

# Detecção de Pedestres com Oclusão

Silvio Gustavo de Oliveira Santos



UNIVERSIDADE  
FEDERAL  
DE PERNAMBUCO



Programa de Pós-graduação em Ciência da Computação

Defesa de Mestrado

Tsang Ing Ren (*Orientador*)

George D. C. Cavalcanti (*Co-Orientador*)

- Introdução
- Detecção de pedestres
- Método proposto
- Experimentos
- Considerações finais

# Introdução

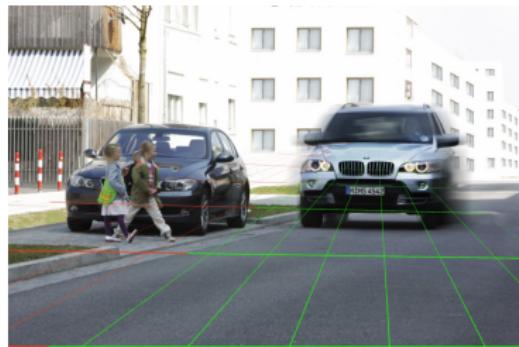
O que são pedestres?

São pessoas que se encontram em pé, podendo estar paradas ou andando



## Por que detectar pedestres?

Possui aplicações em sistemas de segurança de veículos, vigilância e na visão de robôs.



## Problemas

Apesar de grande avanço, os sistemas de detecção de pedestres sofrem alta degradação relacionada com:

- imagens de baixa resolução
- iluminação
- oclusão

# Introdução

- A oclusão ocorre quando partes do corpo do pedestre não estão visíveis
- Possui grande impacto sobre o resultado do sistema.



# Detecção de Pedestres



- Pré-processamento – ajusta e ressalta características
- Extração de características – captura informações que representam e distinguem um pedestre de outros objetos da cena
- Classificação – baseado nas características, classifica a imagem em pedestre ou não pedestre

# Detecção de Pedestres

## Pré-processamento

A etapa de pré-processamento consiste em converter as imagens em escalas de cinza

### Conversão para escalas de cinza

A conversão para escalas de cinza é realizada através de uma soma ponderada das componentes de cor RGB

$$Y = 0.299R + 0.587G + 0.114B$$

### Histogram of Oriented Gradients

Constrói um histograma a partir da magnitude e orientação dos vetores gradiente extraídos das bordas do pedestre

Os vetor gradiente é extraído utilizando os filtros  $G_x$  e  $G_y$ :

$$G_x = [-1 \quad 0 \quad 1]$$

$$G_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

# Detecção de Pedestres

## Extração de características - HOG

As componentes  $I_h$  e  $I_v$  do vetor gradiente são extraídas através da convolução dos filtros  $G_x$  e  $G_y$  com a imagem  $I$

$$I_h = I * G_x$$

$$I_v = I * G_y$$

A partir das componentes verticais e horizontais, obtemos a magnitude e a orientação do vetor gradiente

