

# Projeto Multiresolução de Operadores Morfológicos a Partir de Exemplos

Daniel André Vaquero  
Orientador: Junior Barrera

Departamento de Ciência da Computação  
Instituto de Matemática e Estatística (IME)  
Universidade de São Paulo (USP)

Defesa de Mestrado  
19 de Abril de 2006



# Roteiro

- 1 **Introdução**
  - Motivação
  - Projeto de W-operadores
  - Projeto Multiresolução
  - Objetivos
- 2 **Nossa abordagem**
  - Estimação piramidal da distribuição conjunta
  - Escolha da pirâmide
- 3 **Resultados Experimentais**
  - Classificação de dígitos manuscritos
- 4 **Conclusão**



# Roteiro

- 1 **Introdução**
  - **Motivação**
    - Projeto de W-operadores
    - Projeto Multiresolução
    - Objetivos
- 2 **Nossa abordagem**
  - Estimação piramidal da distribuição conjunta
  - Escolha da pirâmide
- 3 **Resultados Experimentais**
  - Classificação de dígitos manuscritos
- 4 **Conclusão**





# Roteiro

- 1 **Introdução**
  - Motivação
  - **Projeto de W-operadores**
  - Projeto Multiresolução
  - Objetivos
- 2 Nossa abordagem
  - Estimação piramidal da distribuição conjunta
  - Escolha da pirâmide
- 3 Resultados Experimentais
  - Classificação de dígitos manuscritos
- 4 Conclusão



# W-operadores

$$\psi : \{0, 1\}^3 \longrightarrow \{0, 1\}$$

padrão	saída



# W-operadores

 $\psi$ 








padrão	saída
	
	
	
	
	
	
	
	

Imagem de entrada

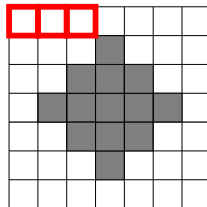
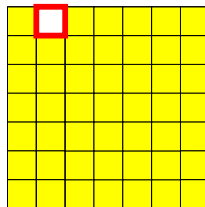


Imagem de saída



# W-operadores

 $\psi$ 





padrão	saída
	
	
	
	
	
	
	
	

Imagem de entrada

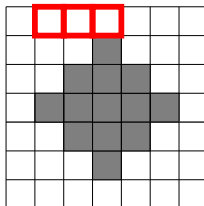
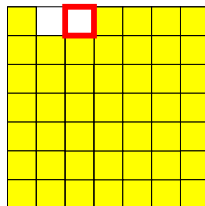


Imagem de saída





# W-operadores

 $\psi$ 

padrão	saída

Imagem de entrada

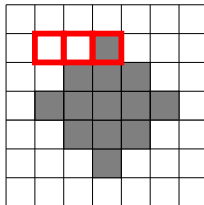
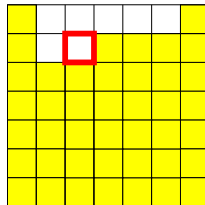


Imagem de saída



# W-operadores

 $\psi$ 

padrão	saída

Imagem de entrada

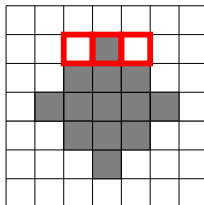
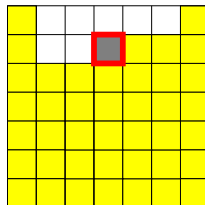


Imagem de saída



# W-operadores

 $\psi$ 

padrão	saída

Imagem de entrada

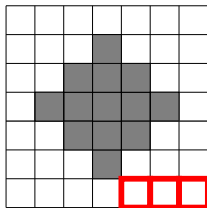
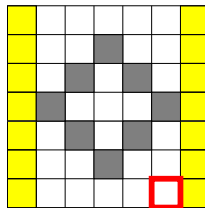


