

Visão Computacional

Prof. Bogdan Tomoyuki Nassu



Nesta aula...

- Introdução.
 - O que é visão computacional?
 - Exemplos de aplicações.
 - Tarefas comuns.
- Informações sobre a disciplina.
 - Enfoque.
 - Pré-requisitos.
 - Avaliação.
 - Bibliografia.
- OpenCV – Open Source Computer Vision.

Professor: Bogdan T. Nassu

- Formação: ciência da computação.
 - Graduação e mestrado: UFPR (2002 e 2005).
 - Doutorado: Universidade de Tokyo (2008).
- Profissional:
 - Área: processamento de imagens e visão computacional.
 - UTFPR (2012~).
 - Professor, DAINF.
 - UFPR (2011~2012).
 - Pós-doutorado (CNPq).
 - Railway Technical Research Institute, Tokyo (2008~2011).
 - Pesquisador.
 - LACTEC (2001~2005).
 - Estagiário, bolsista (assistente de pesquisa).

www.dainf.ct.utfpr.edu.br/~nassu

Visão computacional?

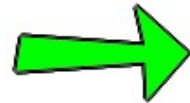
- O que é “ver”?



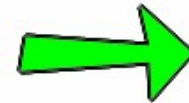
Visão computacional?



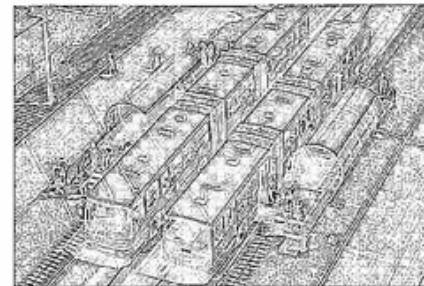
Mundo Real



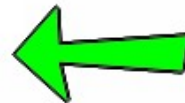
Sensor



Representação



Processamento
/ Interpretação



Modelo /
Conhecimento

Definição

- Shapiro & Stockman.
 - *The goal of computer vision is to make useful decisions about real physical objects and scenes based on sensed images.*
- Szeliski.
 - *In computer vision, we are trying to (...) describe the world that we see in one or more images and to reconstruct its properties.*
- Neves, Vieira Neto e Gonzaga.
 - *A visão computacional [tem] (...) como objetivo a obtenção de algoritmos capazes de interpretar o conteúdo visual de imagens.*
- Wikipedia.
 - *Computer vision (...) includes methods for acquiring, processing, analyzing, and understanding images (...) in order to produce numerical or symbolic information, e.g. in the forms of decisions.*
- Não existe uma definição universalmente aceita!
 - Mesmo o nome “visão computacional” é sujeito a críticas.

Áreas relacionadas

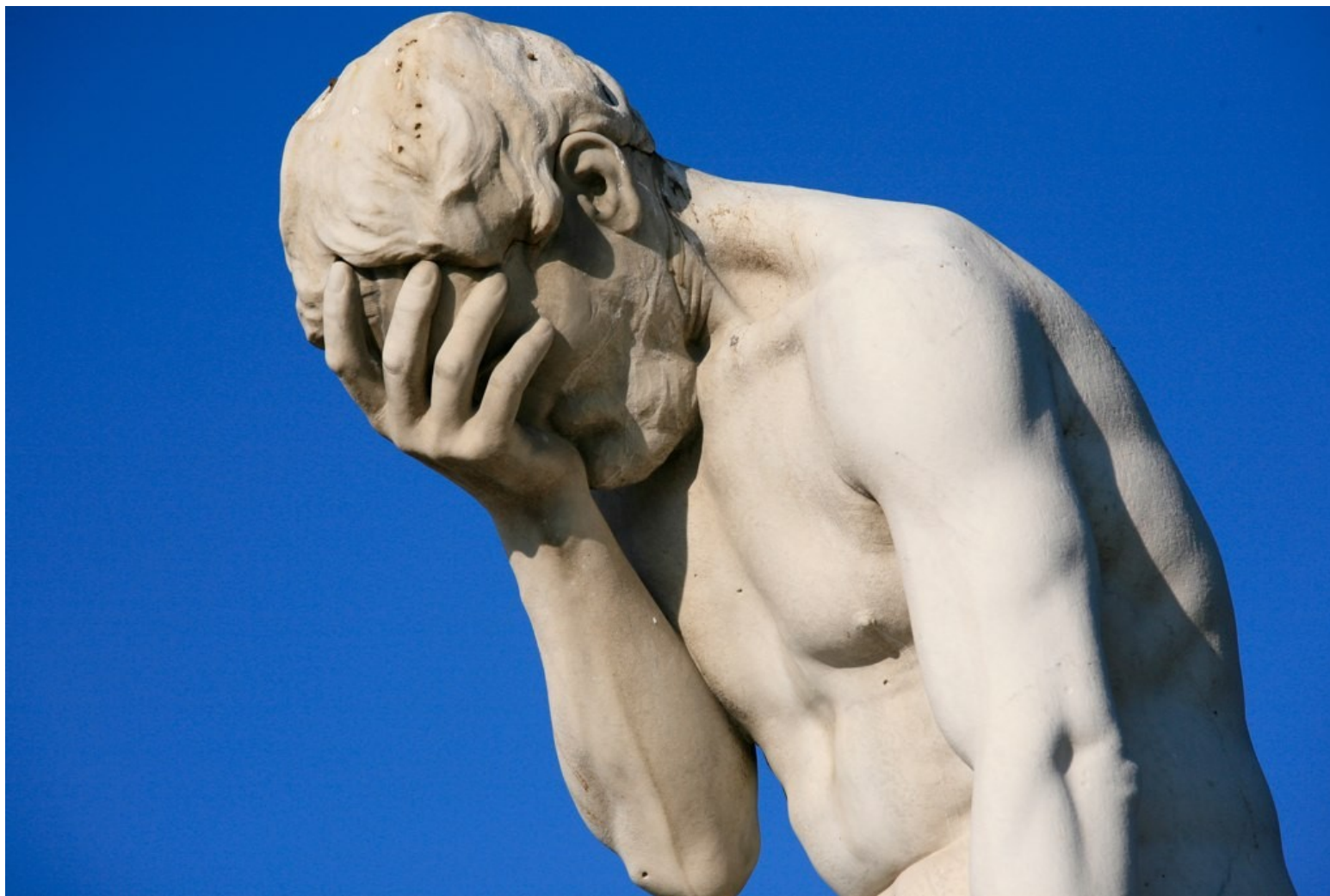
- Processamento de imagens.
 - (Qual a diferença?)
- Inteligência artificial.
 - Aprendizado de máquinas.
 - Reconhecimento de padrões.
- Mundo físico.
 - Hardware.
 - Sensores.
 - Ótica.
- Visão (biologia e neurociência).
- Computação gráfica?!

Uma questão importante...

- “Ver” é uma tarefa trivial para seres humanos e animais.
 - Mas o quão difícil é fazer um computador enxergar?
- Nos primórdios da inteligência artificial, as pesquisas se concentravam em lógica simbólica.
 - Muitos consideravam que o desafio estava em realizar inferências a partir de símbolos (casa, pessoa, porta, etc.), e que obter os símbolos seria comparativamente simples.

Uma questão importante...

- “Obter os símbolos seria comparativamente simples”.

