

# Sistemas de Visão e Percepção Industrial

## Introdução

- 1 A visão artificial como desafio
- 2 Obtenção de imagem
- 3 Representação da imagem digital

# Definições e conceitos gerais

- Processamento de imagem (Image Processing)
  - Transformação ou alteração de uma imagem (e.g. fotografia), em informação de outra natureza que pode não ser necessariamente outra imagem.
  - Inclui operações tão simples como mudar a intensidade global da imagem ou obter os contornos de objetos.
- Visão por computador (Computer Vision)
  - Conjunto de técnicas adicionais para extrair informação sobre a natureza 3D das cenas, como a reconstrução.
- Visão de máquinas (Machine Vision)
  - Inclui a aquisição da imagem e técnicas associadas, como a iluminação, e em aplicações concretas, como máquinas e robôs.
- Visão artificial (Artificial Vision)
  - Campo alargado que inclui todas as ciências e técnicas que permitem o estudo e aplicação de todas as atividades relacionadas com o uso de imagem, sua interpretação e reconstrução de cenas

# Níveis e etapas da visão por computador

## Baixo nível

- 1 Percepção (ou “sensação”)
  - Aquisição ou obtenção de imagem

- 2 Pré-processamento
    - Redução de ruído, realce de detalhes, filtragem, ...
- 

## Médio nível

- 3 Segmentação
    - Partição da imagem em zonas de interesse (objetos)
  - 4 Descrição
    - Determinação de propriedades de objetos (dimensões, forma, ...)
- 

## Alto nível

- 5 Reconhecimento
  - Identificação de objetos descritos a partir de uma base de dados.
- 6 Interpretação
  - Atribuir significados a um conjunto de objetos reconhecidos

# Importância e aplicações da VC

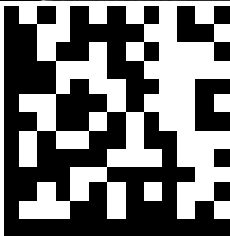
- A visão por computador, como método sensorial, assume um papel cada vez mais importante a vários níveis, entre os quais:
  - Processos industriais envolvendo montagem e/ou identificação de componentes
  - Controlo de qualidade e monitorização.
  - Vigilância e sistemas de segurança.
  - Reconhecimento de caracteres (OCR)
  - Percepção e navegação em robótica
- A VC permite, entre outros,
  - Controlo de qualidade com garantia e redução de custos
  - Automatização e simplificação de tarefas morosas
  - Integração com sistemas de gestão de produção

# Tipologia de aplicações para visão artificial na indústria

- Controlo de qualidade
  - Avaliação de conformidade (análise de cor, de dimensões, ...)
  - Diagnóstico automático (contagens, postura geométrica,...)
- Metrologia
  - Avaliação automática de distâncias, áreas e volumes.
- Monitorização
  - De espaços (segurança).
  - De processos tecnológicos (análise de deformações, geometrias)
- Em geral, nas linhas de montagem:
  - Verificação de componentes.
  - Seleção de componentes.
  - Detecção de geometria e dimensões.

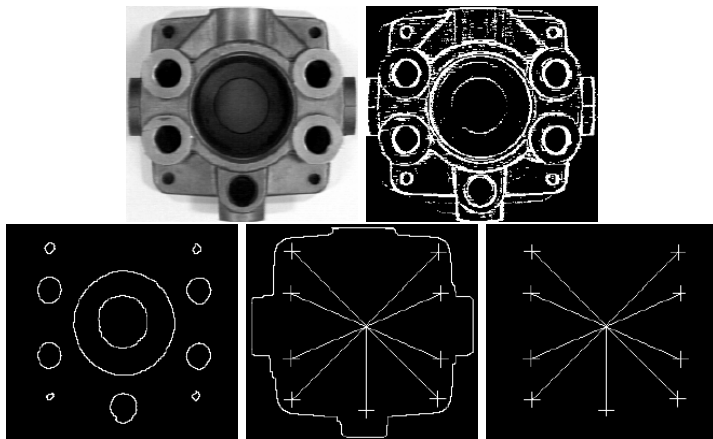
# Exemplos de aplicações – OCR e similares

- Reconhecimento de etiquetas em posições/orientações várias
- Leitura de Códigos Matriciais
- Leitura de Matrículas
- Leitura de algarismos e letras
- etc.



