Introduction

Today’s widespread access to smartphones provides an outlet for the enjoyment of comics. Automatically recognizing both the characters and text within a comic strip could allow everyone to experience comics through text-to-speech (regardless of their ability to physically see the comic strip). To achieve this, tens of thousands of comics need to be analyzed. Our project describes the first steps to accomplish this goal, by segmenting frames and recognizing characters in comic strips.

We focused on the comic strip Garfield. We implemented both a new algorithm, Color Histogram Area Recognition (CHAR), and a standard SIFT matching approach.

Rectification and Frame Isolation

1. Original Image
   - Image taken by Droid camera
   - Angle of rotation varies ± 20 degrees

2. Image Rotation
   - Image is locally thresholded to accentuate dark edges in the image
   - Hough transform used to rotate image from the strongest peak

3. Frame Isolation and Cropping
   - Regions labeled and filtered according to bounding box area and aspect ratio. These frames are then cropped and passed to SIFT or CHAR.

Recognition Methods – CHAR and SIFT

CHAR

1. Frame Color Clustering
   - Frames converted to LAB color space
   - Prominent colors clustered into regions

2. Isolation of distinct cluster regions
   - Color clusters are isolated
   - Open-Close applied to isolate blobs

3. Recognition of Garfield Characters
   - Color histogram of region found
   - Histogram compared with those known from training set to determine character

SIFT

1. Build a Vocabulary Tree
   - Find SIFT keypoints for a database of comics and generate a vocabulary tree

2. Find Set of Best Matches
   - Use vocabulary tree to find the database images that most likely match the input comic strip

3. Locate Characters
   - Find best transform to map database images to input comic
   - Only keep matches that produce a high number of matches after the application of the transform

Experimental Results

- Detection results computed using 100 three-frame comic strips
- Number of characters in test set:

<table>
<thead>
<tr>
<th>Character</th>
<th>Correct Detections (SIFT)</th>
<th>Incorrect Detections (SIFT)</th>
<th>Correct Detections (CHAR)</th>
<th>Incorrect Detections (CHAR)</th>
<th>Mixed Detections (SIFT)</th>
<th>Mixed Detections (CHAR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Garfield</td>
<td>273</td>
<td>165</td>
<td>20</td>
<td>13</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Jon</td>
<td>165</td>
<td>289</td>
<td>73</td>
<td>414</td>
<td>278</td>
<td>78</td>
</tr>
<tr>
<td>Odie</td>
<td>20</td>
<td>73</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Liz</td>
<td>13</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Arlene</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pooky</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- Overview of detection results:

Related Work


Future Work

- Implement edge-based character detection and determine any improvements in functionality
- Test other classification methods (such as svm)
- Implement an algorithm that combines features of both the CHAR and SIFT methods for character recognition