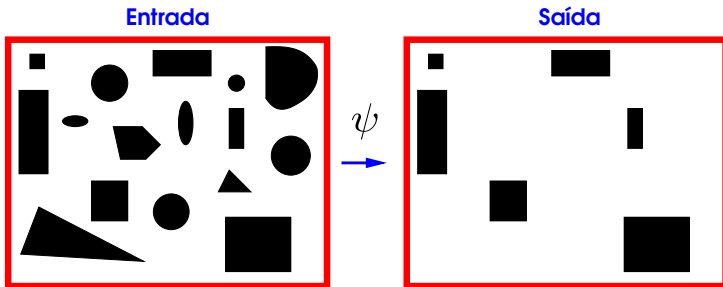
Imagem de saída



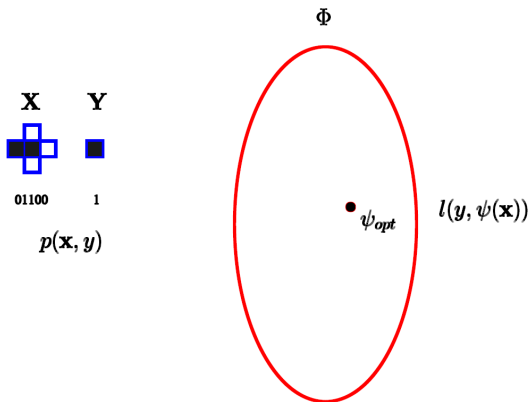
# Projeto a Partir de Exemplos

## Idéia

Aprender um operador a partir de **pares de imagens de entrada/saída**



# O processo de aprendizado



$$\text{minimizar } \sum_{\mathbf{x}, y} P(\mathbf{x}, y) l(y, \psi(\mathbf{x})) = \sum_{\mathbf{x}} P(\mathbf{x}) \sum_{y} P(y|\mathbf{x}) l(y, \psi(\mathbf{x}))$$



# O processo de aprendizado

## Estatísticas

padrão	saída	
	□	■
□ □ □	15	10
■ □ □	10	1
□ ■ □	0	9
□ □ ■	4	1
■ ■ □	1	10
■ □ ■	4	4
□ ■ ■	12	14
■ ■ ■	6	2

## Distrib. condicionais

padrão	P(Y X)	
	□	■
□ □ □	15/25	10/25
■ □ □	10/11	1/11
□ ■ □	0	1
□ □ ■	4/5	1/5
■ ■ □	1/11	10/11
■ □ ■	4/8	4/8
□ ■ ■	12/26	14/26
■ ■ ■	6/8	2/8

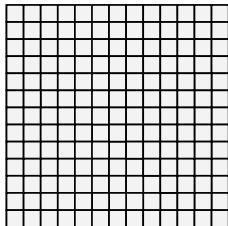
 $\hat{\psi}$ 

padrão	saída
□ □ □	□
■ □ □	□
□ ■ □	■
□ □ ■	□
■ ■ □	■
■ □ ■	□
□ ■ ■	■
■ ■ ■	□





# Problema com janelas grandes

W



Estadísticas



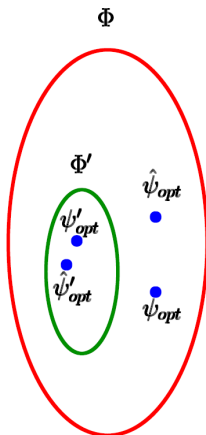
 = observados  
 = não observados

**Problemas:**

- Padrões não observados
- Representação



# Projeto utilizando restrições



Restrição no espaço de operadores

$$\Delta(\psi_{opt}, \psi'_{opt}) + E[\Delta(\hat{\psi}'_{opt}, \psi'_{opt})] \leq E[\Delta(\hat{\psi}_{opt}, \psi_{opt})]$$





# Projeto utilizando restrições

Restrição no conjunto de distribuições de probabilidade

- exemplo: gaussianas
- custo de restrição = 0; porém, modelos inadequados podem levar ao aumento do erro de estimação

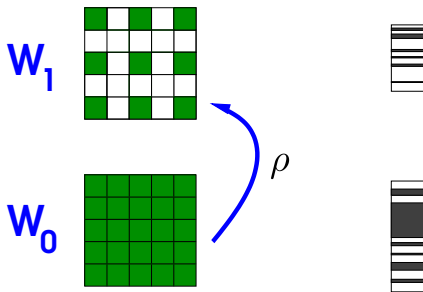


# Roteiro

- 1 **Introdução**
  - Motivação
  - Projeto de W-operadores
  - **Projeto Multiresolução**
  - Objetivos
- 2 Nossa abordagem
  - Estimação piramidal da distribuição conjunta
  - Escolha da pirâmide
- 3 Resultados Experimentais
  - Classificação de dígitos manuscritos
- 4 Conclusão



