

Representação de Imagens Digitais através de Codificação de Vizinhança.



TIAGO BUARQUE ASSUNÇÃO DE CARVALHO

SAAP

ORIENTADOR: TSANG ING REN

**UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA**

RECIFE, 12 DE ABRIL DE 2012

TBAC@CIN.UFPE.BR

Roteiro



- Justificativa
- Revisão da Literatura
- Trabalho prévio
- Publicações
- Metodologia
- Objetivos
- Trabalho desenvolvido no doutorado
- Disciplinas
- Cronograma

Justificativa



- **Representação**
 - Estruturas de dados
- **Os problemas em processamento de imagens**
 - melhoramento, restauração, reconhecimento, análise multiresolução, operações morfológicas etc.
- **Uma representação pode simplificar um ou mais de um problema**
 - menor custo de memória e/ou processamento;
 - maior facilidade de implementação por parte do programador;
 - maior compatibilidade com determinado dispositivo.
 - Maior taxa de compressão e maior taxa de reconhecimento.

Introdução



- 1920, transmissão de imagens digitais
 - semanas → horas
- 1970, padrão digital para transmissão de fax
 - CCITT Group 3
- Imagens digitais
 - Vetorizadas
 - Bitmaps
- Bitmaps
 - Binário – 1 bit
 - Tons de cinza – escala de valores
 - Colorida – vetor RGB

Imagens Binárias



- Desenho
- Texto
- Halftone

Texto

Outro exemplo de imagem de texto:

Tiago B. A. de Carvalho.

Desenho



Tons de cinza



Halftone



Aplicações de Imagens Binárias



- Documentos digitais
 - Bibliotecas digitais
 - Google Books
 - ✦ 7 milhões de livros (2 milhões grátis)
 - ✦ 4 anos digitalizando livros impressos
 - Ebooks
 - ✦ Amazon Kindle, Sony Reader, iPad
- Transmissão de fax
- Biometria de impressões digitais
- Reconhecimento de caracteres manuscritos
- Indexação de formas em vídeo



Representação de Imagens Binárias



- Codificação de imagens
 - Compressão e descrição
- Sem perdas, manter os dados originais
 - Com perdas, sistema visual humano
- Tipos de representação
 - por bitmap
 - interna
 - de contorno

Métodos de Compressão

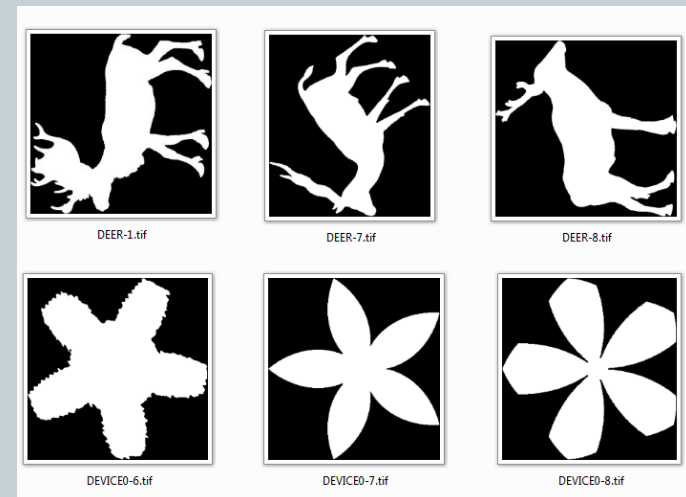
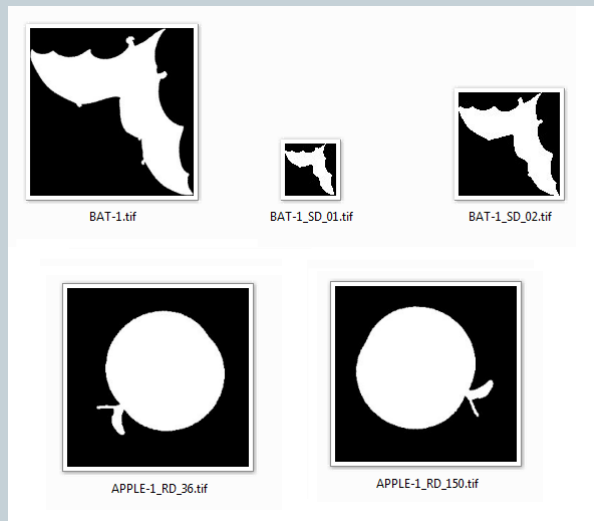