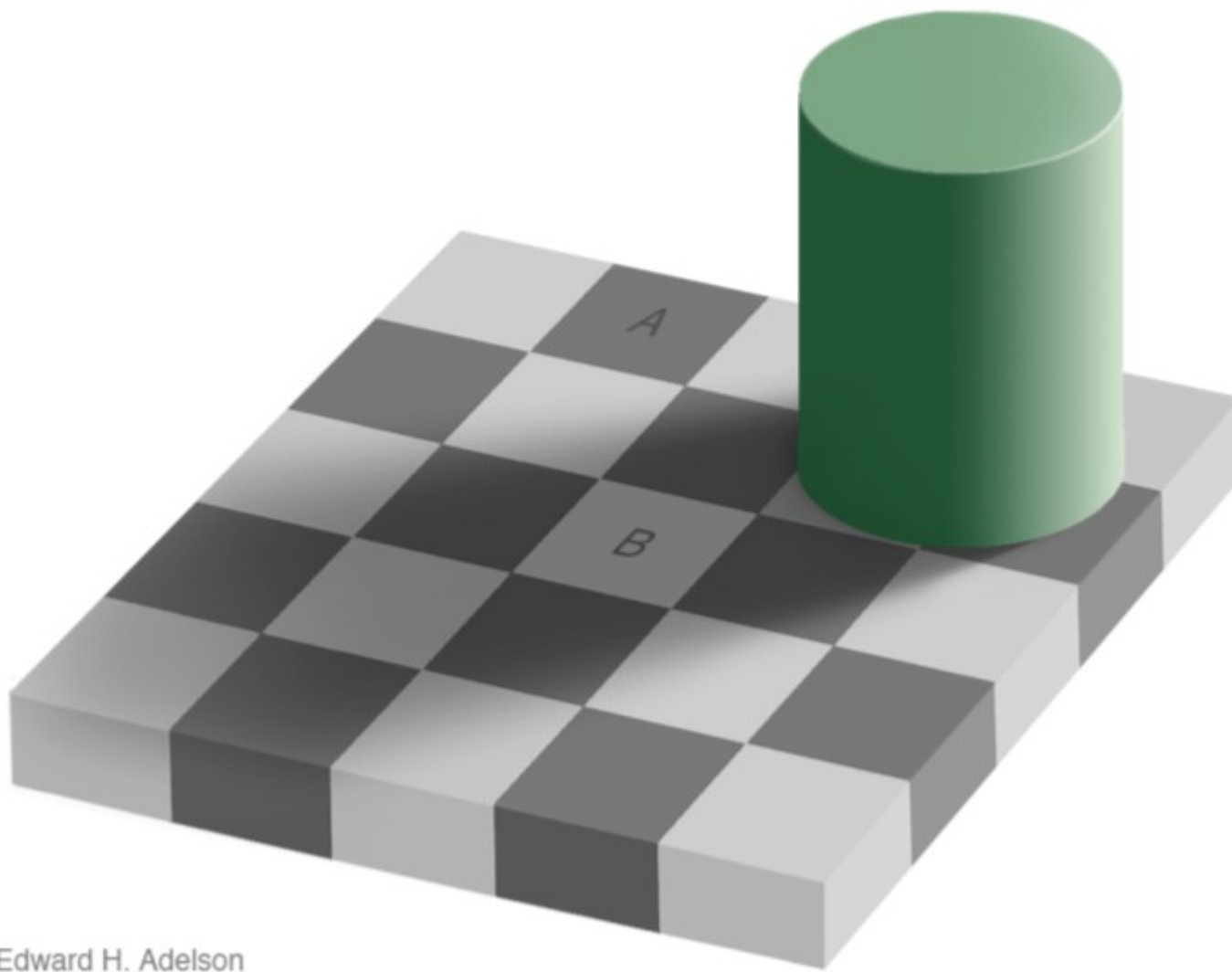


Uma questão importante...

- Enxergar é uma tarefa muito mais “inteligente” do que parece!