# Exemplo de tarefa em Visão Industrial

- Validação de um componente por determinados parâmetros físicos
  - Ex. Verificação de circunferências e seus centros



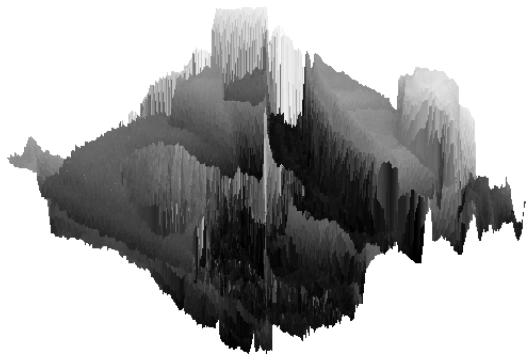


- O número de fornecedores é cada vez maior. Os mais tradicionais são, por exemplo,:
  - OMRON
    - <http://www.omron-industrial.com>
    - exemplo de um produto em <https://industrial.omron.eu/en/products/fq2>
  - EDMUNDOPTICS
    - <http://www.edmundoptics.com>
  - INFAIMON
    - <http://www.infaimon.com>
- Mas há um número crescente e qualquer lista será sempre limitada...
- Uma lista vasta encontra-se, por exemplo, em:
  - <http://www.lengrand.fr/computer-vision-companies>

# A visão artificial como desafio

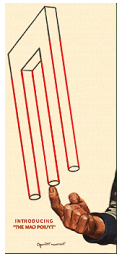
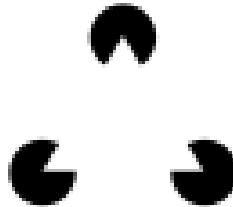
# Complexidade da VC... Porque é difícil?

- Duas representações da mesma imagem
- Ver exemplo em Matlab



# Até onde se poderá contar com a VC?

- Estamos habituados à percepção humana que,
  - muitas vezes, interpreta por processos contextuais.
  - Isso pode ser muito difícil numa máquina que tem um processamento bastante distinto dos sistemas biológicos!
- E as ilusões de ótica?
  - Como se poderão lidar?



- Se o significado de uma imagem depende do ponto de vista ou da orientação, como pode a VC determinar o significado?
  - Nunca o poderá fazer por si só...

## Outras ilusões e a complexidade da interpretação

- Os nossos olhos e cérebro “procuram” a solução mais “confortável” e interpretam erroneamente!



External player



## Como a orientação pode afetar a nossa percepção

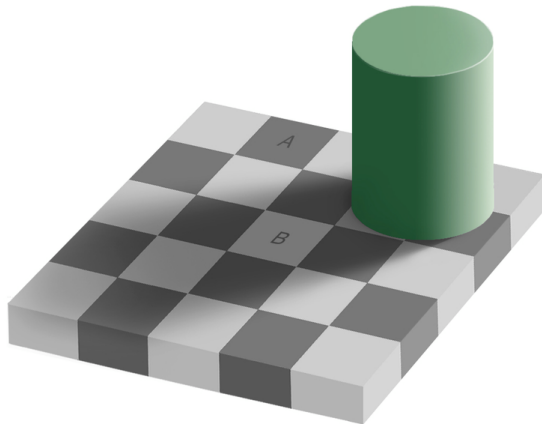


## O velho “jogo” do fundo e do objeto!



- É exatamente a mesma imagem... Rodada 180°!

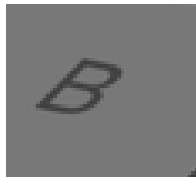
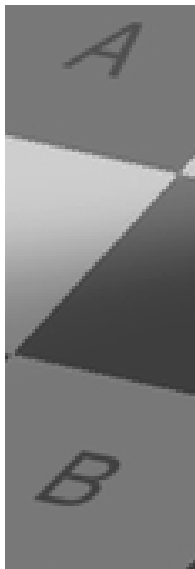
# Extremos da percepção contextual humana



- Os quadrados A e B têm o mesmo valor de intensidade!