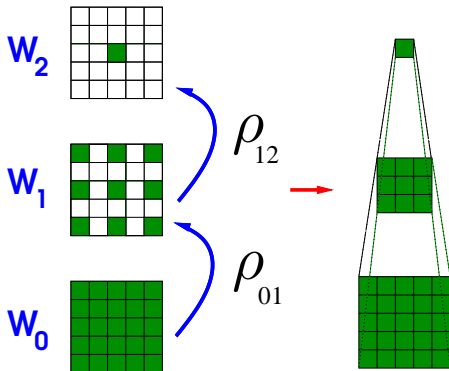
# Restrição de resolução



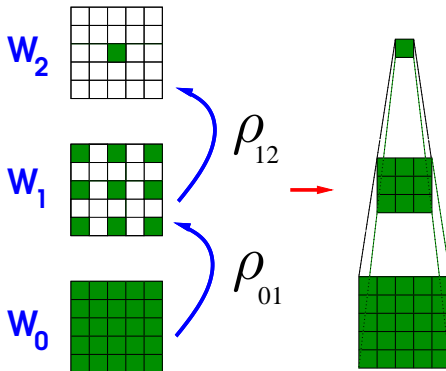
se  $\rho(\mathbf{x}) = \rho(\mathbf{x}')$  então  $\psi(\mathbf{x}) = \psi(\mathbf{x}')$



# Pirâmide de janelas



# Operadores multiresolução



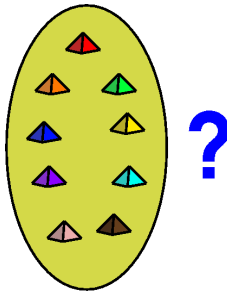
# Roteiro

- 1 **Introdução**
  - Motivação
  - Projeto de W-operadores
  - Projeto Multiresolução
  - **Objetivos**
- 2 Nossa abordagem
  - Estimação piramidal da distribuição conjunta
  - Escolha da pirâmide
- 3 Resultados Experimentais
  - Classificação de dígitos manuscritos
- 4 Conclusão

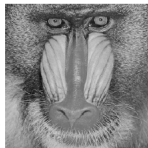
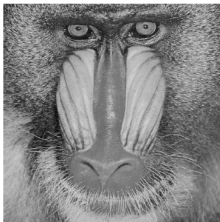


# Escolha da Pirâmide

Cada **pirâmide** induz um **W-operador**. Como escolher uma boa pirâmide?



# Pirâmides de Imagens





# Roteiro

- 1 Introdução
  - Motivação
  - Projeto de W-operadores
  - Projeto Multiresolução
  - Objetivos
- 2 **Nossa abordagem**
  - **Estimação piramidal da distribuição conjunta**
  - Escolha da pirâmide
- 3 Resultados Experimentais
  - Classificação de dígitos manuscritos
- 4 Conclusão



# Estimação Multiresolução

- uso de uma pirâmide para estimar a distribuição conjunta
- $p(Y|\mathbf{X})$  e  $p(\mathbf{X})$



# Estimação Multiresolução









## Estatísticas

padrão	saída	
	□	■
□ □ □	15	10
■ □ □	10	1
□ ■ □	0	0
□ □ ■	4	1
■ ■ □	0	0
■ □ ■	4	4
□ ■ ■	12	14
■ ■ ■	6	2



# Estimação Multiresolução

## Estatísticas

padrão	freqüência
	0
	1
	0
	4
	1
	5
	1
	2



# Estimação Multiresolução

Algoritmo de estimação

**Recebe:**  $\Delta$ ,  $\alpha$ , coleção de pares de exemplo

**Devolve:** Partição  $\mathcal{X} = \{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_n\}$  do espaço de configurações, que representa a distribuição conjunta