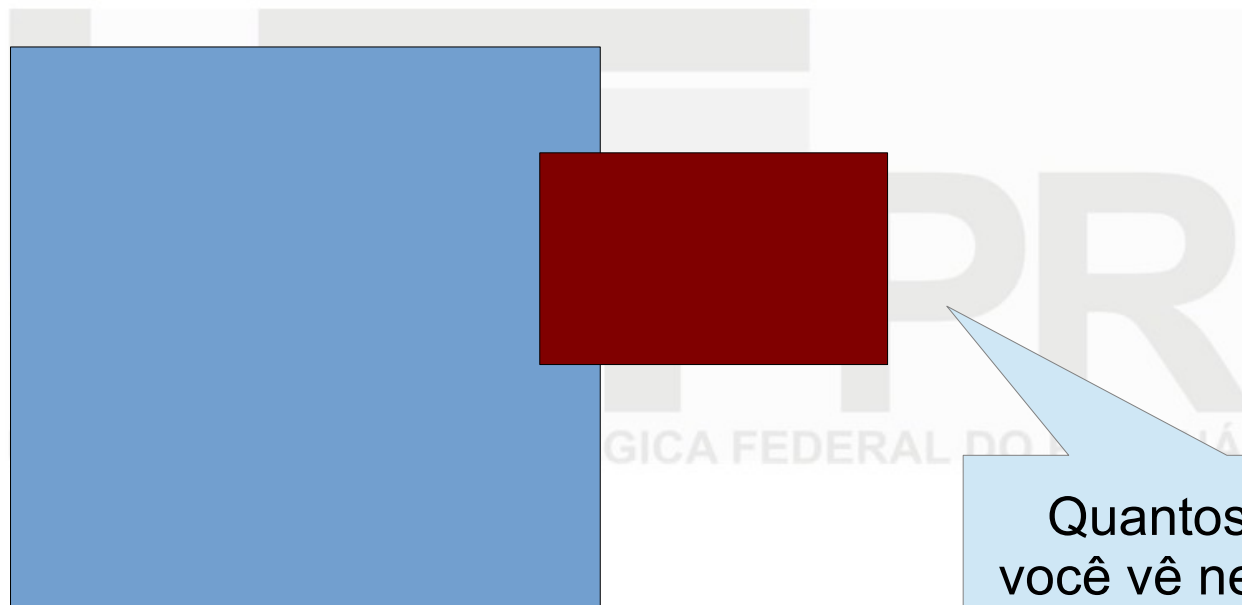


Contexto e interpretação



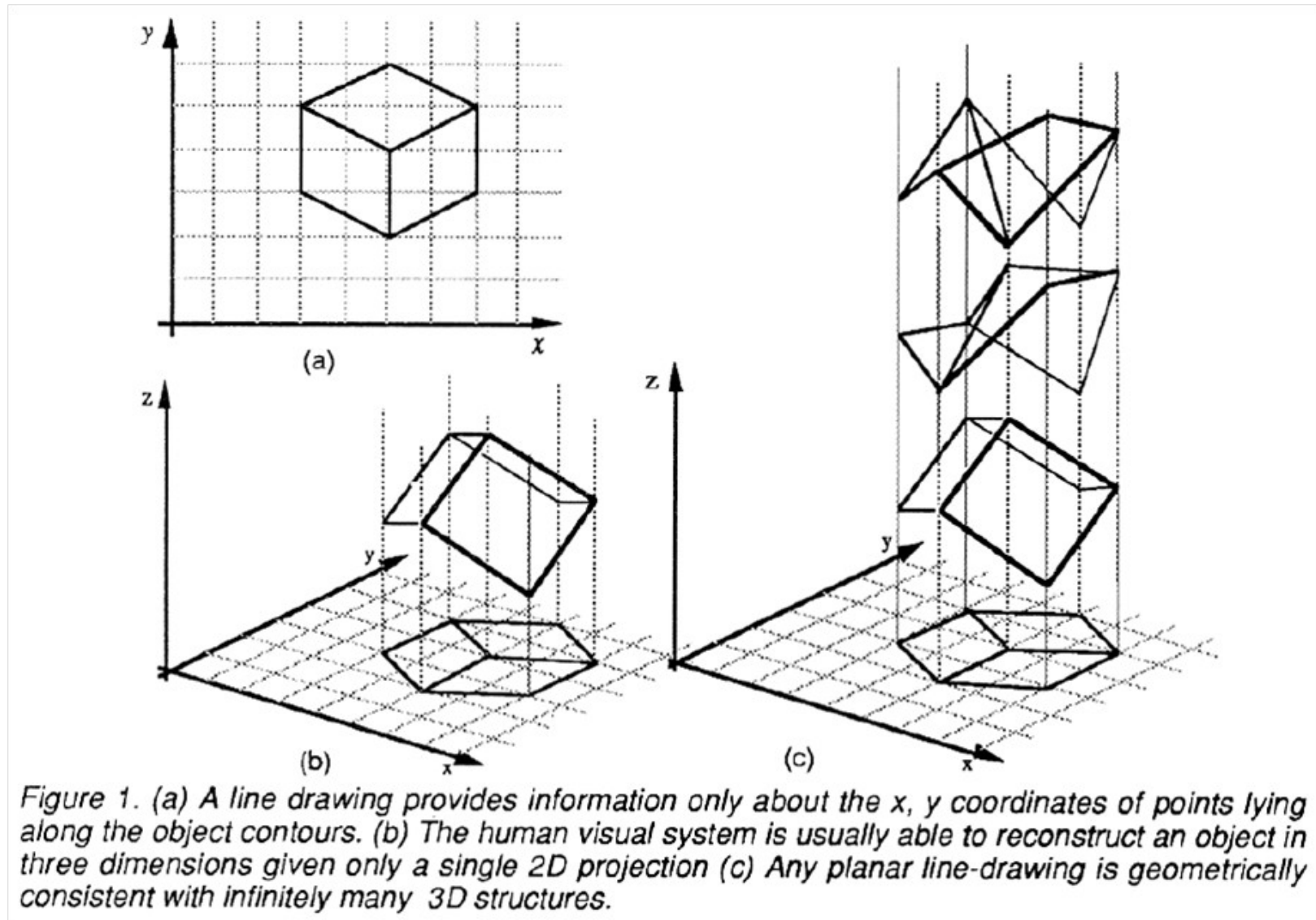
Edward H. Adelson

Contexto e interpretação

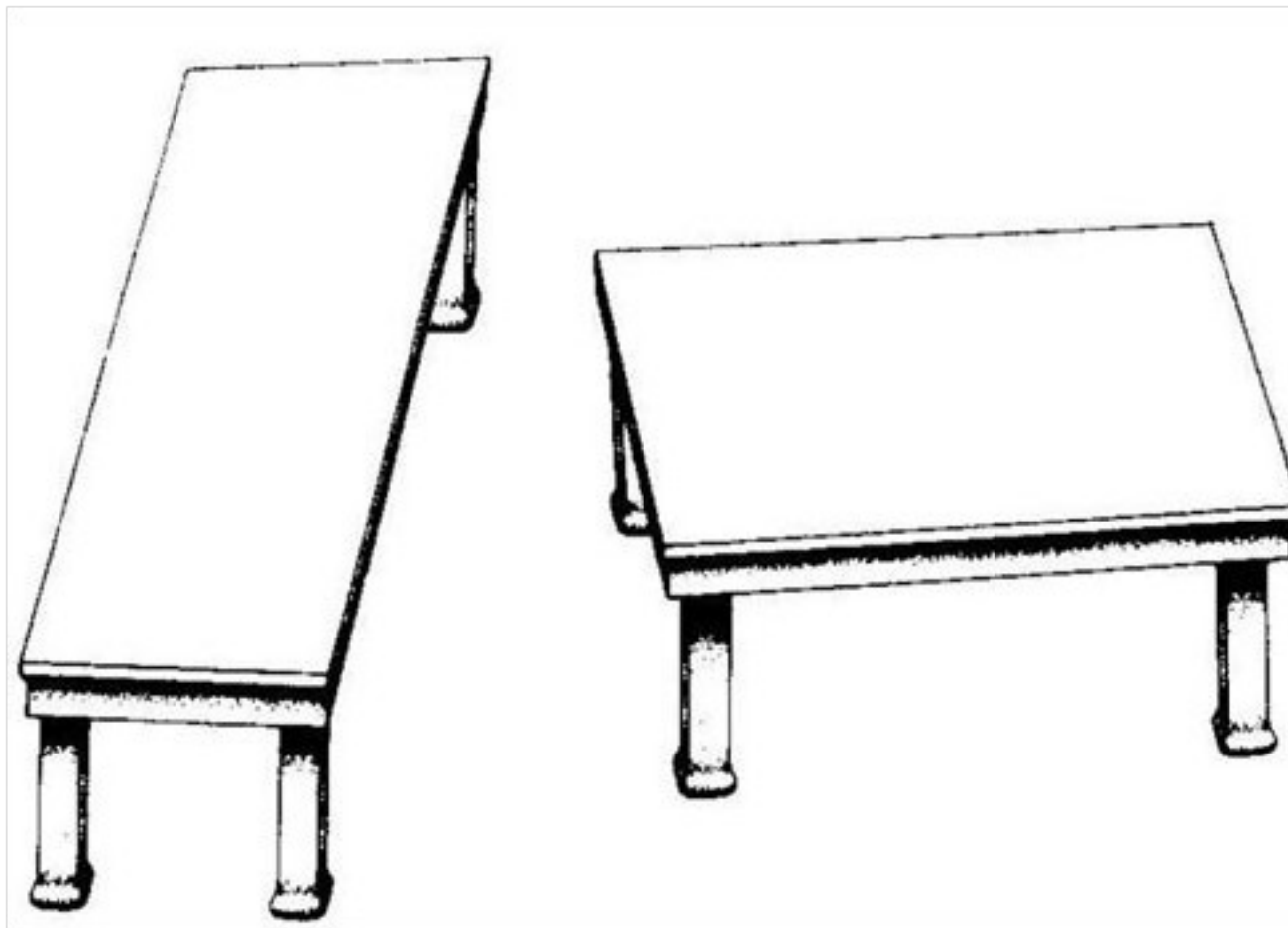


Quantos retângulos
você vê nesta imagem?