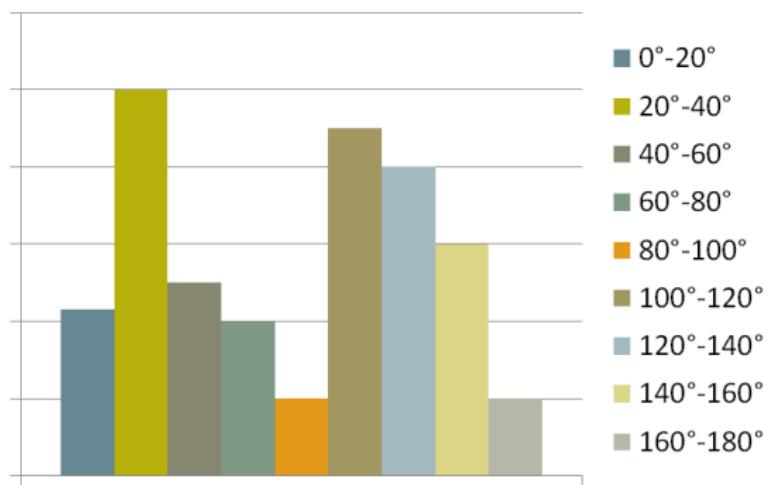
$$M = \sqrt{I_h^2 + I_v^2}$$

$$\theta = \arctan \frac{I_v}{I_h}$$

# Detecção de Pedestres

## Extração de características - HOG

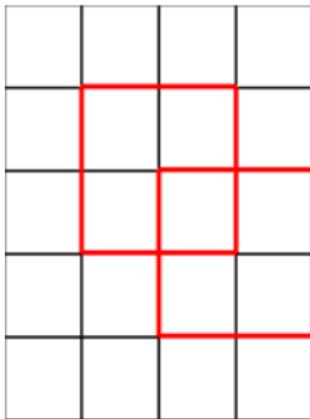
O histograma é então construído a partir da magnitude e orientação de cada vetor



# Detecção de Pedestres

## Extração de características - HOG

- O cálculo do histograma é realizado por célula e normalizado por bloco
- Cada célula corresponde a uma região de 8x8 pixels, formando blocos de 16x16 pixels
- O descritor resultante é formado por todos os histogramas de todos os blocos da imagem



### LBP

O LBP é uma técnica computacionalmente eficiente de extração de textura, sendo invariante à rotação e a mudanças monotônicas do nível de cinza da imagem

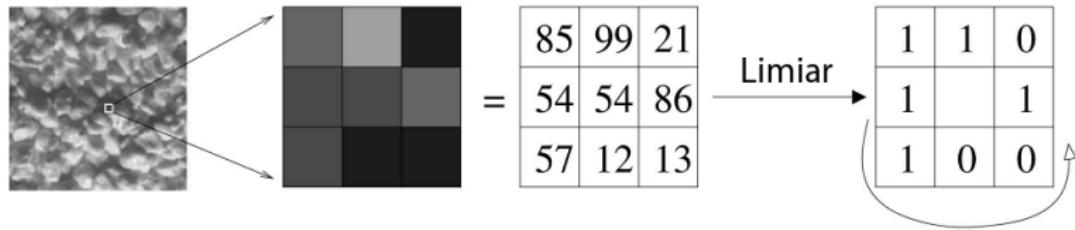
- Calculado através da diferença de cada pixel de uma região 3x3 com o pixel central
- Se a diferença é maior ou igual a 0, é atribuído o valor 1, caso contrário, atribui-se 0

# Detecção de Pedestres

## Extração de características - LBP

$$LBP = \sum_{i=1}^8 2^{(i-1)} f(I(g_i) - I(g_c))$$

$$f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & \text{caso contrário} \end{cases}$$



# Detecção de Pedestres

Extração de características - LBP



### HOG-LBP

O HOG-LBP combina características de gradiente do HOG com as de textura do *Linear Binary Patterns* (LBP)

- Para isso, gera um histograma a partir do LBP na forma de padrões uniformes
- Combina o histograma com o HOG de cada célula da imagem

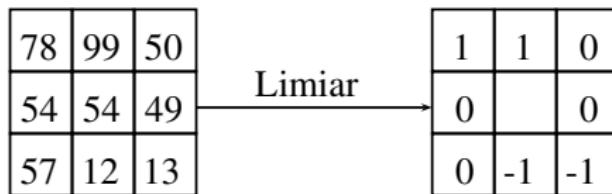
# Detecção de Pedestres

## Extração de características - LTP

### Linear Ternary Patterns

É um extrator de textura baseado no LBP, sendo mais discriminante e menos sensível a ruídos em regiões uniformes

No lugar de utilizar a intensidade do pixel central como limiar, adota um valor  $t$  que define um intervalo em torno do pixel central da região



Código ternário:  
1100(-1)(-1)00

$$[54 - t, 54 + t], t = 5$$

### Local Maximum Edge Binary Patterns

O LMEBP extrai informações baseado na distribuição das bordas em uma imagem e é calculado através da diferença entre um pixel da região  $3 \times 3$  e seus vizinhos, ordenando-os através do valor absoluto da diferença