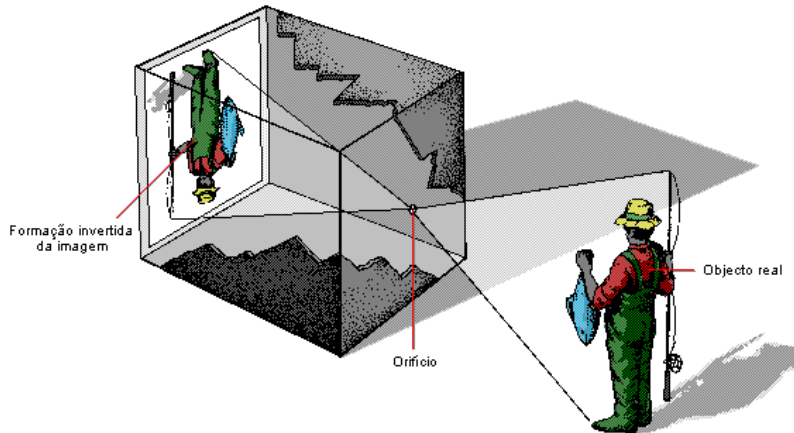
## Vista em detalhe das regiões A e B



# Obtenção de imagem

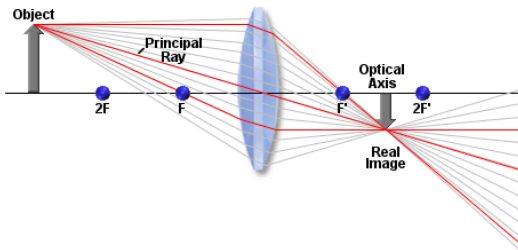
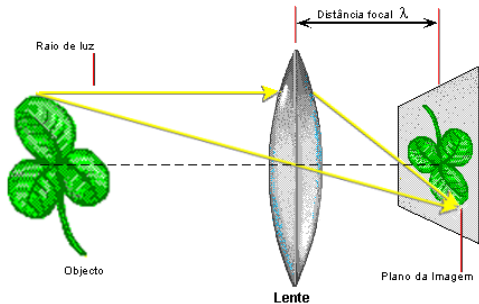
# Formação da imagem na “câmara escura”

- Modelo pin-hole – luz entra por um orifício.
- As dimensões e definição do orifício definem a nitidez da imagem:
  - Demasiado pequeno: deixa passar pouca luz.
  - Demasiado grande: retira a nitidez à imagem formada.



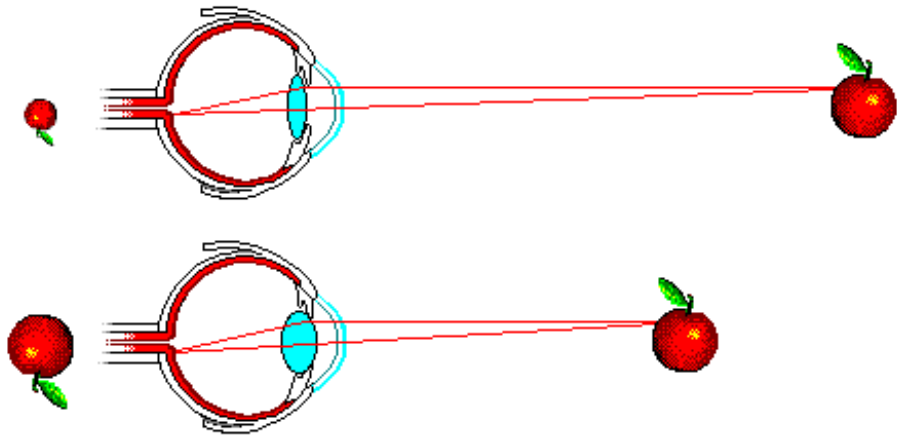
# Formação de imagem através de uma Lente

- Para ultrapassar o problema da pouca luz, usa-se uma abertura maior do orifício, mas recorre-se a uma lente para corrigir o trajeto da luz e dar nitidez à imagem
- Todos os raios de luz que saem de um dado ponto de um objeto na direção de uma lente convergente vão confluir, do outro lado, num outro ponto a uma certa distância do centro dessa lente: trata-se da distância focal.
- O conjunto de todos os pontos para os quais se verifica esse facto constitui o plano focal.
- Assim, a imagem observada nesse plano diz-se focada.