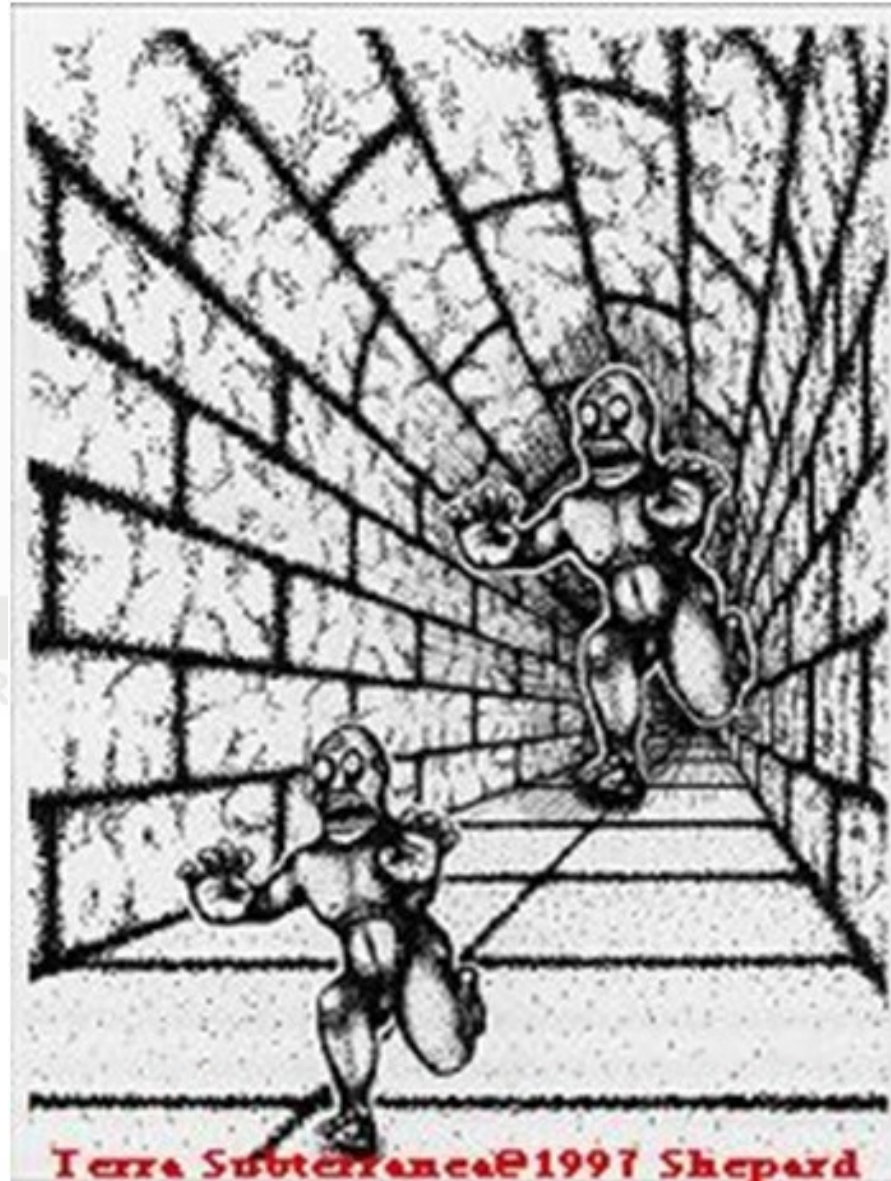
Contexto e interpretação



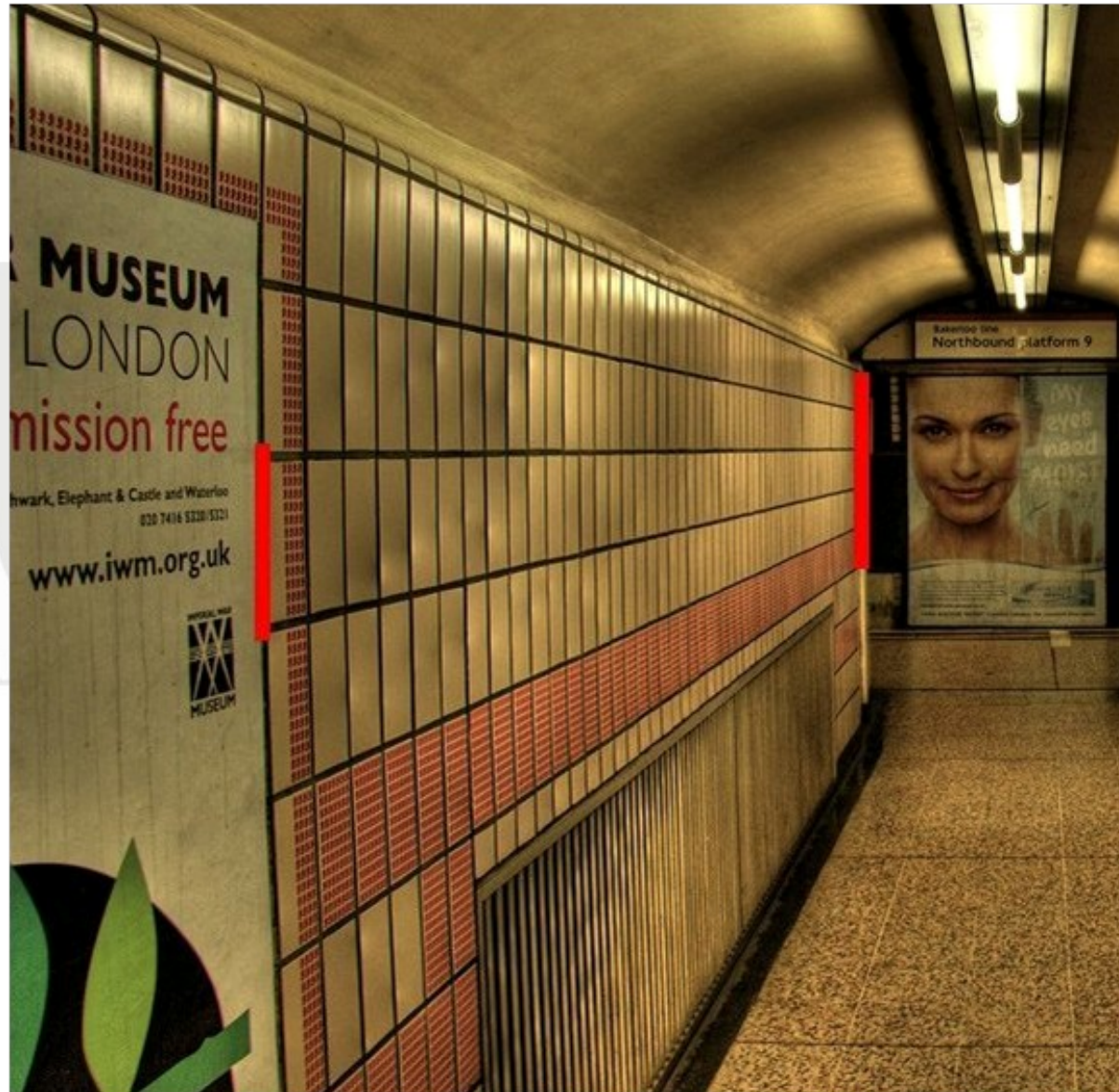
Contexto e interpretação



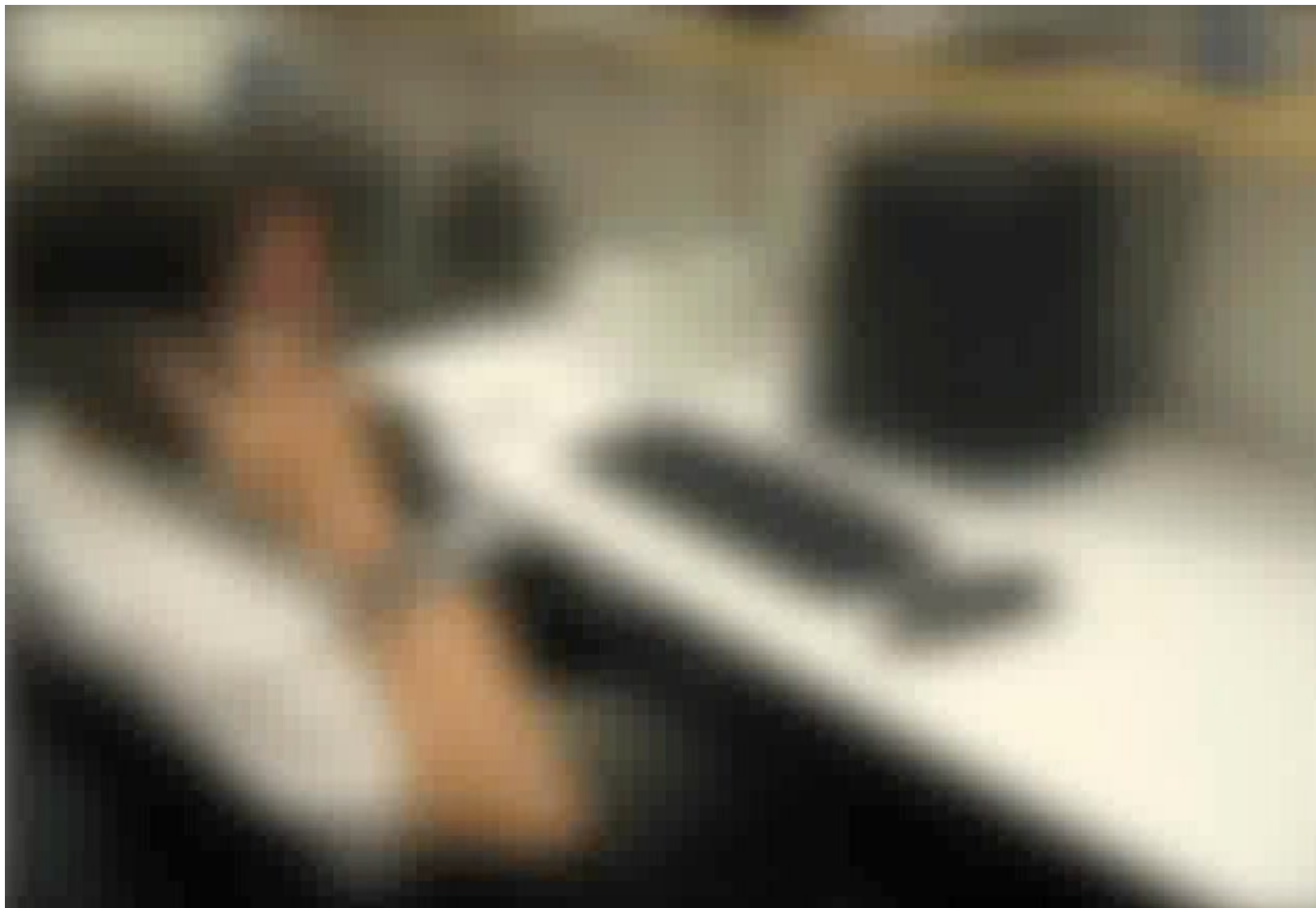
Contexto e interpretação



Contexto e interpretação



Contexto e interpretação



Contexto e interpretação



Contexto e interpretação



http://youtu.be/QWMFpxkGO_s

Moral da história...

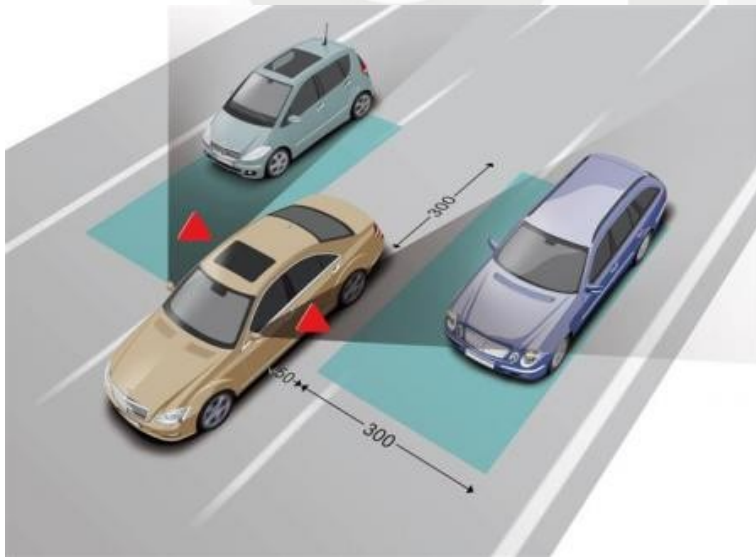
- Quando enxergamos, fazemos muitas simplificações, suposições e adivinhações.
 - Não trabalhamos sobre medições diretas dos valores dos “pixels” que enxergamos.
 - Nos baseamos na nossa experiência para interpretar o contexto e atribuir sentido ao que vemos.
- Aplicações práticas de visão computacional frequentemente codificam conhecimento específico do domínio.
 - Quando “travar”:
 - Como eu enxergo isso? Quais elementos me permitem compreender esta cena? Como poderia isolá-los? Quais são as asserções que eu faço inconscientemente?
 - Cuidado quando for trabalhar com aprendizado de máquinas!

Áreas de aplicação

- Medicina.
- Robótica.
- Automação.
- Documentos e texto.
- Entretenimento.
- Sistemas de transporte inteligentes.
- Etc.

Exemplo: ITS

- Sistemas de assistência ao motorista.
 - Visão noturna.
 - Assistente para permanência na pista.
 - Assistente para pontos cegos.
 - Monitoração da atenção.
 - Frenagem automática.
 - Assistente para baliza.