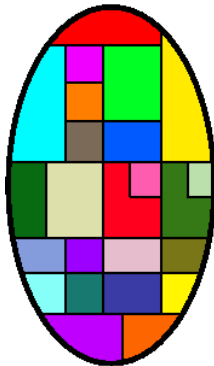
$$\forall \mathbf{x}_i \in \mathcal{X}_i, \quad P(Y|\mathbf{x}_i) = P(Y|\mathcal{X}_i)$$

e

$$\Gamma(\mathcal{X}_i) = \sum_{\mathbf{x} \in \mathcal{X}_i} P(\mathbf{x})$$



# Estimação Multiresolução

 $W_0$ 

# Estimação Multiresolução



100 ●	000
010	110 ● ● ● ■
001 ■	101 ● ● ● ■ ■
011 ●	111 ■ ■

$\mathcal{X}^0$



$\mathcal{X}$

$\alpha = 2$



# Estimação Multiresolução



100 ●	000
010	110 ● ● ● ■
001 ■	101 ● ● ● ■ ■
011 ●	111 ■ ■

$x^0$

110 ● ● $\frac{4}{14}$ ● ■ $\frac{1}{14}$
101 ● ● $\frac{5}{14}$ ● ■ $\frac{1}{14}$
111 ■ $\frac{2}{14}$ ■ $\frac{1}{14}$

