RGBD Face Detection with Kinect Sensor

ZhongJie Bi
Outline

• The Existing State-of-the-art Face Detector
• Problems with this Face Detector
• Proposed solution to the problems
• Result and ongoing tasks
The Existing State-of-the-art Face Detector

- Uses OpenCV Cascade Classifier
- A Grey-scale Face Detector
- Based on ‘Robust Real-time Object Detection’ by Paul Viola & Michael Jones
Robust Real-time Object Detection

• Based on the value of simple features
  ➢ The sum of the pixels within the white rectangles are subtracted from the sum of pixels in the grey ones.
Robust Real-time Object Detection

• Based on the value of simple features
  ➢ The sum of the pixels within the white rectangles are subtracted from the sum of pixels in the grey ones.
Robust Real-time Object Detection

• Based on the value of simple features
  ➢ The sum of the pixels within the white rectangles are subtracted from the sum of pixels in the grey ones.
  ➢ Simple features as they are, there are still tens of thousands to evaluate in a particular detection window.
Robust Real-time Object Detection

• Based on the value of simple features
  ➢ The sum of the pixels within the white rectangles are subtracted from the sum of pixels in the grey ones.
  ➢ Simple features as they are, there are still tens of thousands to evaluate in a particular detection window.
  ➢ Question: How to evaluate these features fast?

A

B

C

D
Robust Real-time Object Detection

• How to evaluate simple features fast?
  ➢ The idea of **Integral Image**

\[
ii(x, y) = \sum_{x < x', y < y'} i(x', y')
\]

\[
s(x,y) = s(x,y-1) + i(x,y)
\]

\[
ii(x,y) = ii(x-1,y) + s(x,y)
\]
Robust Real-time Object Detection

• How to evaluate simple features fast?
  ➢ The idea of *Integral Image* – 3 operations per area sum
  \[ ii(x, y) = \sum_{x < x', y < y'} i(x', y') \]
  \[ D = ii(4) + ii(1) - ii(2) - ii(3) \]
Robust Real-time Object Detection

• Learning Classification Functions
  ➢ With the idea of integrated image, there are still tens of thousands of features to evaluate in each image sub-window, hence the computation of all features is computationally prohibitive.
  ➢ A combination of only a small number of these features can yield an effective classifier.
  ➢ Challenge: Find these discriminant features.
  ➢ Solution: A variant of AdaBoost for aggressive feature selection (The 1st and 2nd feature selected are intuitive.)
Robust Real-time Object Detection

- Speed-up through the Attentional Cascade
  - Simpler classifiers are used to reject the majority of sub-windows before more complex classifiers are called upon to achieve low false positive rates.
Robust Real-time Object Detection

• Speed-up through the Attentional Cascade
  ➢ Simpler classifiers are used to reject the majority of sub-windows before more complex classifiers are called upon to achieve low false positive rates.
  ➢ Cascade is trained like a decision tree.
  ➢ A well-trained cascade of classifiers can achieve good detection performance while eliminating the need for further processing of negative sub-windows.
The Existing State-of-the-art Face Detector

• OpenCV Application of Cascade Classifier
  ➢ Detection at multiple scales is achieved by scaling the detector.
  ➢ The amount of shift between subsequent sub-windows is determined by some constant number of pixels as well as the current scale.
  ➢ Multiple detections of a same face were combined based on overlapping bounding region.
Problems with this Face Detector

- **OpenCV Application of Cascade Classifier**
  - Detection at multiple scales is achieved by scaling the detector. – **Problem:** scale factor affects performance.
  - The amount of shift between subsequent sub-windows is determined by some constant number of pixels as well as the current scale.
  - Multiple detections of a same face were combined based on overlapping bounding region. – **Problem:** confirmation parameter affects performance.
Problem: False Negative
Proposed Solution

• Change parameters to increase sensitivity of the detector.
  ➢ At the cost of increasing false positive rate.
Proposed Solution
Proposed Solution

• Change parameters to increase sensitivity of the detector.
  ➢ At the cost of increasing false positive rate.
  ➢ How to decrease false negative while counteract the increase in false positive rate?
Proposed Solution

• Change parameters to increase sensitivity of the detector.

• Make use of color information from the RGB Sensor.
  
  ➢ Different skin color differ mostly in intensity, not RGB proportion.
  
  ➢ Normalized RGB can be used to distinguish skin (even of different skin color) from non-skin parts in an image:
    \[ r = \frac{R}{R+G+B} \] lies in the range of 0.32~0.50 for skins
    \[ g = \frac{G}{R+G+B} \] lies in the range of 0.27~0.48 for skins
Proposed Solution

Test Passed

\[ r = 0.33 \sim 0.35 \]
\[ g = 0.32 \sim 0.33 \]
Proposed Solution
Proposed Solution

• Change parameters to increase sensitivity of the detector.

• Make use of color information from the RGB sensor.

• Make use of depth information from the depth sensor.
  ➢ Face depth vary within a range of 50mm.
  ➢ Depth is discontinuous at the edge of a face.
Challenges of using Depth Info.

- Disparity caused by the distance of sensors.

Coordinates of the face area in the RGB image doesn’t correspond to the face area in the depth image. Also, disparity varies with depth: the further away, the smaller the disparity.
Challenges of using Depth Info.

- Disparity caused by the distance of sensors.

Solution: Extend the sub window \( \frac{1}{2} \) facesize to the left, 1 facesize to the bottom. The new sub-window is 3 times the size of the face area.
Challenges of using Depth Info.

• Disparity caused by the distance of sensors.
• Holes caused by shadows (Black Area).
Challenges of using Depth Info.

- Disparity caused by the distance of sensors.
- Holes caused by shadows (Black Area).

Solution: Regard holes (with value -1) as background (with value 4095).
Pruning false alarms based on depth information.

\[
\text{Proportion} = \frac{\# \text{ of Dots within min(depth) } \sim \text{ min } + 50\text{mm}}{\# \text{ of Dots in the extended subwindow}} \times 3
\]

- The proportion index must be within the range of 0.5~1. Any sub-window with proportion index less than 0.5 or larger than 1 (in the case of constant depth, the value is 3) will be rejected as false detection.
Proposed Solution

This area is probably too far away for sensor to get depth values other than its maximum 4095, and hence the constant depth value in this area.

Proportion=0.54 >0.5, Test Passed

Proportion=3 >1

Proportion=0.432 <0.5
Conclusion

• Change parameters to increase sensitivity of the detector \(\rightarrow\) less false negatives but more false alarms
• Prune false alarms based on additional information: skin color and depth
• Quantitative analysis will be elaborated in the technical report.
RGBD Face Detection with Kinect Sensor

ZhongJie Bi

Thank you.