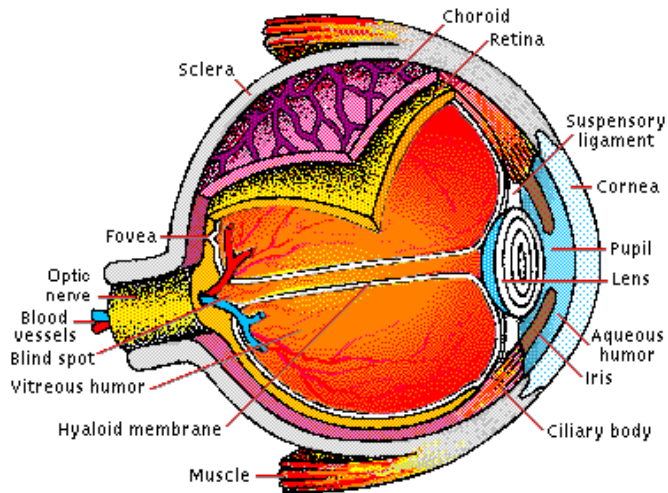


# Formação da imagem no olho humano

- A imagem forma-se no plano da retina por projeção da luz vinda do mundo 3D através de uma lente (cristalino) que muda de dimensões para ajustar a distância focal.



# Elementos do olho Humano



- A retina é a superfície sensível à luz
- Os músculos deformam a lente (cristalino) para mudar o plano focal e assim ajustar-se à distância dos objetos
- A íris controla a quantidade de luz que entra, mudando o diâmetro da pupila

# Componentes na formação de imagem

- Abertura do orifício (diafragma)
- Distância focal (geometria da(s) lente(s))
- Quantidade de luz que atinge o sensor.
  - Esta última é determinante na formação da imagem no elemento sensível, e é afetada por:
    - **Abertura** da lente (orifício do canal de passagem)
    - **Tempo de exposição** para obtenção da imagem
    - **Sensibilidade** à luz do elemento sensível
    - Qualidade do material da lente no que respeita a sua transparência à luz.

# Abertura de uma lente

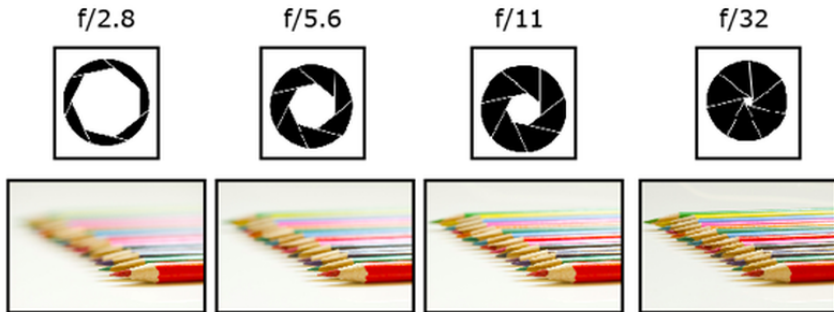
- A abertura é controlada pelo diafragma.
- Afeta a quantidade de luz que passa para o elemento sensível
- Afeta a gama de distâncias (profundidade de campo) dos objetos que ainda resulta em imagem aceitavelmente focada.
  - Grandes aberturas ( $f/1.4$  ou  $f/2.8$ ) deixam passar muita luz mas implicam baixa profundidade de campo.
  - Pequenas aberturas ( $f/11$  ou  $f/16$  ou menos) deixam passar pouca luz mas permitem grande profundidade de campo (praticamente toda a cena focada!)
    - Este último caso é próximo do modelo pin-hole!





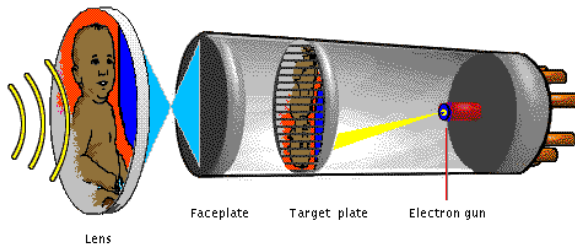
# Abertura e profundidade de campo

- Quanto maior for a abertura, menor será a profundidade de campo.
- Isto é, menor quantidade da cena real surge focada na imagem.
- Exemplo da mesma cena captada com diferentes aberturas.
  - Para se ter a mesma exposição (quantidade de luz), ter-se-á ajustado o tempo de exposição (velocidade do obturador ou *shutter*)