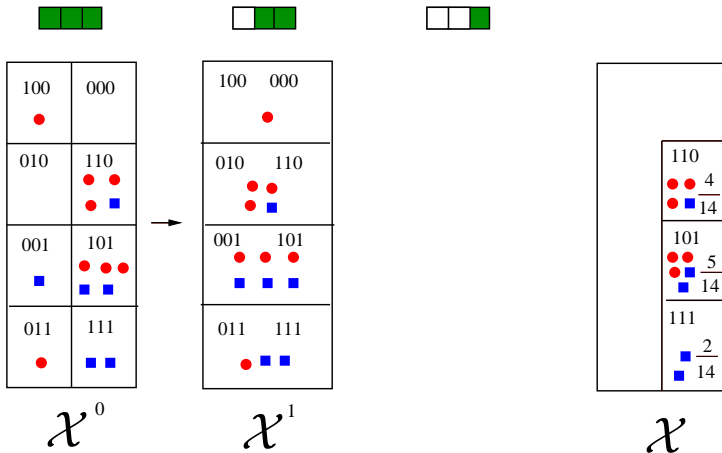
$x$

$\alpha = 2$





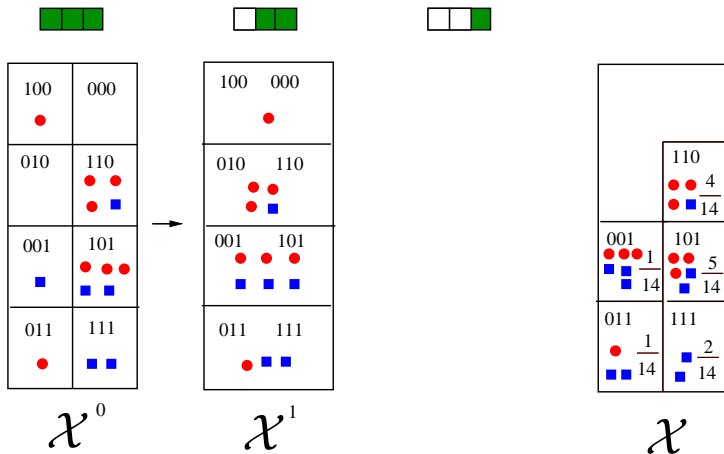
# Estimação Multiresolução



$$\alpha = 2$$

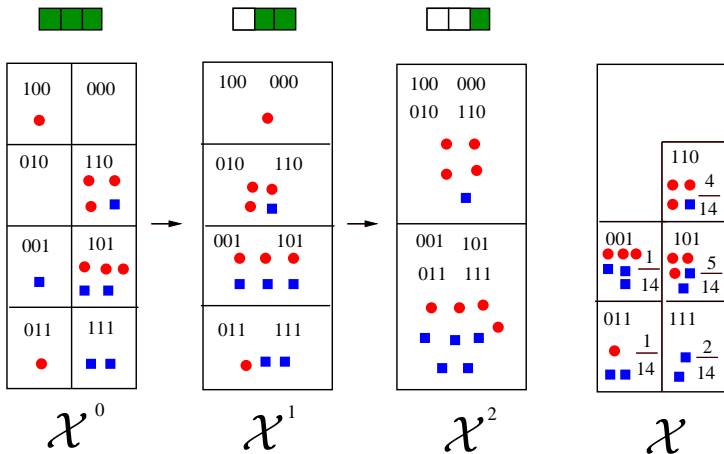


# Estimação Multiresolução



$$\alpha = 2$$

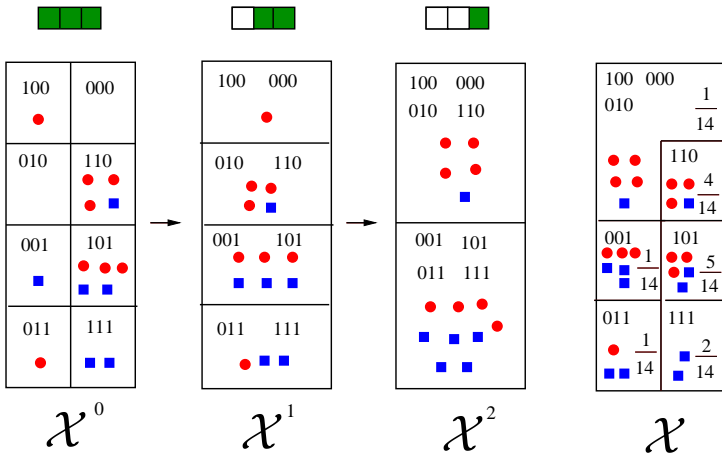
# Estimação Multiresolução



$$\alpha = 2$$



















# Estimação Multiresolução



$\alpha = 2$



# Estimação Multiresolução

100 000 010	$\frac{1}{14}$								
	<table border="1"> <tr> <td>110</td> <td><math>\frac{4}{14}</math></td> </tr> <tr> <td></td> <td><math>\frac{1}{14}</math></td> </tr> </table>	110	$\frac{4}{14}$		$\frac{1}{14}$				
110	$\frac{4}{14}$								
	$\frac{1}{14}$								
<table border="1"> <tr> <td>001</td> <td><math>\frac{1}{14}</math></td> </tr> <tr> <td></td> <td><math>\frac{1}{14}</math></td> </tr> </table>	001	$\frac{1}{14}$		$\frac{1}{14}$	<table border="1"> <tr> <td>101</td> <td><math>\frac{5}{14}</math></td> </tr> <tr> <td></td> <td><math>\frac{1}{14}</math></td> </tr> </table>	101	$\frac{5}{14}$		$\frac{1}{14}$
001	$\frac{1}{14}$								
	$\frac{1}{14}$								
101	$\frac{5}{14}$								
	$\frac{1}{14}$								
<table border="1"> <tr> <td>011</td> <td><math>\frac{1}{14}</math></td> </tr> <tr> <td></td> <td><math>\frac{1}{14}</math></td> </tr> </table>	011	$\frac{1}{14}$		$\frac{1}{14}$	<table border="1"> <tr> <td>111</td> <td><math>\frac{2}{14}</math></td> </tr> <tr> <td></td> <td><math>\frac{1}{14}</math></td> </tr> </table>	111	$\frac{2}{14}$		$\frac{1}{14}$
011	$\frac{1}{14}$								
	$\frac{1}{14}$								
111	$\frac{2}{14}$								
	$\frac{1}{14}$								

$\mathcal{X}$

- Resultado: partição de  $\{0, 1\}^{W_0}$
- Representada por:
  - uma seqüência de tabelas
  - uma pirâmide de janelas



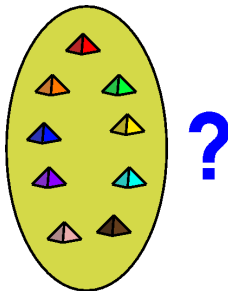
# Roteiro

- 1 Introdução
  - Motivação
  - Projeto de W-operadores
  - Projeto Multiresolução
  - Objetivos
- 2 **Nossa abordagem**
  - Estimação piramidal da distribuição conjunta
  - **Escolha da pirâmide**
- 3 Resultados Experimentais
  - Classificação de dígitos manuscritos
- 4 Conclusão



## Escolha da Pirâmide

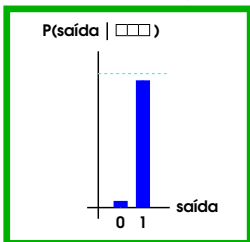
Cada **pirâmide** induz uma **distribuição condicional**, a partir da qual é projetado o **W-operador**. Como escolher uma boa pirâmide?



