# Gender Recognition from Body

*Liangliang Cao, Mert Dikmen, Yun Fu, and Thomas Huang*

*Dept. Electrical and Computer Engineering,*
*University of Illinois at Urbana-Champaign*

## Problem

***Human visual system can recognize the gender from body shape despite the color and shape varieties.***

***Can computer do that?***

### *Why are we interested in this problem:*

***Application Demands:***

- *security surveillance in building entrances and parking lots*
- *customer statistics collection in supermarkets, cafes, McDonald's*

***Why not gender from face:***

- *Sometimes it is intrusive to take face photos.*
- *When the resolution is too low, faces are hard to detect and track*
- *Faces are often occluded by glasses, masks and facial hair.*

***Why not gait:***

- *It is hard to capture and describe human gait if the background is cluttered.*
- *Professional gait captures are expensive.*

### *Why this problem is difficult?*

***Both females and males can dress the same color or not.***

***People of the same gender may choose clothes/ hairs of different styles.***

***The background of body images is cluttered .***

### *Why do we think this problem is still solvable?*

*Although not absolutely sure, Some parts of human body give hints on whether they are more likely to be female or male.*

- *For example, not all the females have long hair, but one person with long hair is more likely to be female than male.*
- *Similarly, a skirt is probably worn by a female, while a suit is more likely to be on a male.*

We can combine these hints and draw a more reliable conclusion!

## Database

***There are no public databases for gender from body.***

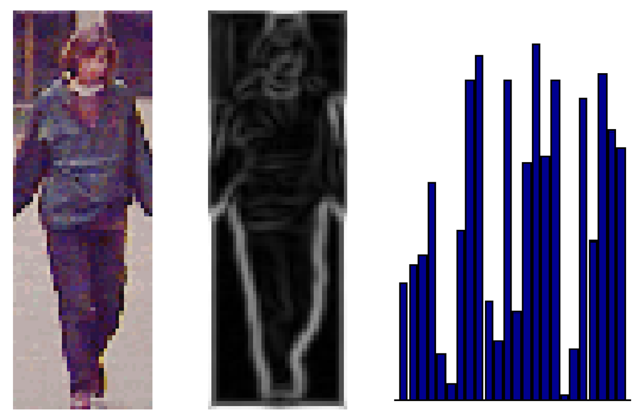
***We use pedestrian database and label them into male and female classes.***

- *Choose public MIT pedestrian database*
- *Labeled by the agreement of two researchers.*
- *Give up those on which we cannot agree or even cannot judge.*

### *Result in Database:*

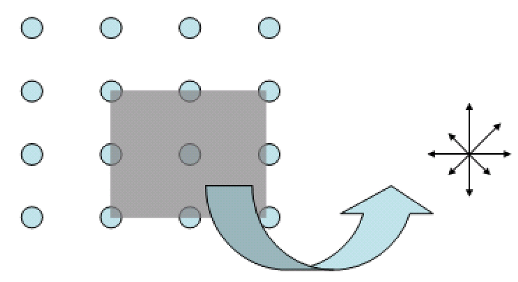
- *888 images in total.*
- *600 male + 288 female.*
- *53% back view images and 47% frontal view images.*

## The Feature



**(a)** **(b)** **(c)**

Different features: (a) raw pixels, (b) edge map, (c) histogram of oriented gradients



Patch-based representation by grid sampling

## The Algorithm

**Algorithm 1** : Adaboost Algorithm.

**Input:** $n_+$ male images and $n_-$ female images. The image labels $y_i$ is 1 for male, and $-1$ for female, and the image feature $\mathbf{x}_i = (x_i(1), x_i(2), ..., x_i(d))^T$, where $i$ is the image index with $1 \leq i \leq n_+ + n_-$ and $d$ is the feature dimension.

1: Initialize the weights of training samples $D_1(i) = 0.5/|n_+|$ if $y_i = 1$, or $D_1(i) = 0.5/|n_-|$, if $y_i = -1$.
2: **for** $t = 1, ..., T$ **do**
3: Based on $D_t$, select the weak classifier $h_t(\mathbf{x}_i) = h_t(x_i(k_t))$, where $k_t$ is the optimal feature position.
4: Compute the training error $\varepsilon_t$ of weak classifier $h_t$.
5: Compute $\alpha_t = 0.5 \log \frac{1-\varepsilon_t}{\varepsilon_t}$.
6: Update $D_{t+1}(i) = \frac{D_t(i)}{Z_t} \exp(-\alpha_t y_i h_t(\mathbf{x}_i))$, where $Z_t$ is a normalized factor.
7: **end for**

**Output:** The final classifier is $H(\mathbf{x}) = \text{sign}(\sum_{t=1}^{T} \alpha_t h_t(\mathbf{x}))$.

**Algorithm 2** : Random Forest Algorithm.

1: Draw bootstrap samples from the original data.
2: **for each** bootstrap samples **do**
3: Randomly draw a subset of all the predictors.
4: Using the selected subset, grow an unpruned classification tree.
5: **end for**