# Detecção de Pedestres

## Extração de características - LMEBP

Imagem

2	5	3	9	1
6	7 <sub>(6)</sub>	6 <sub>(7)</sub>	1 <sub>(8)</sub>	5
2	3 <sub>(5)</sub>	6 <sub>(0)</sub>	8 <sub>(1)</sub>	2
3	8 <sub>(4)</sub>	4 <sub>(3)</sub>	2 <sub>(2)</sub>	9
1	2	3	2	1

Diferenças Locais

1	0	-5
<b>g</b>	<b>h</b>	<b>a</b>
-3		2
<b>c</b>		<b>f</b>
2	-2	-4
<b>e</b>	<b>d</b>	<b>b</b>

<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>e</b>	<b>f</b>	<b>g</b>	<b>h</b>
<b>-5</b>	<b>-4</b>	<b>-3</b>	<b>-2</b>	<b>2</b>	<b>2</b>	<b>1</b>	<b>0</b>
<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>

# Detecção de Pedestres

## Extração de características - LMEBP

	0	1	2	3	4	5	6	7	8	LMEBP	
a	0	0	1	1	0	1	0	0	1	105	1º LMEBP
b	0	0	1	1	0	1	0	1	1	107	2º LMEBP
c	0	0	1	1	0	1	0	0	1	105	3º LMEBP
d	0	0	1	0	0	1	0	0	1	073	4º LMEBP
e	1	0	1	0	0	1	0	1	1	331	5º LMEBP
f	1	0	0	0	0	1	0	1	1	267	6º LMEBP
g	1	0	1	0	0	0	0	0	1	321	7º LMEBP
h	1	1	1	0	0	1	0	1	1	459	8º LMEBP

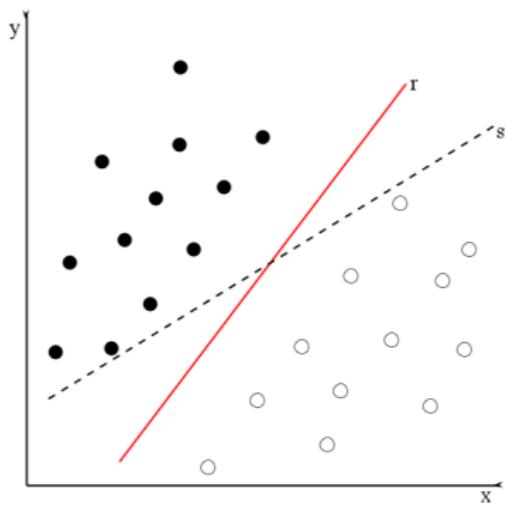
O descriptor é formado por um histograma que dá a frequência de cada código LMEBP na imagem

# Detecção de Pedestres

## Classificação

### Máquina de Vetores de Suporte (SVM)

É uma técnica supervisionada que encontra um hiperplano que separa as classes dos dados de treinamento com a maior margem possível



# Método proposto

Base de imagens com oclusão sintética

- Existem várias bases de imagens de pedestres, porém a taxa de acerto sobre grande parte delas é muito alta
- Por outro lado, quando apresentadas imagens com oclusão, as taxas de acerto dos classificadores treinados nessas bases caem consideravelmente
- A partir da base INRIA, criamos uma base com oclusão sintética, para estudar como os classificadores se comportam quando os pedestres sofrem oclusão de diferentes tamanhos e direções

# Método proposto

Base de imagens com oclusão sintética

- Formada mesclando imagens de pedestres da base INRIA com imagens de não pedestres
- Oclusões criadas nas direções ascendente, descendente e lateral (esquerda para a direita)