- Compressão sem perdas
- Imagem Binária
 - CCITT Group 3
 - CCITT Group 4
 - JBIG
 - JBIG2
- Também usados para imagens binárias
 - GIF
 - PNG

Reconhecimento de Forma



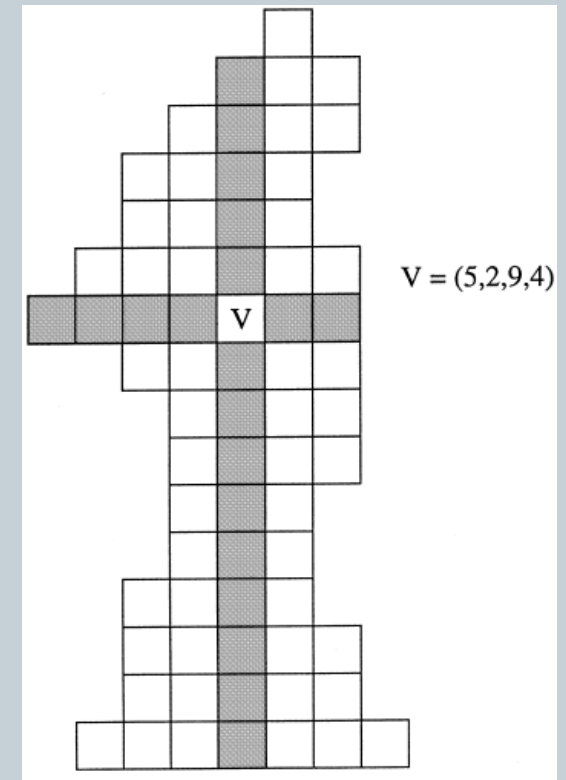
- Silhuetas de objetos bidimensionais
- Recuperação de imagens e indexação de vídeo
- MPEG-7 (descrição de conteúdo multimídia)
- Core Experiment CE-Shape-1 (A1, A2, B, C)



Codificação de Vizinhaça Original



- Tsang et al. 1999, 2006a, 2006b
 - Um vetor por pixel preto
 - Codificação do vetor em um número
 - Histograma de códigos
 - Extração de características
 - Caracteres manuscritos
 - Operadores de vizinhaça
 - Não é possível reconstruir a imagem



Codificação de Vizinhaça Proposta



- Código de Vizinhaça, XC
 - Centro (x,y) e Vizinhaça (n, l, s, o)
 - Permite a reconstrução de imagem original
- Exemplo (x, y, n, l, s, o)
 - $(2, 1, 0, 0, 3, 0)$
 - $(1, 2, 0, 2, 0, 0)$
 - $(2, 2, 1, 1, 2, 1)$
 - ...