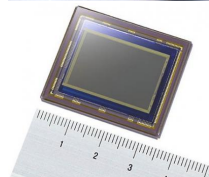
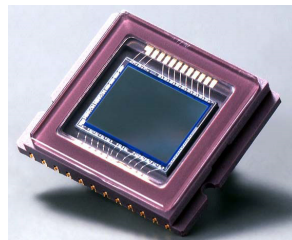
# Obtenção clássica da imagem

- A imagem foi inicialmente obtida por emulsões químicas fotosensíveis depositadas numa chapa ou película colocada no plano focal
  - A fotografia e o cinema tradicionais
- A imagem de televisão foi a primeira a decompor diretamente uma imagem em linhas horizontais suficientemente próximas para criar a ilusão de continuidade vertical
  - Os “tubos” das câmaras de filmar e os cinescópios dos televisores analógicos. As 405 e 625 linhas foram “standards” que vigoraram muitas décadas.
  - Na era digital o número de formatos é muito variado e o número de linhas por imagem pode ir até aos milhares (resoluções 4k e 8k!)



# Sensores modernos para imagem

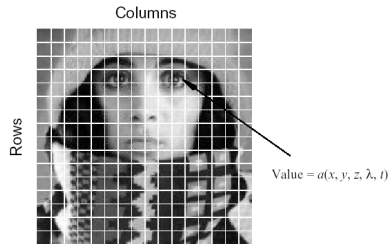
- Os sensores modernos de imagem baseiam-se nos Charge Coupled Devices (CCD)
  - Baixo custo e baixo consumo de energia
  - Reduzidas dimensões (diagonal: desde 2.5 mm até 22.5 mm, ou mais ...)
  - Altas sensibilidades (0.5 lux ou menos)
  - Percepção discretizada em linhas e em colunas (de centenas a milhares)
    - Atualmente medem-se em megapixels
- Há uma tecnologia mais recente de sensores
  - Sensores CMOS - desvantagens e vantagens mútuas;
  - ver comparação por exemplo em:
    - <https://www.phase1vision.com/blog/difference-between-cmos-and-ccd>



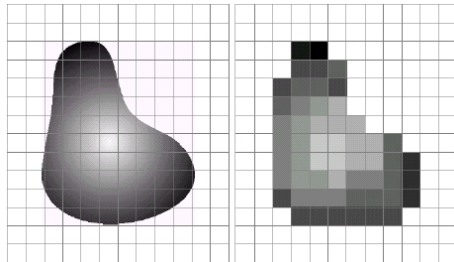
# Representação da imagem digital

# A imagem digital

- Constituída por elementos discretos:
  - o pixel (=Picture Element)
- Processo de amostragem da imagem
  - Amostragem espacial
    - Número limitado (finito) de pontos (pixels) nas duas dimensões
  - Amostragem de amplitude ou intensidade
    - Cada ponto é representado por um número que traduz os níveis de cinzento (intensidade luminosa global)
    - Nas imagens a cores, cada ponto é representado por mais do que um número (normalmente 3)
    - Surge o fenómeno de quantização (número limitado de níveis)



Objeto real — Imagem discretizada



# Representação de uma imagem digital

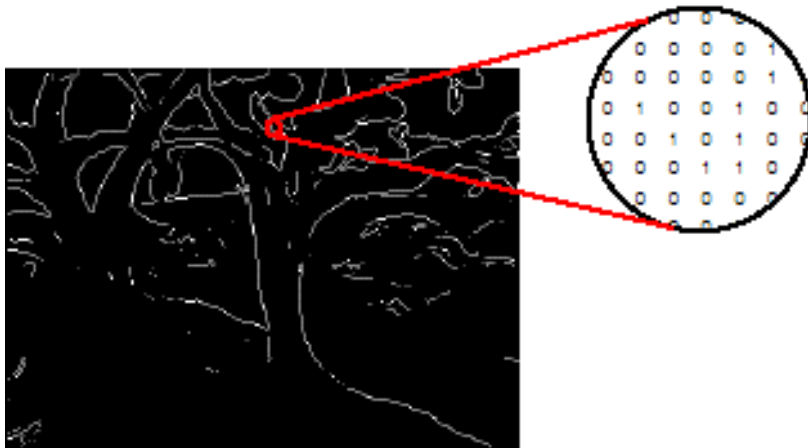
- Uma imagem como uma função  $z=f(x,y)$ 
  - $z$  - intensidade da luz no ponto  $(x,y)$
  - $x,y$  - linha e coluna na imagem respetivamente
- Uma imagem amostrada é representável por uma matriz de dimensão  $(N \times M)$

