Efficient and robust image descriptor for GUI object classification

Anastasia Dubrovina, Pavel Kisilev, Daniel Freedman, Sagi Schein, Ruth Bergman

1 Technion - IIT, Israel, 2 IBM Research Labs, Haifa, Israel, 3 HP Labs, Israel
1 nastya@cs.technion.ac.il, 2,3 {pavel.kisilev, daniel.freedman} @gmail.com, 4,5 {sagi.schein, ruth.bergman} @hp.com

PROBLEM DESCRIPTION

GUI object classification is essential for image-based software automation tools. As such, it has to be robust to certain types of transformations:

- color,
- size / ratio of height and width,
- text or image,
- inaccurate segmentation;
- change of OS or “skin”;
- etc.

Five types of pre-segmented objects used in our experiments (snapshots from applications running on Windows Vista operating system.)

Existing GUI object classification techniques include:

- Template matching.
- Classification with standard features (SIFT, SURF, etc.)

The above are not robust to GUI objects transformations, and may fail for small GUI objects.

SIFT features detected for several push button objects. No features were detected for smaller objects.

CONTRIBUTION

We present a novel image descriptor developed specifically for Graphical User Interface (GUI) objects.

- Robust to various typical changes in GUI object appearance.
- Exploits GUI object structure and outperforms existing image descriptors.

THE PROPOSED IMAGE DESCRIPTOR

Input: grayscale object image \( I(x,y) \)

1. Fourier-Mellin transform-based feature

1.1. Convert \( I(x,y) \) into two 1D images

\[
I^x(x,y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} I(u,v) \frac{\partial I}{\partial u} du \\
I^y(x,y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} I(u,v) \frac{\partial I}{\partial v} du 
\]

1.2. Apply 1D Fourier-Mellin transform, to eliminate shift \( x_0 \) and scaling \( a \).

\[
F_a(x) = F_{I^x}(x) \cdot D_a(x) \\
F_y(x) = F_{I^y}(x) \cdot D_a(x) \\
D_a(x) = \exp(i \alpha x + i \log a)
\]

2. Histogram of gradient directions

4-bin histogram of the gradient directions, \( \angle(\nabla I) \), wrapped to be in the range \([0, \pi)\).

3. Percentage of white pixels

\#pixels \((I > I_0)\) / \#pixels of \(I\)

\(I_0 = 245\) was used.

Concatenate the above into a single vector-valued GUI object descriptor.

GUI OBJECT CLASSIFICATION

- Evaluation dataset consisted of 258 pre-segmented GUI objects: check boxes, edit boxes, list box openers, push buttons and radio buttons (see illustration on the left).

- Classification was performed using Kernel-SVM with \( K(x_1, x_2) = \exp(-\gamma \|x_1 - x_2\|^2) \).

- The dataset was randomly divided into training and testing sets of varying sizes.

- The graph on the left shows percentages of correctly classified objects, each averaged over 10 independent experiments.

- Comparison to modified SIFT features:
  - Feature points - pixels with prominent gradient values.
  - SIFT features at fixed scale of \( 8 \times 8 \) pixels.
  - BOVW image descriptor.

Robustness to GUI object transformations

- Classifier was trained correctly segmented objects.
- All the transformed objects on the left were correctly classified.