# Método proposto

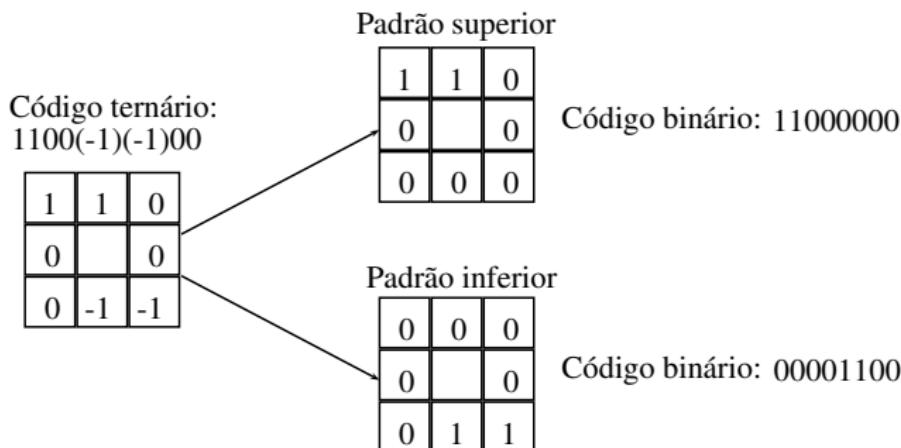
Base de imagens com oclusão sintética

- Foram formados 9 conjuntos com oclusão ascendente, 9 conjuntos com oclusão descendente e 4 conjuntos com oclusão lateral
- Cada conjunto possui tamanho de oclusão 8 pixels maior que os demais conjuntos da mesma direção
- Também foi criado um conjunto com 50% de imagens sem oclusão e 50% de imagens com oclusões de diferentes tamanhos



### HOG-LTP

Combinação das características HOG com as do LTP, de modo similar ao HOG-LBP. Para isso, o cálculo dos padrões uniformes foi realizado a partir das componentes superior e inferior do LTP

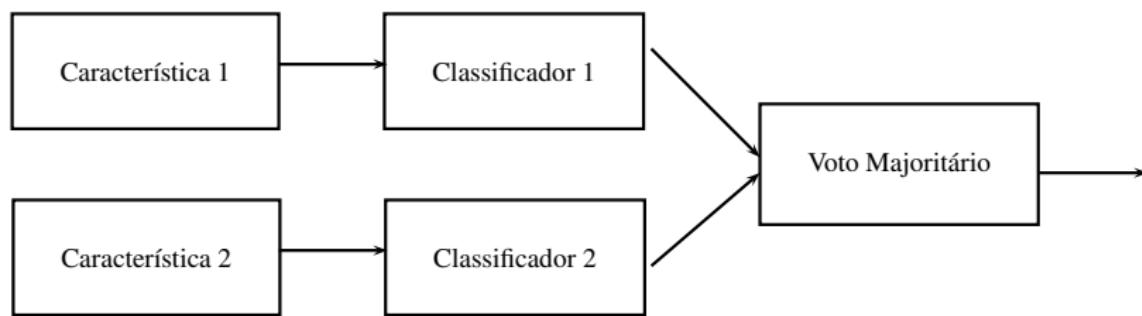


### HOG-LMEBP

Combinação das características HOG com as do LMEBP, através da concatenação dos histogramas do LMEBP com o HOG calculado para cada célula

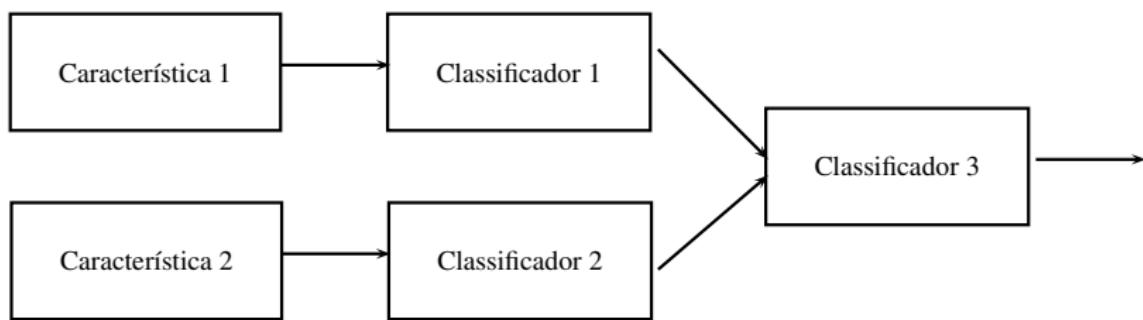
### Voto Majoritário

Os votos dos classificadores 1 e 2 são combinados e a classe que possuir mais votos é a escolhida. Caso os 2 classificadores discordem, a imagem é classificada como pedestre