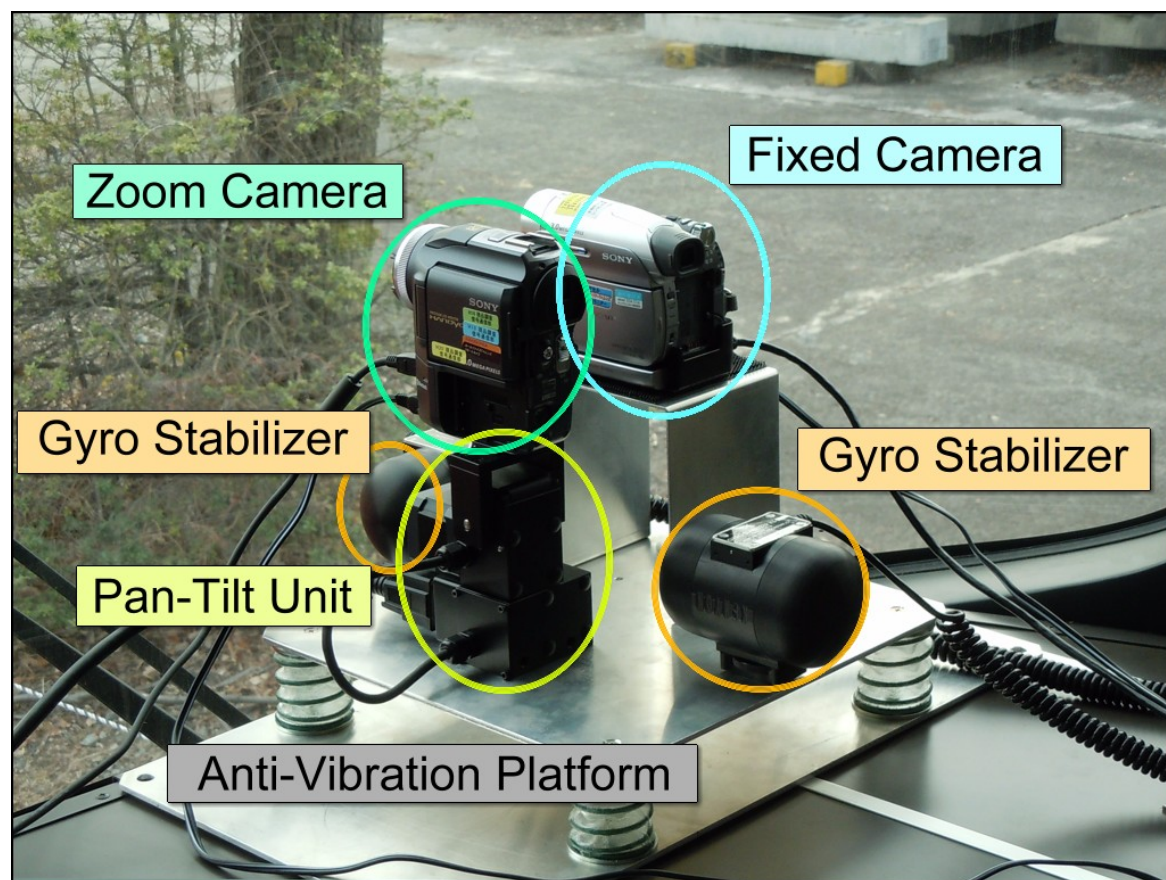
Exemplo: veículos autônomos

- DARPA Grand Challenge (2004).
 - 240km no deserto nos EUA.
 - Melhor resultado: 11.8km (!).
- DARPA Grand Challenge (2005).
 - 212km.
 - Melhor tempo: 6h54m.
- DARPA Urban Challenge (2007).
 - Simulação de trecho urbano.
- VisLab Intercontinental Autonomous Challenge (2010).
 - De Parma até Shanghai em 3 meses.
- Estado atual.



Exemplo: extração de trilhos

- Railway Technical Research Institute.
 - Direcionamento de uma câmera sobre uma unidade pan/tilt.

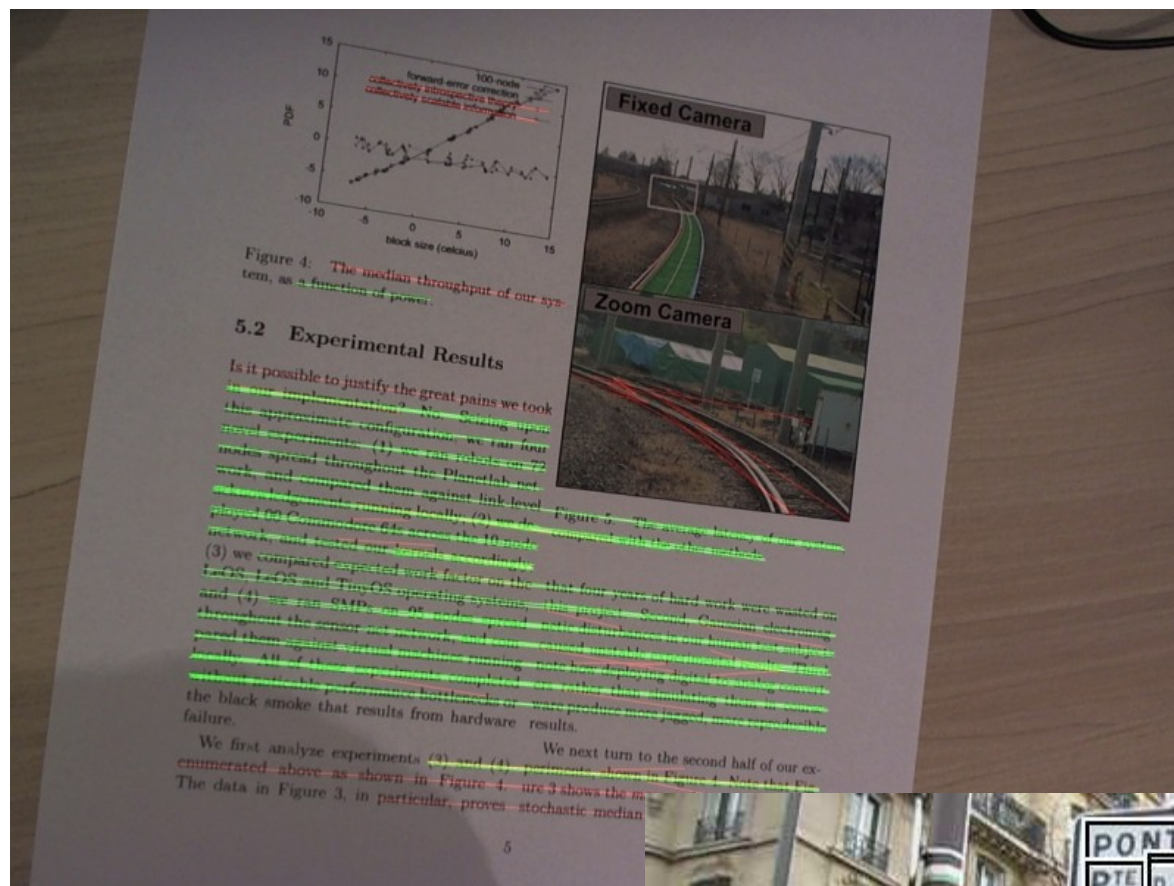


<http://www.youtube.com/user/NassuRTRI>

Exemplo: reconhecimento de sinalização



Exemplo: detecção de texto



Exemplo: detecção de faces



Tag Your Friends

This will quickly label your photos and notify the friends you tag. [Learn more](#)



Who is this?



Who is this?



Who is this?



Who is this?

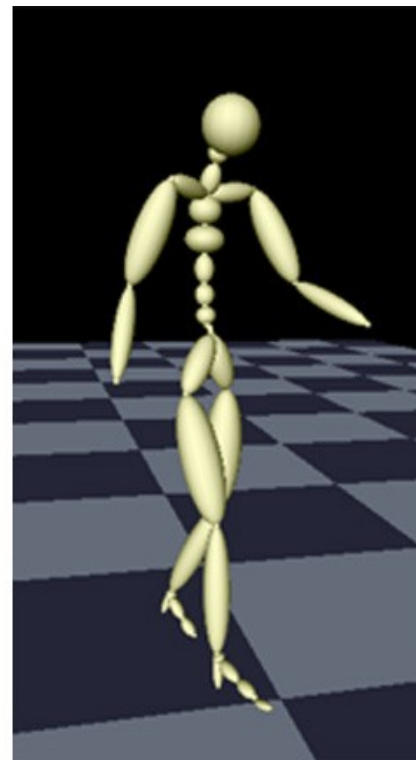
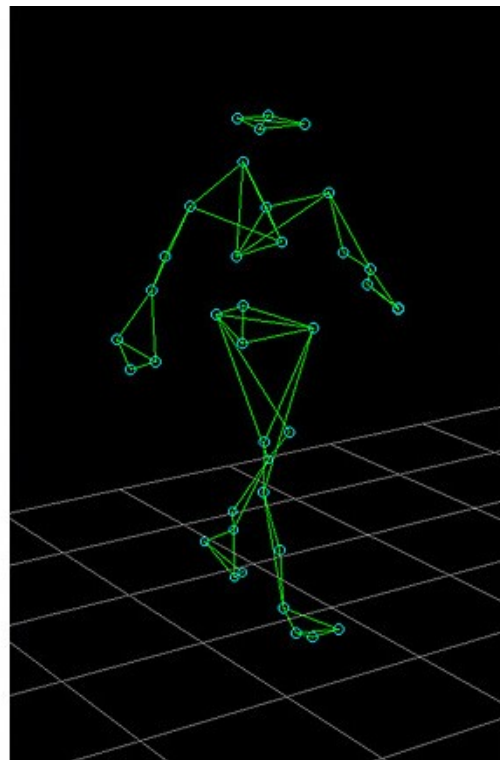


Who is this?



