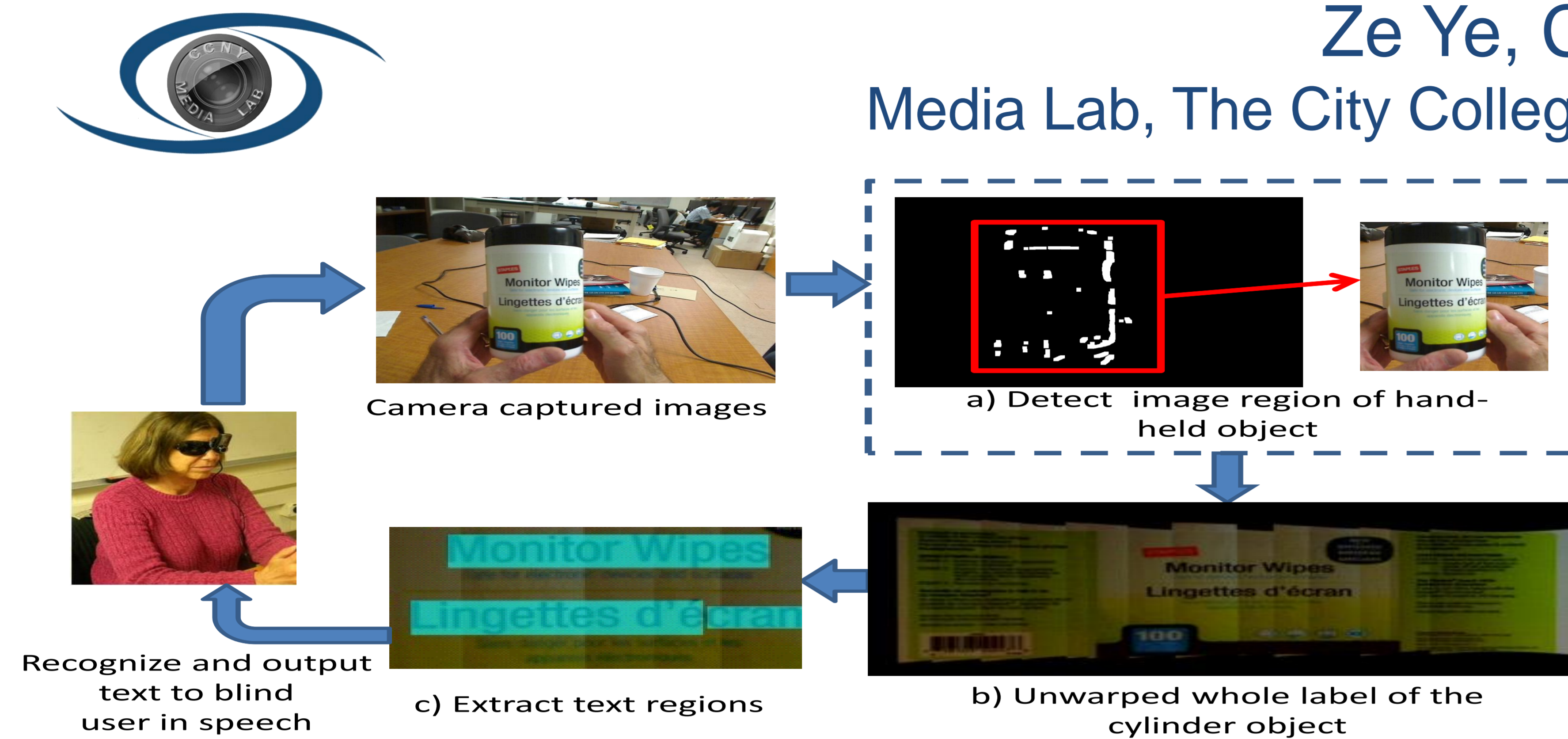


READING LABELS OF CYLINDER OBJECTS FOR BLIND PERSONS

Ze Ye, Chucai Yi, and Yingli Tian

Media Lab, The City College of New York, The City University of New York



Introduction

- A camera-based assistive framework to help blind persons to read text labels from cylinder objects in their daily life.
- Image stitching to mosaic all images taken by camera and text extraction to read the involved text information from complex backgrounds.
- Challenge
 - Objects in camera-based images is surrounded by various background outliers, and text characters usually appear in multiple scales, distortions, fonts, colors, and orientations.
 - The existing OCR software was designed for recognizing text information from scanned documents and cannot handle non-flat surfaces.
 - Only small portion of the label which faces to the camera is visible.

Unwarp Cylinder Object Surface

- Unwarping
 - The orientation of cylinder objects is normalized into vertical based on camera perspective projection.
 - Each pixel in the cylinder surface will be projected to its tangent plane.
- Stitching
 - Preprocessing the camera-based images by histogram equalization, threshold selection and image down-sampling to half size.
 - SIFT-based matching points will be calculated between two overlapped images of different portion of the object label. We then use RANSAC algorithm to estimate the transformation matrix in order to stitching all the flattened images together.

Automatic Text Localization

- Layout Analysis
 - Color uniformity
 - Horizontal alignment
- Character Structure Modeling and Learning
 - Geometrical properties, e.g., size, aspect ratio
 - Stroke orientation, stroke width, and stroke ratio
 - Gradient distribution
 - Cascade Adaboost learning model to obtain text classifier from training set of text patches.

Dataset and Experimental Results

- A dataset of camera-based cylinder objects, containing 300 images from 16 cylinder objects.
- Not all the objects appear in upright position. Some of the cylinders have more than 70 degrees between the camera optical axis and the object axis.
- All objects are selected captured as testing images to evaluate our system. Text localization algorithm is performed on the scene images of Robust Reading Dataset.
- Blue rectangles were drawn out after text localization. This algorithm has high accuracy for large words that have 3 or more than 3 characters.

Supported by: NSF, NIH, FHWA, and Microsoft Research.