### Classificador em cascata

Os níveis de confiança associados às características apresentadas aos classificadores 1 e 2 alimentam um 3º classificador responsável pela classificação



- Imagens de pedestre
  - 2272 imagens de treinamento
  - 1126 imagens de teste
- Imagens de não pedestre
  - 5000 imagens de treinamento
  - 4000 imagens de teste
- Classificador SVM linear

# Experimentos

## Combinação de características

Classificadores treinados e testados na base INRIA

Descriptor	Acurácia Pedestre (%)	Acurácia não Pedestre (%)
HOG	95,20	98,25
HOG-LBP	97,86	96,67
HOG-LTP	96,89	98,55
HOG-LMEBP	93,25	87,02

# Experimentos

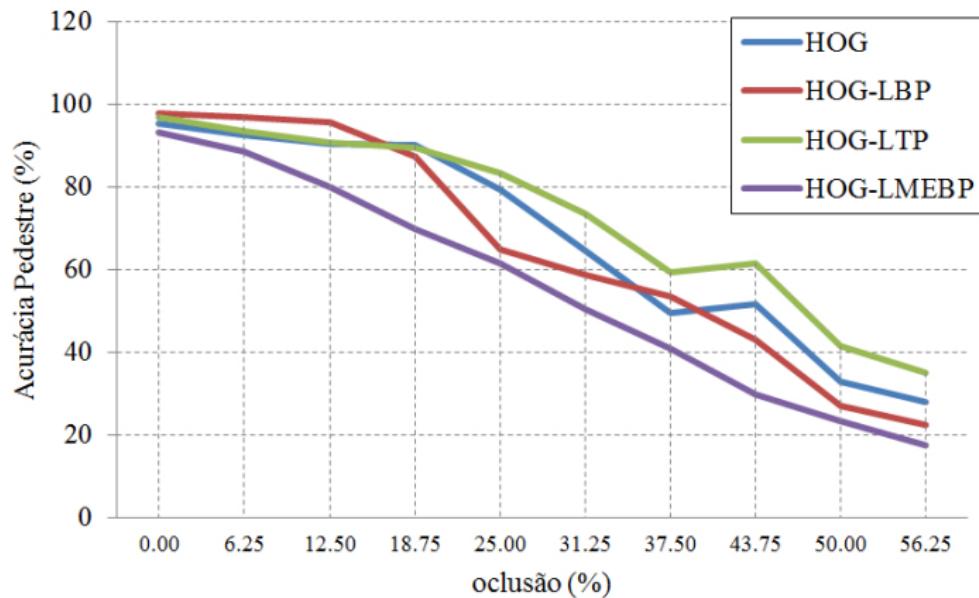
## Combinação de características

- Taxa de acerto elevada
- Imagens em condições ideais
- Como se comportam os classificadores quando os pedestres estão ocluídos?
  - Treinamento na base INRIA sem oclusão
  - Teste na base com oclusão sintética

