

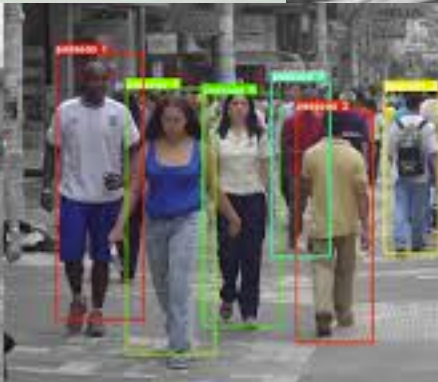
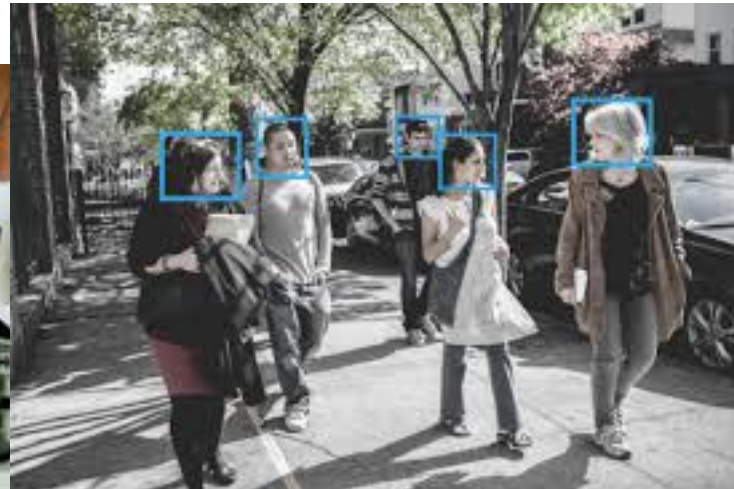
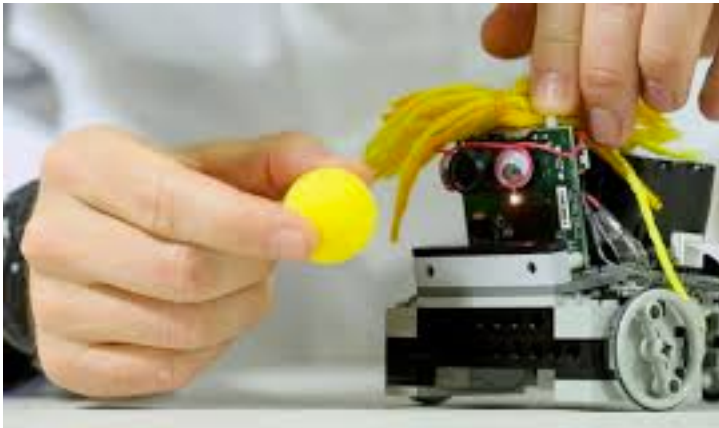


Introdução a Visão Computacional

Prof. Dr. Geraldo Braz Junior

DEINF - UFMA

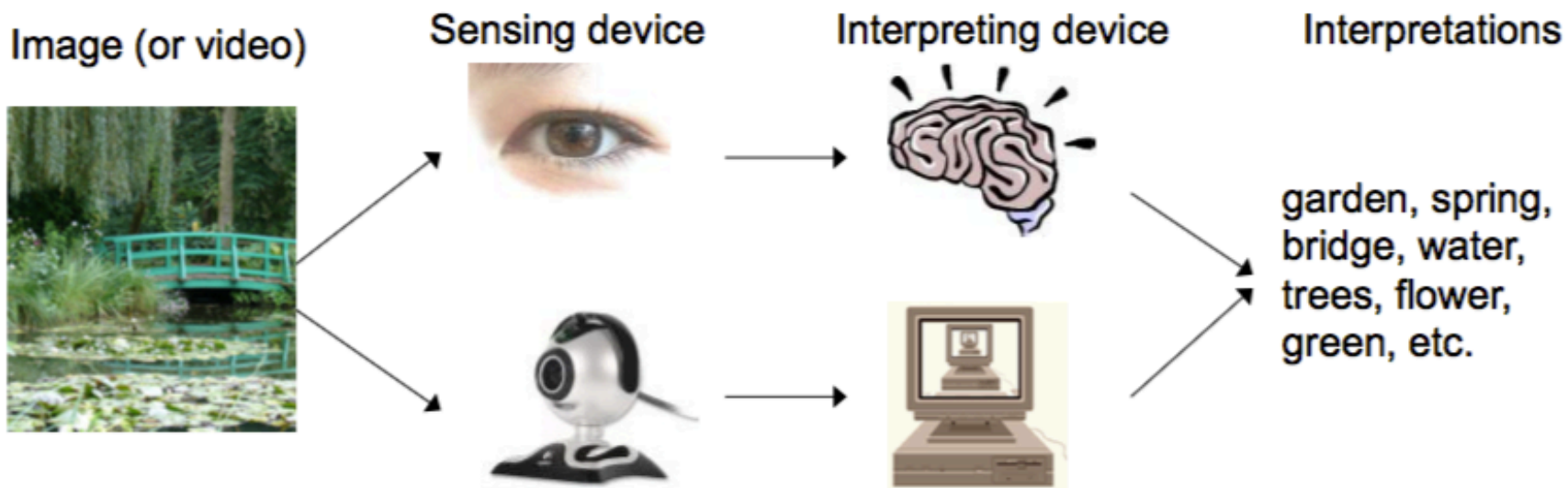
O que é Visão Computacional? Você já viu?