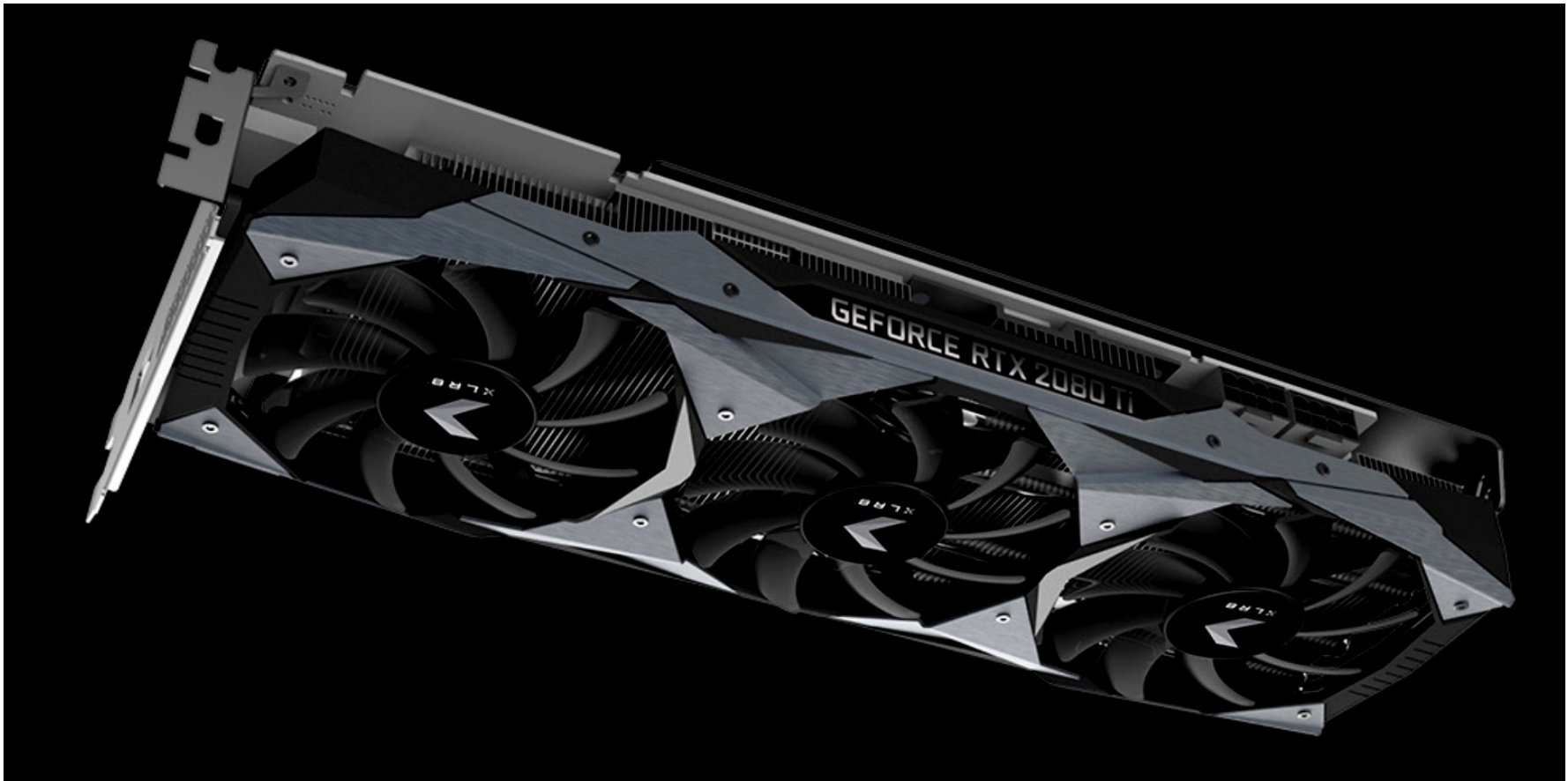
Who is this?

Exemplo: captura de movimento



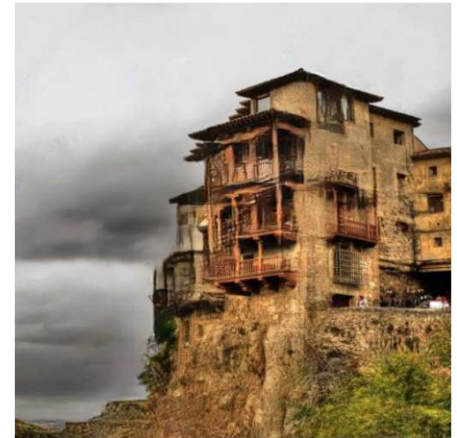
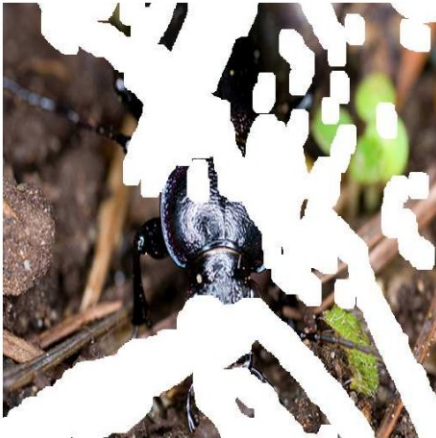
... e mais

- Aprendizado de máquinas + processamento de imagens = *área quente*.



... e mais

- Aprendizado de máquinas + processamento de imagens = *área quente*.



Algumas tarefas comuns...

- Algumas tarefas comuns em visão computacional.
 - Detecção.
 - Localização.
 - Reconhecimento.
 - Classificação.
 - Segmentação.
 - Rastreamento.

Detecção

- Determinar se um objeto de interesse está presente na imagem.



“Existe ao menos uma placa nesta imagem”.

Localização

- Determinar a posição dos objetos de interesse na imagem.
 - Frequentemente, “detecção” é usado também para localização.



“Existem duas placas nesta imagem, e estas são as suas posições”.

Reconhecimento

- Determinar o estado ou identidade dos elementos presentes na imagem, em um universo de possíveis respostas.
 - As “possíveis respostas” costumam estar em uma lista.