**Output:** The trained random forest is the majority voting of the decision trees.

**Algorithm 3** : Part-based Gender Recognition.

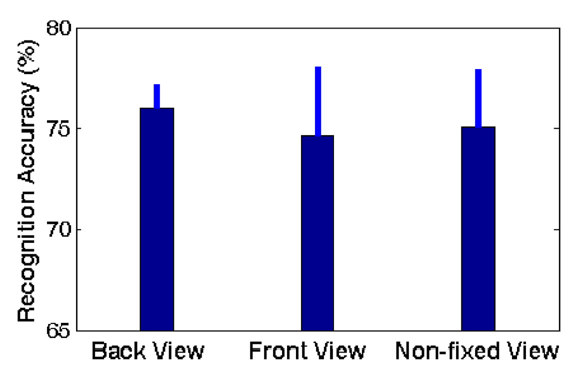
1: For each image $i$, extract part-based features $\mathbf{x}_i$, where $\mathbf{x}_i = \{x_i(p)\}$, with index $1 \leq p \leq P$.
2: **for** $t = 1, ..., T$ **do**
3: Select a body part $p_t$.
4: Get weak classifier $h_t(\mathbf{x}) = h_t(x(p_t))$.
5: Compute the voting weight $\alpha_t$ as Adaboost.
6: **end for**

**Output:** The final classifier is $H(I) = \text{sign}(\sum_{t=1}^{T} \alpha_t h_t)$.

## Result

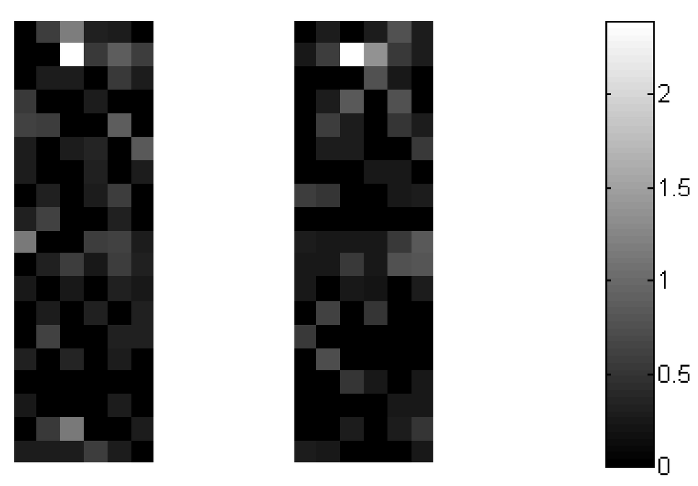
| Methods | Frontal View | Back View |
|---|---|---|
| Ada + Raw | $71.7 \pm 3.6\%$ | $61.5 \pm 4.3\%$ |
| RF + Raw | $72.8 \pm 3.7\%$ | $60.2 \pm 4.4\%$ |
| Ada + Edge Map | $71.1 \pm 3.6\%$ | $70.1 \pm 3.7\%$ |
| RF + Edge Map | $73.1 \pm 3.7\%$ | $63.9 \pm 4.4\%$ |
| Ada + HOG | $70.9 \pm 4.4\%$ | $63.0 \pm 4.1\%$ |
| RF + HOG | $73.1 \pm 3.5\%$ | $65.8 \pm 4.2\%$ |
| Ada + Part | $71.9 \pm 3.4\%$ | $71.2 \pm 3.3\%$ |
| RF + Part | $73.2 \pm 2.4\%$ | $65.4 \pm 0.0\%$ |
| PBGR | **$76.0 \pm 1.2\%$** | **$74.6 \pm 3.4\%$** |

Comparison of different features and algorithms. Each method is measured by the average accuracy and the standard deviation.



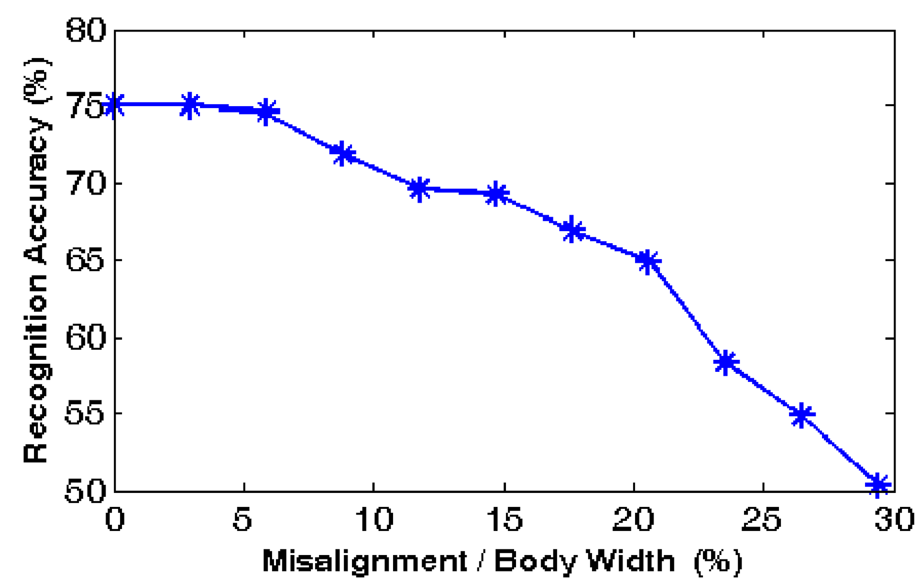

Gender recognition without knowing views. The thick bars stand for the average accuracy for each experiment, and the thin bars are the standard deviation of the 5-fold cross validation.

Recognition examples. Red and blue rectangles indicate the subjects are recognized as female and male, respectively. Our approach fails to recognize the last example probably because of the unusual pose of the subject.




**(a)** **(b)**

Voting weights for different patches. (a) weights from the frontal view classifier. (b) weights from the back view classifier.




Accuracy of gender recognition from misaligned images. Here a misaligned image is shifted horizontally with a distance so that the center of human body is no longer the center of the image.