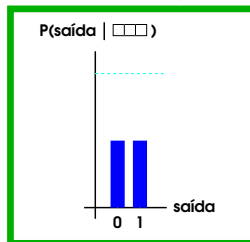
# Entropia

$$H(X) = - \sum_{x \in \mathcal{A}_X} P(x) \log_2(P(x))$$

## Baixa Entropia



## Alta Entropia



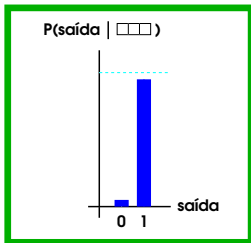


# Escolha da Pirâmide

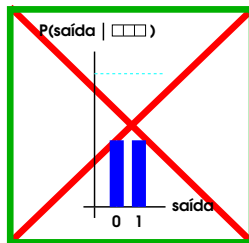
## Entropia condicional

$$H(Y|X) = \sum_{x \in \mathcal{A}_X} P(x) \cdot H(Y|x)$$







### Baixa Entropia



### Alta Entropia



# Cálculo da Entropia Condicional

100 000 010  $\frac{1}{14}$	110  $\frac{4}{14}$
001  $\frac{1}{14}$	101  $\frac{5}{14}$
011  $\frac{1}{14}$	111  $\frac{2}{14}$

 $\mathcal{X}$ 

$$\begin{aligned}
 H(Y|\mathbf{X}) &= \sum_{\mathbf{x} \in \mathcal{A}_{\mathbf{x}}} P(\mathbf{x}) \cdot H(Y|\mathbf{x}) \\
 &= \sum_{i=1}^n \sum_{\mathbf{x} \in \mathcal{X}_i} P(\mathbf{x}) \cdot H(Y|\mathbf{x}) \\
 &= \sum_{i=1}^n \left( \sum_{\mathbf{x} \in \mathcal{X}_i} P(\mathbf{x}) \right) \cdot H(Y|\mathcal{X}_i) \\
 &= \sum_{i=1}^n \Gamma(\mathcal{X}_i) \cdot H(Y|\mathcal{X}_i)
 \end{aligned}$$



# Roteiro

- 1 Introdução
  - Motivação
  - Projeto de W-operadores
  - Projeto Multiresolução
  - Objetivos
- 2 Nossa abordagem
  - Estimação piramidal da distribuição conjunta
  - Escolha da pirâmide
- 3 **Resultados Experimentais**
  - **Classificação de dígitos manuscritos**
- 4 Conclusão



# Classificação de dígitos manuscritos

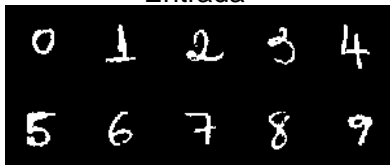


# Treinamento

Conjunto de treinamento

Aproximadamente 47000 dígitos

Entrada



Saída



$$\psi : \{0, 1\}^W \rightarrow \{0, \dots, 9\}$$



# Projeto do Operador

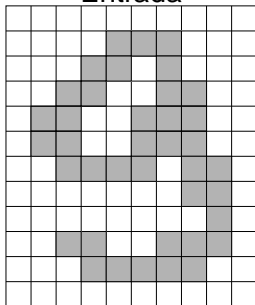
## Critérios de decisão

- Máxima verossimilhança: escolher  $y$  que maximiza  $P(\mathbf{x}|Y = y)$
- Moda: escolher  $y$  que maximiza  $P(y|\mathbf{x})$