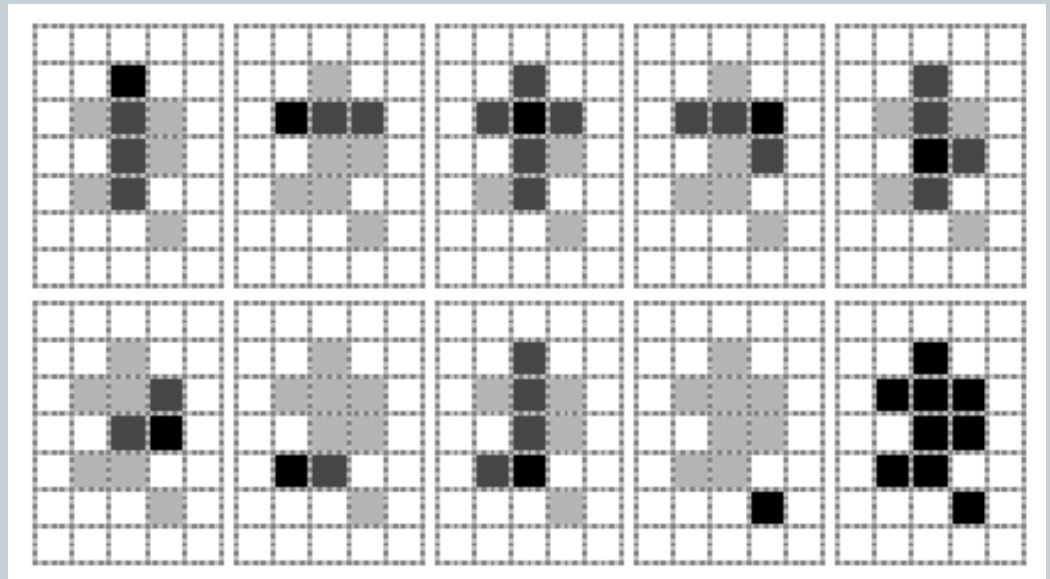


Imagem Binária



- Uma imagem binária é uma matriz P
 - $P = [p_{xy}]$
 - pixel p_{xy} em $\{0, 1\} = \{\text{preto, branco}\}$

$$P = \begin{bmatrix} p_{00} & p_{10} & p_{20} & \cdots & p_{(w-1)0} \\ p_{01} & p_{11} & p_{21} & \cdots & p_{(w-1)1} \\ p_{02} & p_{12} & p_{22} & \cdots & p_{(w-1)2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{0(h-1)} & p_{1(h-1)} & p_{2(h-1)} & \cdots & p_{(w-1)(h-1)} \end{bmatrix}$$

$$p_{xy} = \begin{cases} 1, & \text{se } (x,y) \in \Pi \\ 0, & \text{se } (x,y) \notin \Pi \end{cases}$$

$$\Pi = \{(x_1, y_1), (x_2, y_2), \dots, (x_b, y_b)\} \in \mathcal{P}(\mathbb{N} \times \mathbb{N}) \setminus \emptyset.$$

Funções Braço (b)



- Segunda parte do vetor XC
- Representam a vizinhança
 - Cada elemento do vetor vizinhança é um braço
- Associa o centro e o comprimento do braço a pontos

$$b(x, y, c) = \{(x_1, y_1), (x_2, y_2), \dots, (x_a, y_a)\} \in \wp(\mathbb{N} \times \mathbb{N}) \setminus \emptyset$$

$$B = \{b | b : \mathbb{N}^3 \rightarrow \wp(\mathbb{N} \times \mathbb{N}) \setminus \emptyset\}.$$

Código de Vizinhaça (XC)



- Vetor XC, fi

$$\varphi = (x, y, c_1, c_2, \dots, c_{d-2}) \in \mathbb{N}^d, d > 2.$$

- Configuração, beta

$$\beta = (b_1, b_2, \dots, b_{d-2}) \in \mathbb{B}^{d-2},$$

- Função vizinhaça do XC, u

$$u(\varphi, \beta) = \bigcup_{i=1}^{d-2} b_i(x, y, c_i),$$

De XCs para Imagem Binária



- Dimensões da imagem
- Configuração beta
- Conjunto de códigos XC, Fi
- Função vizinhança, v

$$v(\Phi, \beta) = \bigcup_{i=1}^m u(\varphi_i, \beta).$$

De Imagem Binária para XCs



- Definir um XC para cada pixel preto da imagem
 - Centro do XC = (x, y) , posição do pixel preto
 - Calcular o tamanho máximo para todos os braços, fazer com que cada braço codifique o máximo possível de pontos

$$\varphi_{xy} = \arg \max_{\varphi=(x,y,c_1,c_2,\dots,c_{d-2})} (|u(\varphi, \beta)|), u(\varphi, \beta) \subseteq \Pi$$