# Experimentos

## Combinação de características

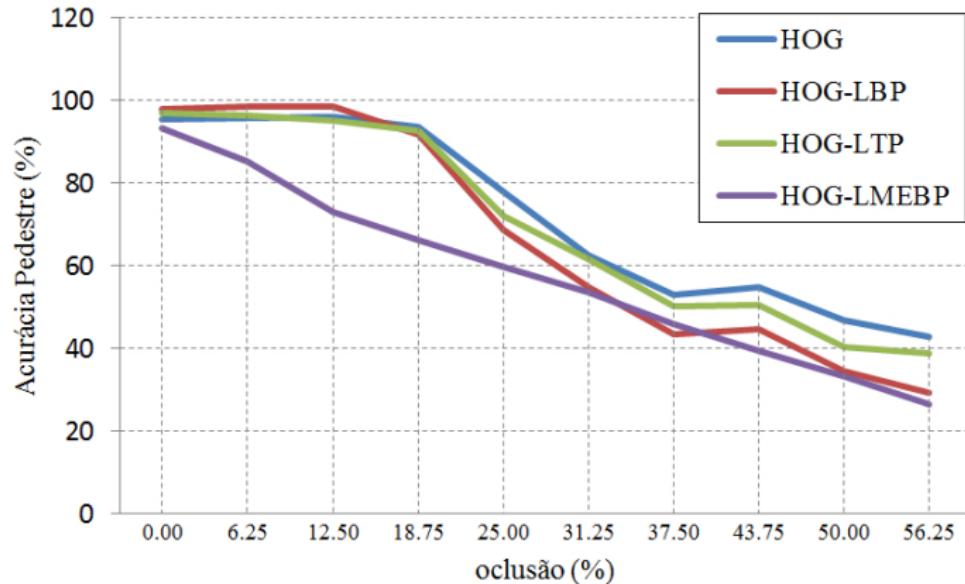
### Oclusão ascendente



# Experimentos

## Combinação de características

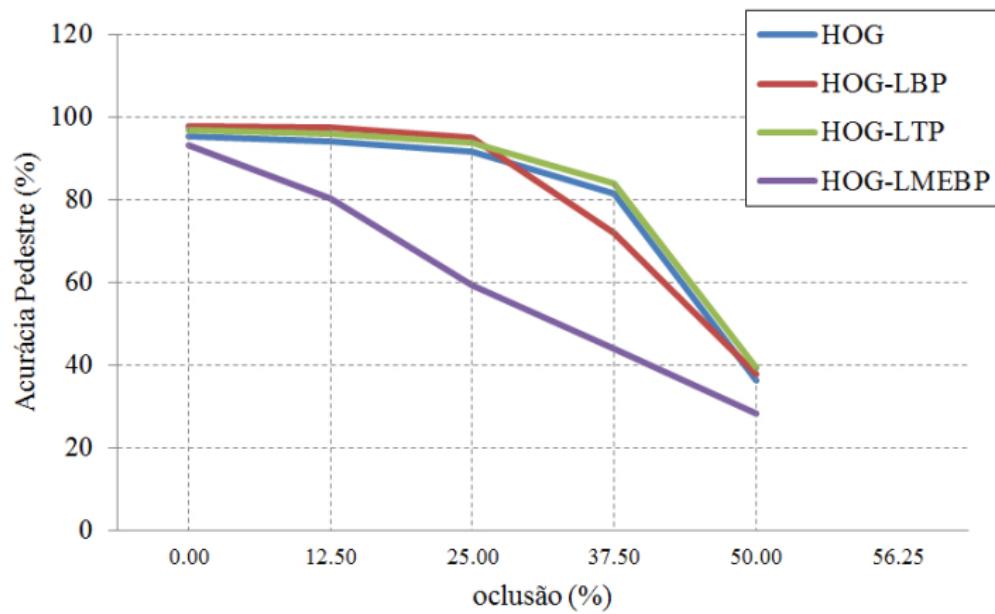
### Oclusão descendente



# Experimentos

## Combinação de características

### Oclusão lateral



# Experimentos

## Combinação de características

### Oclusão misturada

Tipo de Oclusão	HOG (%)	HOG-LBP (%)	HOG-LTP (%)	HOG-LMEBP (%)
Ascendente	80.63	80.81	83.21	70.33
Descendente	82.32	81.61	82.80	72.38
Lateral	85.34	85.87	88.27	72.02