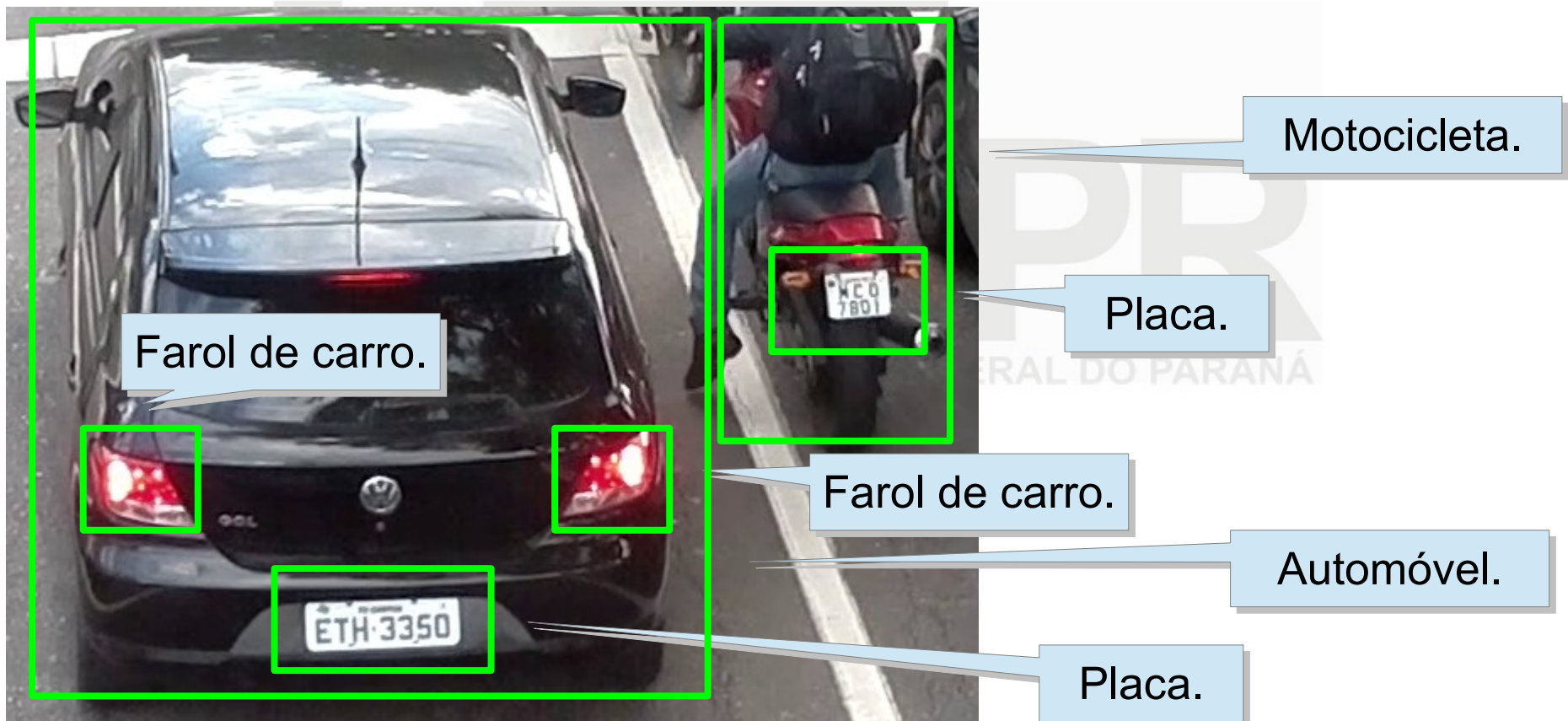


Frequentemente, o reconhecimento é feito sobre regiões, após uma detecção.

“ETH-3350”

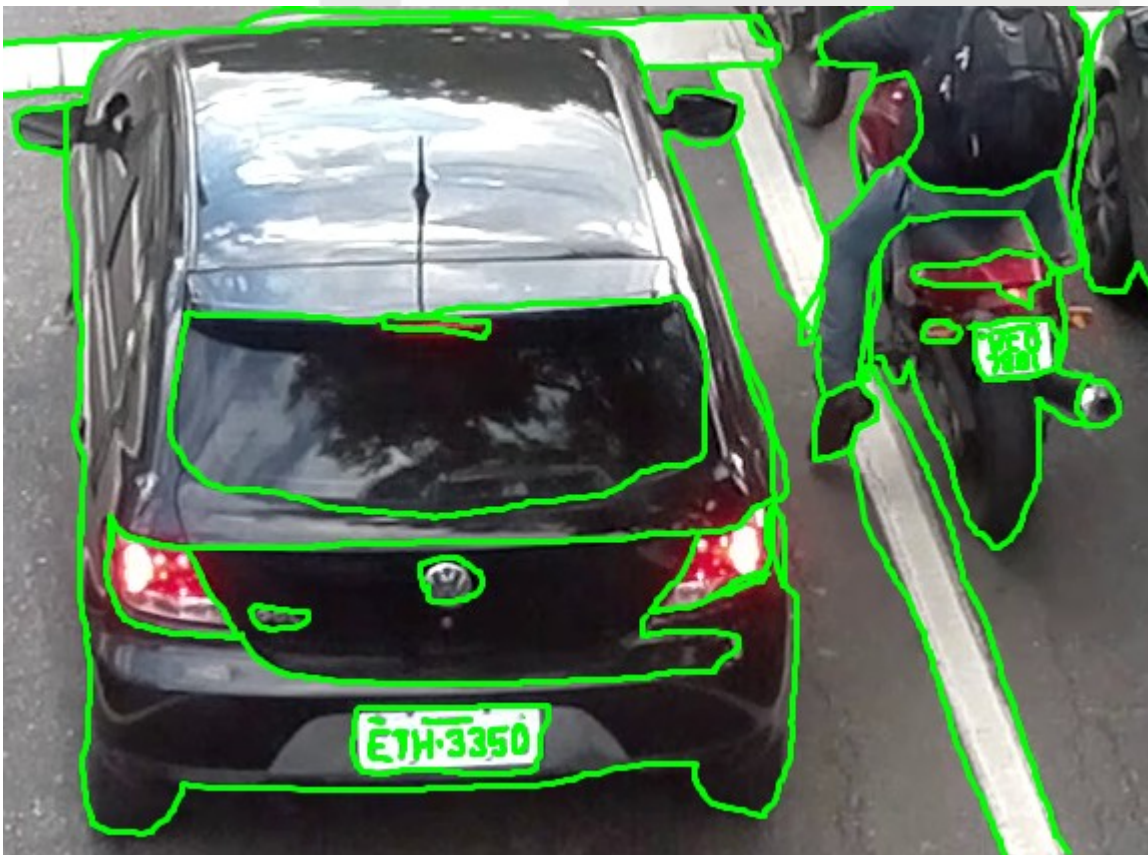
Classificação

- Selecionar em uma lista os rótulos adequados para uma imagem.
 - Classificação e reconhecimento podem ser problemas idênticos.



Segmentação

- Separar a imagem em regiões (pixels).
 - “Baixo nível” x “alto nível”.



Rastreamento

- Acompanhar a trajetória de objetos ou regiões em vídeos.
 - Só faz sentido em sequências de imagens!



Desafios

- Como dar sentido a uma matriz de números?
- Variações, variações...
 - Existem muitos modelos de veículos.
 - Efeito de perspectiva modifica o aspecto das letras na placa.
 - Variações de iluminação modificam os valores dos pixels.
 - Veículos do mesmo modelo podem ter cores diferentes.
 - Lataria pode estar amassada.
 - ...
- Como lidar com isso?

Soluções

- Extrair primitivas.
 - Bordas, cantos, retas, círculos, regiões homogêneas, etc.
- Trabalhar sobre propriedades estatísticas.
 - “Aprendizado de máquinas” → criar modelos estatísticos.
- Usar descritores.
 - Um descritor é um “resumo” das propriedades de uma imagem.
- Controlar o ambiente para reduzir variações.
 - Câmera em posição fixa x câmera móvel.
 - Iluminação conhecida x iluminação natural.
 - Rosto visto de frente x posição arbitrária.

Pré-requisitos

- Pré-requisitos:

- Programação.
- Estruturas de dados.
- Processamento digital de imagens.
- Um pouco de estatística e geometria analítica.

Bibliografia

- Kaehler & Bradski, *Learning OpenCV 3*, O'Reilly, 2017.
- Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2011.
 - szeliski.org/Book/
- Goodfellow, Bengio & Courville, *Deep Learning*, MIT Press, 2016.
 - www.deeplearningbook.org
- Gonzalez, *Processamento Digital de Imagens*, 3ª ed., Pearson, 2011.
- Artigos de congressos e periódicos da área.

OpenCV



- Open Source Computer Vision.
 - Biblioteca open source com rotinas para visão computacional e aprendizado de máquinas.
 - Licença BSD.
 - Escrita em C++.
 - Interfaces para C++, Python e Java.
 - OS: Windows, Linux, Android e Mac OS.

opencv.org