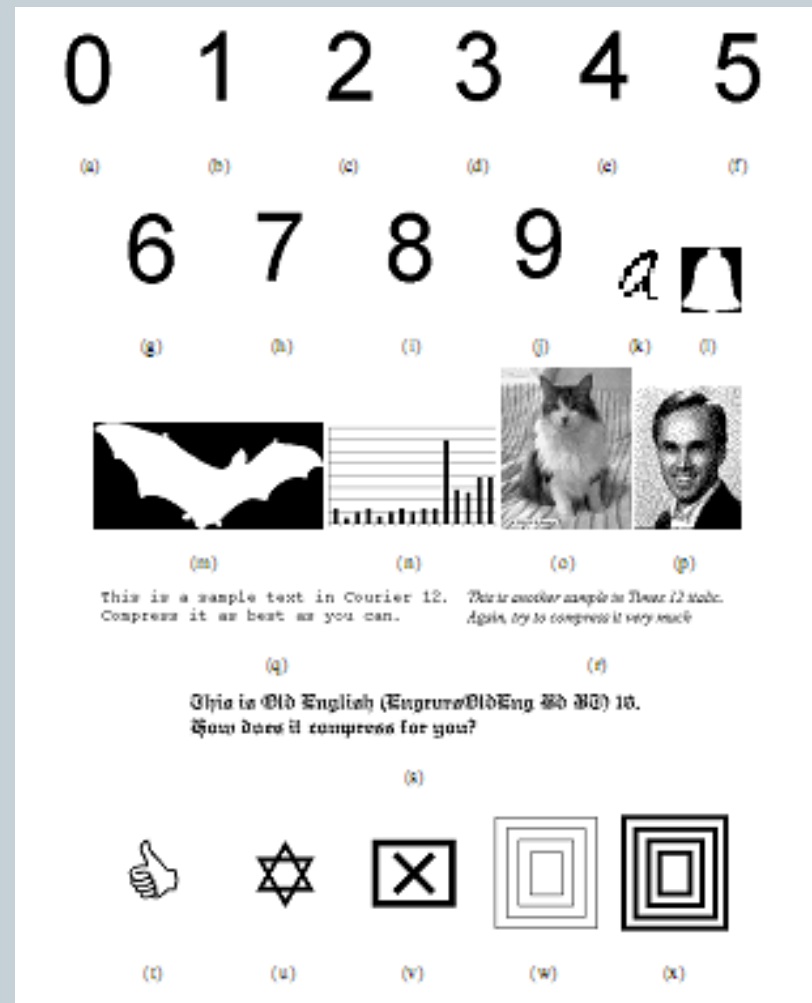
Algoritmo para a Redução de XCs



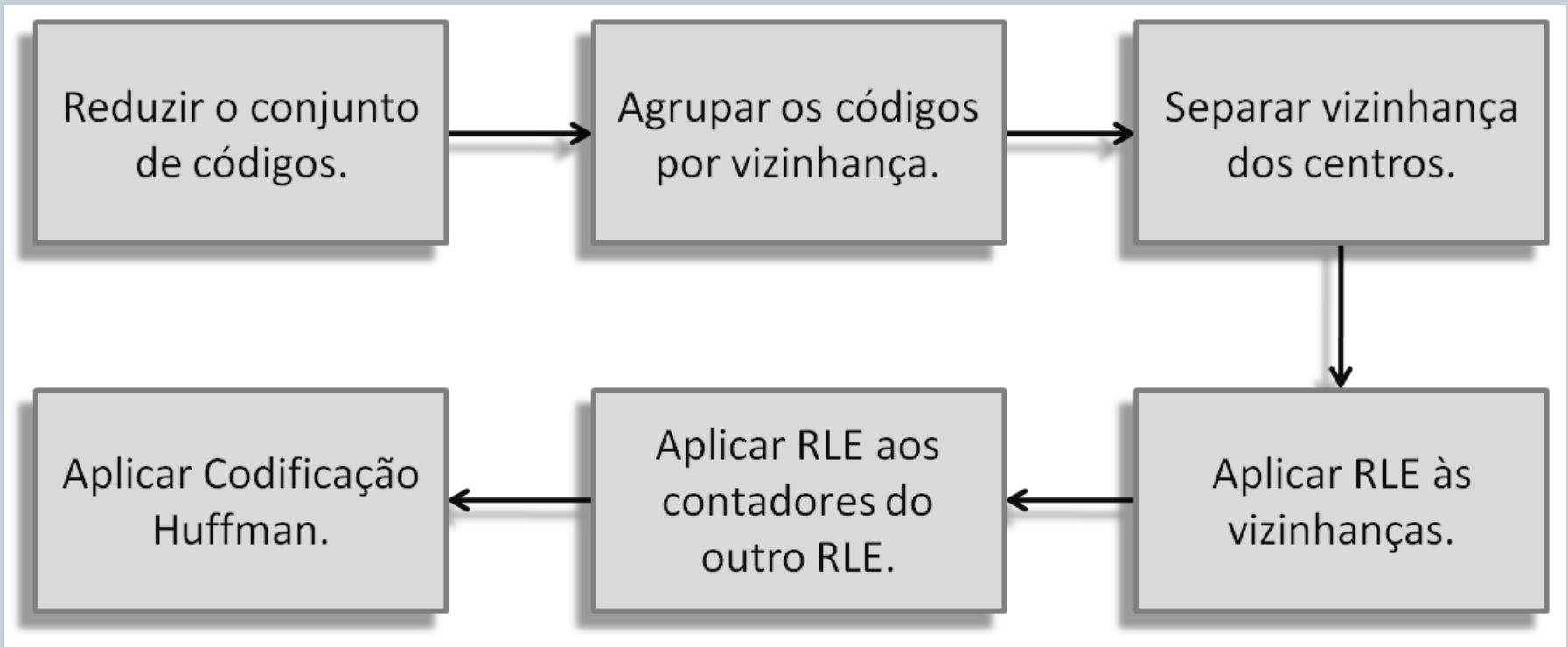
- Algoritmos para a redução do conjunto de códigos de vizinhança necessários para representar uma imagem binária sem perdas
- REDAG
 - Redução do Conjunto de Códigos Via Algoritmo Genético
- RED1
 - Primeiro Algoritmo Determinístico para Redução do Conjunto de Códigos
- RED2
 - Segundo Algoritmo Determinístico para Redução do Conjunto de Códigos