# Experimentos

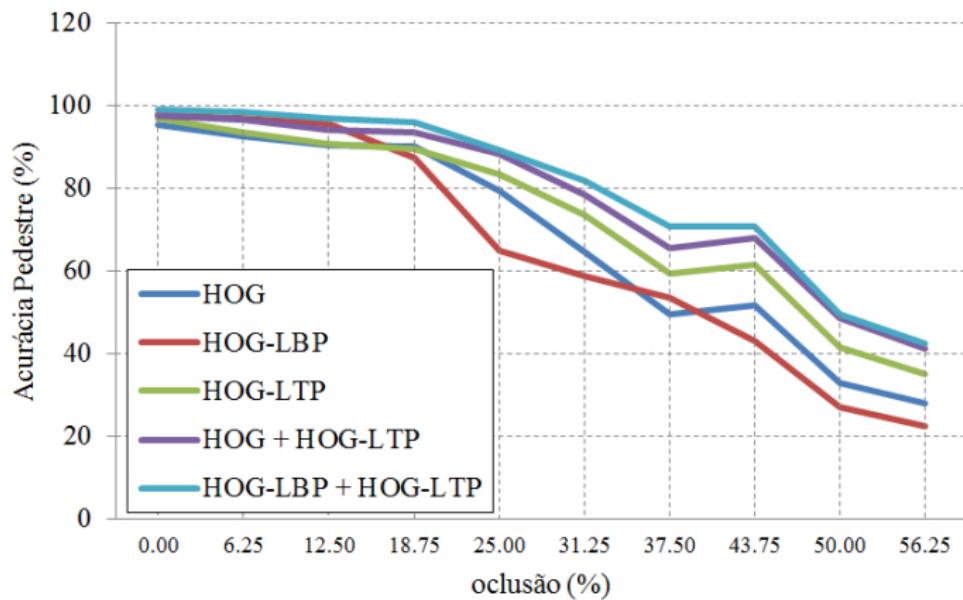
## Combinação de características

- Acurácia diminui a medida que oclusão aumenta
- Com grande área ocluída, o HOG e o HOG-LTP apresentam melhores resultados que o HOG-LBP
- Queda brusca na acurácia ao ocultar cabeça e ombros