O que se espera?

- Visão é sobre entender imagens
 - Coloridas ou em tons de cinza
 - Pequenas sequências ou videos
 - Tomadas com uma câmera fixa ou em movimento
 - De uma câmera calibrada ou não
- Somente imagens?
- Com o objetivo básico de extrair informação útil dessas imagens de acordo com a aplicação.

Como esperamos o funcionamento?



Visão x Relacionamento

Visão Geral



Áreas: Engenharia, Robótica, Ciência da Computação, Psicologia, Matemática, Física, Biologia

Qual o objetivo: entender



What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Source: S. Narasimhan

Percepção

- “O que nós **experimentamos**, aparentemente diretamente, é na verdade bem diferente do que nossos órgãos estão **registrando**” [adaptado de Perception: From Sense to Object by J. Wilding]
- “Cada imagem é a **imagem de uma coisa** apenas para aquele que **sabe como lê-la**, e que esteja habilitado com a ajuda da imagem para formar uma idéia da coisa” [adaptado de Handbook of Physiological Optics by H. Helmholtz]

Visão Humana

- A visão humana é soberbamente eficiente
 - Detecta rapidamente padrões em uma cena



Quantos recortes
tem animais?

Segmentação/ Percepção



Segmentação / Percepção

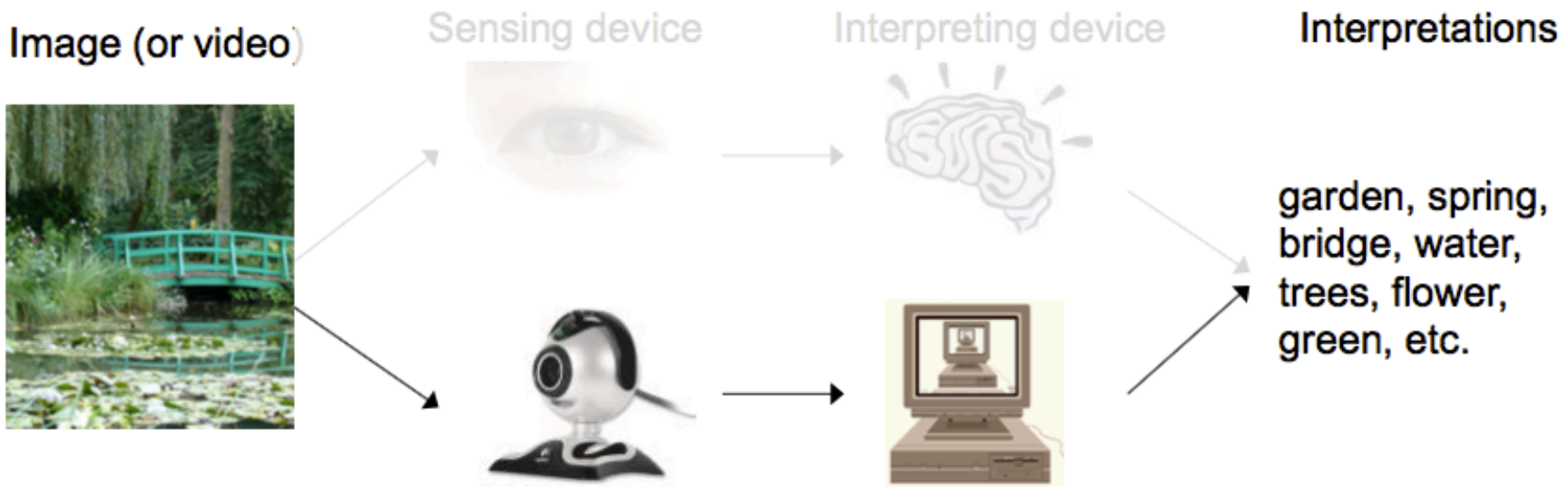


... mas variações ...