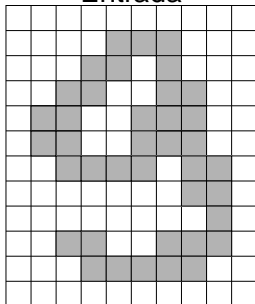
# Classificação

Entrada

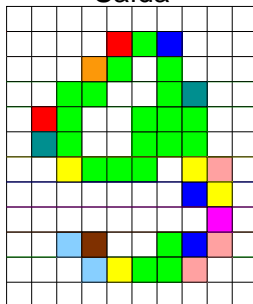


# Classificação

Entrada



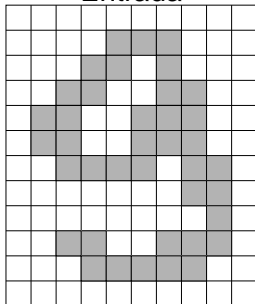
Saída



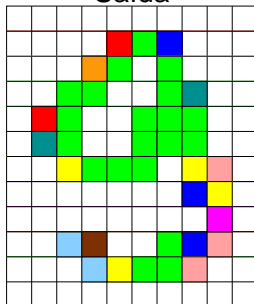


# Classificação

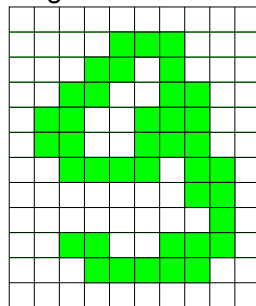
Entrada



Saída



Dígito classificado



# Resultados

## Entropias condicionais

Pirâmide	Entropia
14	0,276
11	0,317
10	0,441
13	0,477
5	0,480
12	0,527
6	0,543
1	0,572
9	0,580
8	0,600
2	0,611
3	0,652
7	0,680
4	0,807

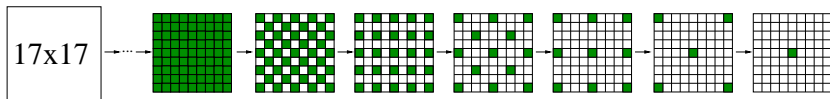
## Taxa de acerto (%)

Pirâmide	Acerto
14	89,4
11	88,6
10	85,6
13	74,7
12	73,8
8	71,5
7	70,4
9	68,8
6	68,4
3	54,9
5	54,8
2	51,7
1	51,7
4	50,5



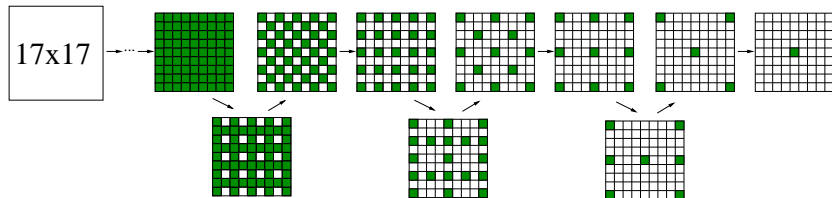
# Resultados

Melhor pirâmide (número 14): variação do esquema de amostragem **quincunx**



# Resultados

Melhor pirâmide (número 14): variação do esquema de amostragem **quincunx**



# Resultados

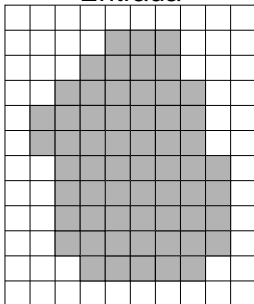
## Taxas de acerto (%) - Pirâmide 14

Método	$\alpha$	Acerto
<b>M.V.</b>	<b>2</b>	<b>89,4</b>
Moda	2	88,1
Dougherty et. al (2000)	2	88,0
Dougherty et. al (2000)	1	88,2

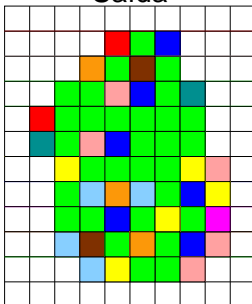


# Novo experimento

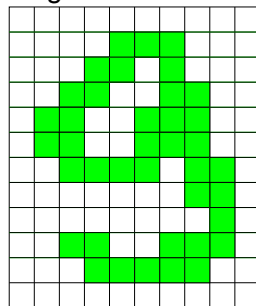
Entrada



Saída



Dígito classificado



# Resultados

## Novo experimento

- padrões no fecho convexo de cada dígito
- imagens em menor resolução
- dígitos normalizados quanto ao tamanho



# Resultados

## Entropias condicionais

Pirâmide	Entropia
14	0,187
11	0,217
13	0,263
12	0,300
5	0,350
8	0,386
6	0,393
7	0,427
3	0,431
10	0,444
1	0,467
2	0,475
9	0,493
4	0,543