# Experimentos

Combinação de classificadores - Voto majoritário

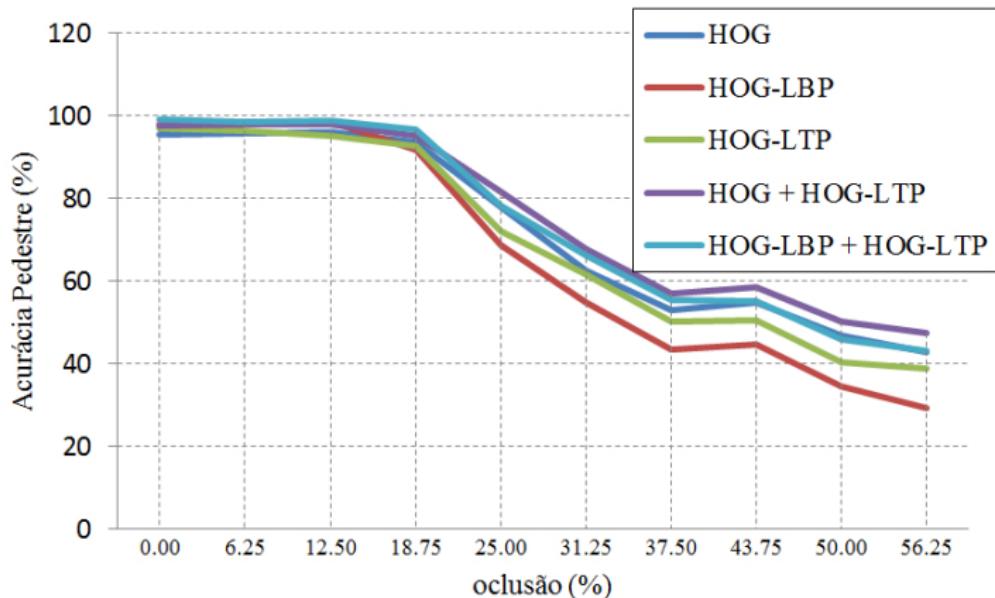
## Oclusão ascendente



# Experimentos

Combinação de classificadores - Voto majoritário

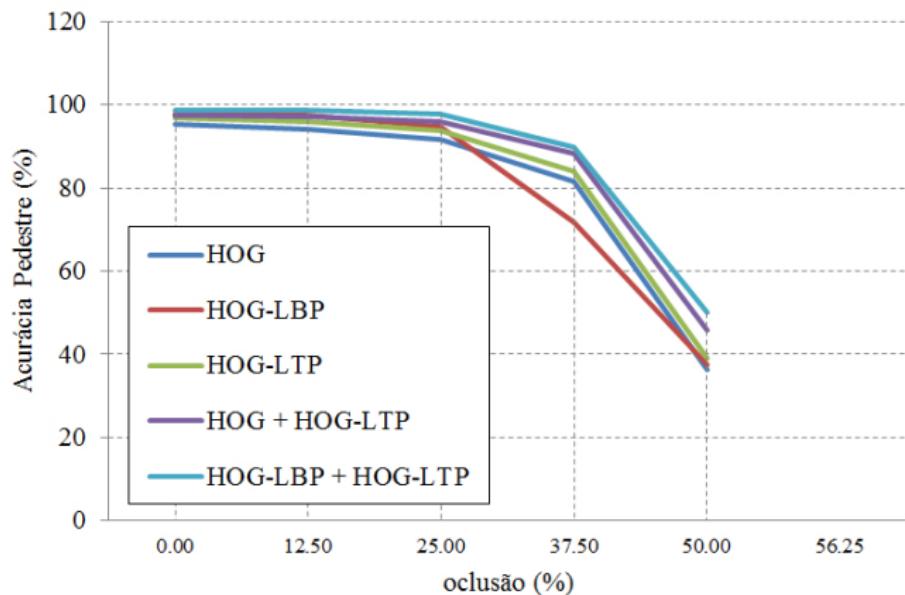
## Oclusão descendente



# Experimentos

Combinação de classificadores - Voto majoritário

## Oclusão lateral



# Experimentos

Combinação de classificadores – Voto majoritário

## Oclusão misturada

Tipo de Oclusão	HOG + HOG-LTP (%)	HOG-LBP + HOG-LTP (%)
Acendente	86.50	88.72
Descendente	85,79	86,15
Lateral	90.05	92.01

# Experimentos

## Combinação de classificadores – Voto majoritário

- Acurácia continuou diminuindo a medida que oclusão aumenta
- Os classificadores combinados apresentaram desempenho superior aos isolados
- As curvas de desempenho seguiram o mesmo padrão dos classificadores isolados

## Impacto sobre imagens de não pedestres

	HOG (%)	HOG-LBP (%)	HOG-LTP (%)	HOG-LMEBP (%)
Acurácia	98.25	96.67	98.55	87.02

	HOG + HOG-LTP (%)	HOG-LBP + HOG-LTP (%)
Acurácia	97.15	95.37

- A oclusão degrada bastante o resultado do sistema
- A cabeça e ombros contêm informações importantes para a detecção
- A simetria do pedestre garante baixa queda na acurácia devido à oclusão lateral
- As propostas do HOG-LTP e da combinação de classificadores por voto majoritário se mostraram menos prejudicadas pela oclusão

- Estudo de extratores de características para detecção de pedestres
- Combinação de características e classificadores
- Criação de base de imagens com oclusão sintética
- Avaliação do desempenho utilizando imagens com oclusão

Santos, S. G. O., Tsang, I. R., Cavalcanti, G. D. C., Jyh, I. R., Sijbers, J. (2013).

**Pedestrian Detection Under Progressive Occlusion**

*In: IEEE International Conference on System, Man and Cybernetics (SMC).*

# Detecção de Pedestres com Oclusão

Silvio Gustavo de Oliveira Santos



UNIVERSIDADE  
FEDERAL  
DE PERNAMBUCO



Programa de Pós-graduação em Ciência da Computação

Defesa de Mestrado

Tsang Ing Ren (*Orientador*)

George D. C. Cavalcanti (*Co-Orientador*)