Experimentos – Imagens de Teste

- 24 imagens
- Desenho
 - 2 silhuetas
 - 10 Dígitos
 - 7 símbolos
- Texto
 - 3 fontes
- Halftone
 - 2 fotos



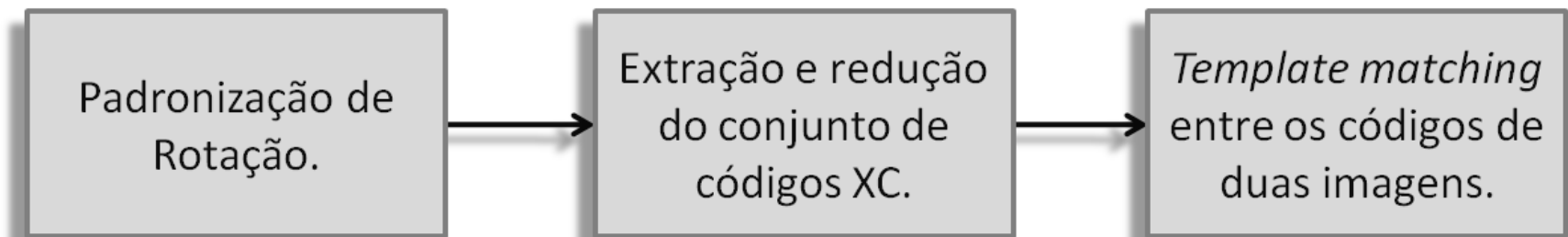
Compressão com XC



Reconhecimento de Forma com XC



- **Distância entre duas formas A e B**
 1. Padronizar a rotação das duas imagens A e B
 2. Gerar mais 3 imagens para B, rotacionado e girando 180°
 3. Gerar o conjunto de códigos reduzido para as 5 imagens
 4. Calcular a distância de A para cada versão de B
 5. Retornar a menor das 4 distâncias calculadas



Conclusões



- **Contribuições**
 - Formalização da codificação de vizinhança
 - Funções braço
 - Redução do conjunto de XCs
 - Compressão
 - Descritor de forma

Publicação do Trabalho



- No Mestrado

- Tiago Buarque Assunção de Carvalho, D. J. Tenório, I. R. Tsang, G. D. C. Cavalcanti, and I. J. Tsang. **Neighborhood coding for bilevel image compression and shape recognition.** In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2010

- Após o mestrado

- de Carvalho, T. B. A. ; TENORIO, D. J. ; CAVALCANTI, G. D. C. ; TSANG, I. R. . **Codificação de Vizinhança para Compressão de Imagens e Reconhecimento de Forma.** In: 23rd SIBGRAPI Conference on Graphics, Patterns and Images, 2010, Gramado - RS. 23rd SIBGRAPI Conference on Graphics, Patterns and Images, 2010.
- TENORIO, D. J. ; de Carvalho, T. B. A. ; CAVALCANTI, G. D. C. ; TSANG, I. R. . **Lossless Binary Image Compression Using Neighborhood Coding.** In: 23rd SIBGRAPI Conference on Graphics, Patterns and Images, 2010, Gramado - RS. 23rd SIBGRAPI Conference on Graphics, Patterns and Images, 2010.