## Taxa de acerto (%)

Pirâmide	Acerto
13	92,5
12	92,0
8	91,0
7	90,7
14	90,5
11	90,0
6	84,6
3	84,1
5	83,7
9	83,7
10	83,4
4	83,4
2	80,8
1	80,6





# Resultados

## Taxas de acerto (%) - Pirâmide 13

Método	$\alpha$	Acerto
<b>M.V.</b>	<b>2</b>	<b>92,5</b>
Moda	2	91,3
Dougherty et. al (2000)	2	91,2
Dougherty et. al (2000)	1	91,3



# Base de dados MNIST

## MNIST DB

- teste para avaliação de algoritmos de análise de formas
- conjunto de treinamento = 60000 dígitos
- conjunto de teste = 10000 dígitos
- dígitos normalizados quanto ao tamanho



# Base de dados MNIST

## Entropias condicionais

Pirâmide	Entropia
14	0,182
11	0,204
10	0,327
13	0,367
12	0,405
5	0,405
6	0,431
8	0,476
1	0,493
9	0,499
2	0,522
3	0,531
7	0,535
4	0,693

## Taxa de acerto (%)

Pirâmide	Acerto
14	96,71
11	96,51
10	95,21
13	92,60
12	92,18
8	91,79
7	91,47
6	90,75
9	90,67
5	84,84
3	84,19
1	82,05
2	81,76
4	75,94



# Resultados

## Taxas de acerto (%) - Pirâmide 14

Método	$\alpha$	Acerto
M.V.	2	96,71
Moda	2	96,16
Dougherty et. al (2000)	2	96,22
<b>Dougherty et. al (2000)</b>	<b>1</b>	<b>96,83</b>



# Resultados

## Comparação com outros métodos

Algoritmo	Acerto (%)
Rede neural convolucional LeNet-5	99,05
Rede neural convolucional LeNet-4	98,9
SVM, kernel Gaussiano	98,6
<i>k</i> -vizinhos mais próximos, L3 (Wilder)	97,17
Rede neural de 3 camadas, 500+150 unidades ocultas	97,05
Rede neural de 3 camadas, 300+100 unidades ocultas	96,95
<i>k</i> -vizinhos mais próximos, L2 (Wilder)	96,91
<b><i>W</i>-operadores multiresolução</b>	<b>96,83</b>
40 PCA + classificador quadrático	96,7
1000 RBF + classificador linear	96,4
Rede neural de 2 camadas, 1000 unidades ocultas	95,5
Rede neural de 2 camadas, 300 unidades ocultas, MSE	95,3
<i>k</i> -vizinhos mais próximos, L2 (LeCun)	95,0
Classificador linear	88,0



# Números variados de amostras



# Mapeamentos de Pirâmides de Imagens

- Diversos filtros
- Amostragem **quincunx**
- Esqueletos dos dígitos



# Conclusão

- Entropia condicional: foi eficaz na maioria dos experimentos
  - valor de saída
  - conjunto limitado de pirâmides
  - erros de estimação
- Boa capacidade de generalização ( $2^{23} \ll 2^{289}$ )
- Representação eficiente
- Pequenas variações nos erros obtidos para diferentes critérios
- Pirâmides de imagens: amostragem quincunx





# Contribuições

- Algoritmo de estimação da distribuição conjunta
- Uso da entropia condicional na escolha da pirâmide
- Técnica de busca por distribuições de massa concentrada
- Mapeamentos da teoria de pirâmides de imagens
- Aplicação ao reconhecimento de dígitos manuscritos



## Pesquisa Futura

- Combinar restrições no espaço de operadores com restrições no espaço de distribuições de probabilidade
- Aperfeiçoamento do critério de escolha da pirâmide
- Escolha da resolução em que é feita a estimação
- Parâmetro  $\alpha$ : determinação automática, valores diferentes em cada resolução



## Pesquisa Futura

- Outros tipos de mapeamentos de resolução (exemplo: *wavelets*)
- Seleção de características para escolher as janelas da pirâmide
- Aplicações: imagens em níveis de cinza, coloridas
- Refinamento do classificador de dígitos