$$f(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & \cdots & f(0, M-1) \\ f(1, 0) & f(1, 1) & \cdots & f(1, M-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(N-1, 0) & f(N-1, 1) & \cdots & f(N-1, M-1) \end{bmatrix}$$

- Nalguns sistemas a numeração de linhas e colunas começa em 1 e não em 0 (caso do Matlab).

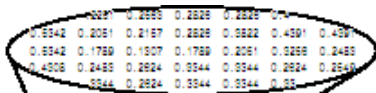
# Imagem Binária

- Imagem Binária:
  - Matriz a 2 dimensões
  - Cada pixel assume o valor 0/1 (on/off)
    - Nalgumas representações 0 ou 255



# Imagem a níveis de cinzento

- Imagem de Intensidade (Níveis de cinzento):
  - Matriz 2 dimensões
  - Representação do nível de intensidade ou nível de cinzento
    - 0-255 (8 bits) ou 0-65535 (16 bits)
    - Ou então: valor decimal entre 0 e 1



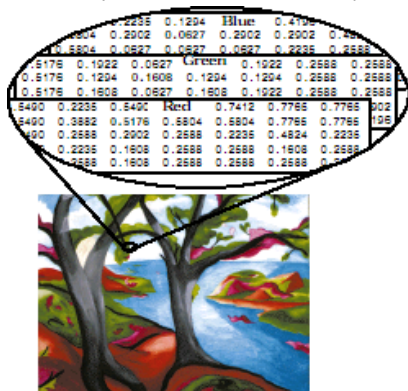
0.2291	0.2553	0.2828	0.2828	0.4
0.5342	0.2051	0.2157	0.2828	0.5522
0.5342	0.1789	0.1907	0.1789	0.2051
0.4908	0.2455	0.2824	0.5344	0.5344
0.5344	0.2824	0.5344	0.5344	0.53





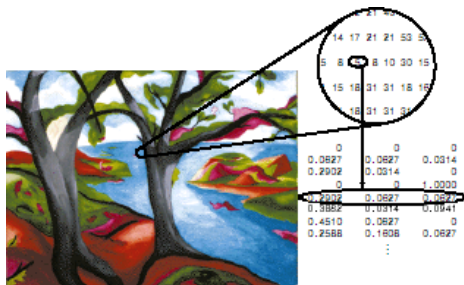
# Imagem a cores

- Imagem a Cores (modelo RGB):
  - Hiper-Matriz 3 dimensões.
  - Em geral imagem a 24 bits (representando R 8 Bits, G 8 Bits, B 8 Bits)
  - A cor é composta pela adição das cores primárias: Vermelho, Verde, Azul
- Há porém outros modelos de cor (abordado mais tarde)



# Imagens a cores indexadas

- Há alguns tipos de representação de imagens em que o valor pixel é dado por um índice que aponta para uma tabela de cores possíveis (as palettes de cores)
- O formato GIF de ficheiro de imagens é um exemplo desta representação
- De qualquer modo, para todos os efeitos, cada pixel tem as três componentes de cor



# Resolução numa Imagem

- Resolução espacial
  - Número de linhas e colunas
    - pixels verticais e horizontais
  - Efeito mosaico
- Resolução de amplitude ou profundidade
  - Número de níveis de cinzento ou cores por pixel – nas imagens a cores há mais que um valor por pixel (é usual três)
  - O número de bits e efeito de “solarização”
  - Nas imagens binárias, os pixels são representados por um só bit



150 x 200



75 x 100



38 x 50



256



4



1

- Imagem - níveis de cinzento
- Variar resolução espacial
- Variar resolução de profundidade
- Carregar imagem a cores RGB
  - Ilustrar os conceitos das componentes
  - Mostrar as três componentes à vez
  - Fazer variar uma das componentes
  - Mudar a intensidade de uma imagem
  - Etc.