Metodologia



- **Revisão bibliográfica**
 - Construção de survey
- **Pesquisa**
 - Proposta de soluções
 - Desenvolvimento das propostas
- **Publicações**
 - Periódicos

Periódicos e Conferências



- **Periódicos**

- IEEE Transactions on Image Processing (qualis A1),
- IEEE Transactions on Pattern Analysis and Machine Intelligence (qualis A1),
- Computer Vision and Image Understanding (qualis A1),
- **Pattern Recognition Letters (qualis A2)**,
- Image and Vision Computing (qualis A2) e
- ACM Computing Surveys (qualis A1).

- **Conferências**

- ICIP (qualis A2) IEEE International Conference on Image Processing,
- ICASSP (qualis B2) IEEE International Conference on Acoustics, Speech, and Signal,
- CIVR (qualis B2) ACM International Conference on Image and Video Retrieval e
- ICIAP (qualis B2) – International Conference on Image Analysis and Processing.

Objetivo Geral



- Aplicar a codificação de vizinhança para um amplo conjunto de problemas de Processamento de Imagem
- Demonstrar, através dessas aplicações, que a codificação de vizinhança é um método de representação útil
- Averiguar os demais métodos de representação para imagens binária (*Survey*)
- Possivelmente estender a codificação para outros domínios de imagens

Alguns objetivos específicos



- Aprofundar o que foi desenvolvido no doutorado
 - Complexidade dos algoritmos propostos
 - Novas funções braços -> maior compressão
 - ✦ Taxonomia de imagens -> Segmentação de *layout* de documentos
 - Método mais eficiente para redução dos códigos -> ótimo?
- *Survey* de representação de imagens binárias
 - ACM Computing Surveys (qualis A1)
- Explorar domínios de aplicação
 - Armazenamento de imagens de impressões digitais
 - E de mapas de relevância
- Operações morfológicas
- Etc.

Trabalho desenvolvido no doutorado



- Algoritmo para extrair os códigos que já entrega um conjunto reduzido de códigos
 - Qualidade do resultado da redução comparada com os algoritmos anteriores
 - Redução do custo de processamento em relação aos algoritmos anteriores de $O(n^2)$ para $O(n)$ -> falta provar
- Operações morfológicas
 - Borda, esqueleto etc.
 - Já existiam em trabalhos anteriores para o conjunto redundante de códigos mas não se aplicavam em conjuntos reduzidos de códigos
 - ✦ Dificuldades em trabalhar com a informação esparsa
 - ✦ Vantagem: menor custo computacional do que os métodos tradicionais de operações morfológicas em matrizes.

Disciplinas Cursadas



- **Dispensa de 3 disciplinas do Mestrado**
- **2011.1**
 - Processamento de Imagens
- **2011.2**
 - Trabalho individual
- **2012.1**
 - Aprendizagem de Máquina
- **2012.2**
 - Seminários

Cronograma



- **Março de 2012**
 - Cursar disciplinas.
 - Operadores morfológicos nos códigos reduzidos
- **Abril de 2012**
 - Cursar disciplinas.
 - Finalização de um artigo para *Pattern Recognition Letters*.
 - Submissão deste artigo.
- **Maio a setembro de 2012**
 - Cursar disciplinas.
 - Operadores morfológicos nos códigos reduzidos
 - Levantar artigos recentes e *surveys* relacionados ao tema e escrever um *survey* para *ACM Computing Surveys*.
- **Agosto de 2012**
 - Cursar disciplinas.
 - Submissão do *survey* para *ACM Computing Surveys*.
- **Setembro de 2012 a Fevereiro de 2013**
 - Cursar a disciplina de seminários
 - Desenvolvimento de técnicas de processamento de imagens utilizando codificação de vizinhança, visando mostrar a versatilidade deste método de representação de imagens.
 - ✦ Operadores morfológicos nos códigos reduzidos; rotação rápida