podem ser dificilmente percebidas pela visão humana

Visão Computacional

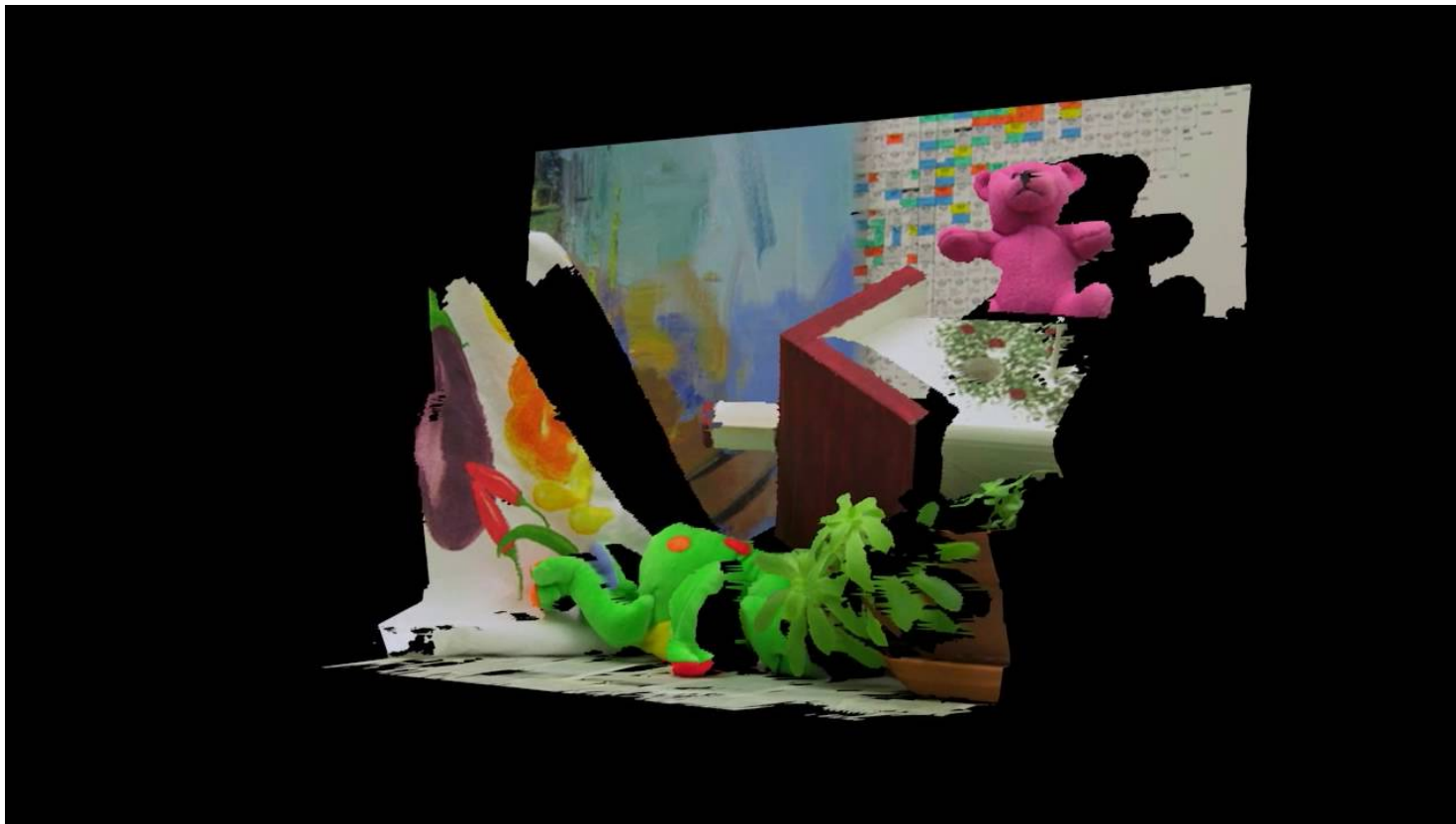


Quais informações podemos extrair?

- Métricas (2D, 3D)
 - Características Estruturais
- Semântica
- E se não soubermos modelar as métricas?

Aplicações: Detecção

- Graph Cut based Continuous Stereo Matching using Locally Shared Labels (CVPR 2014)



Aplicações: Tracking

- Joint Tracking and Segmentation of Multiple Targets (CVPR 2015)



**thus, also performing instance-based
video segmentation.**

Aplicações: Reconstrução 3D

- DynamicFusion: Reconstruction and Tracking of Non-rigid Scenes in Real-Time (CVPR 2015)



Live Input Depth Map



Live Model Output



Live RGB Image (unused)



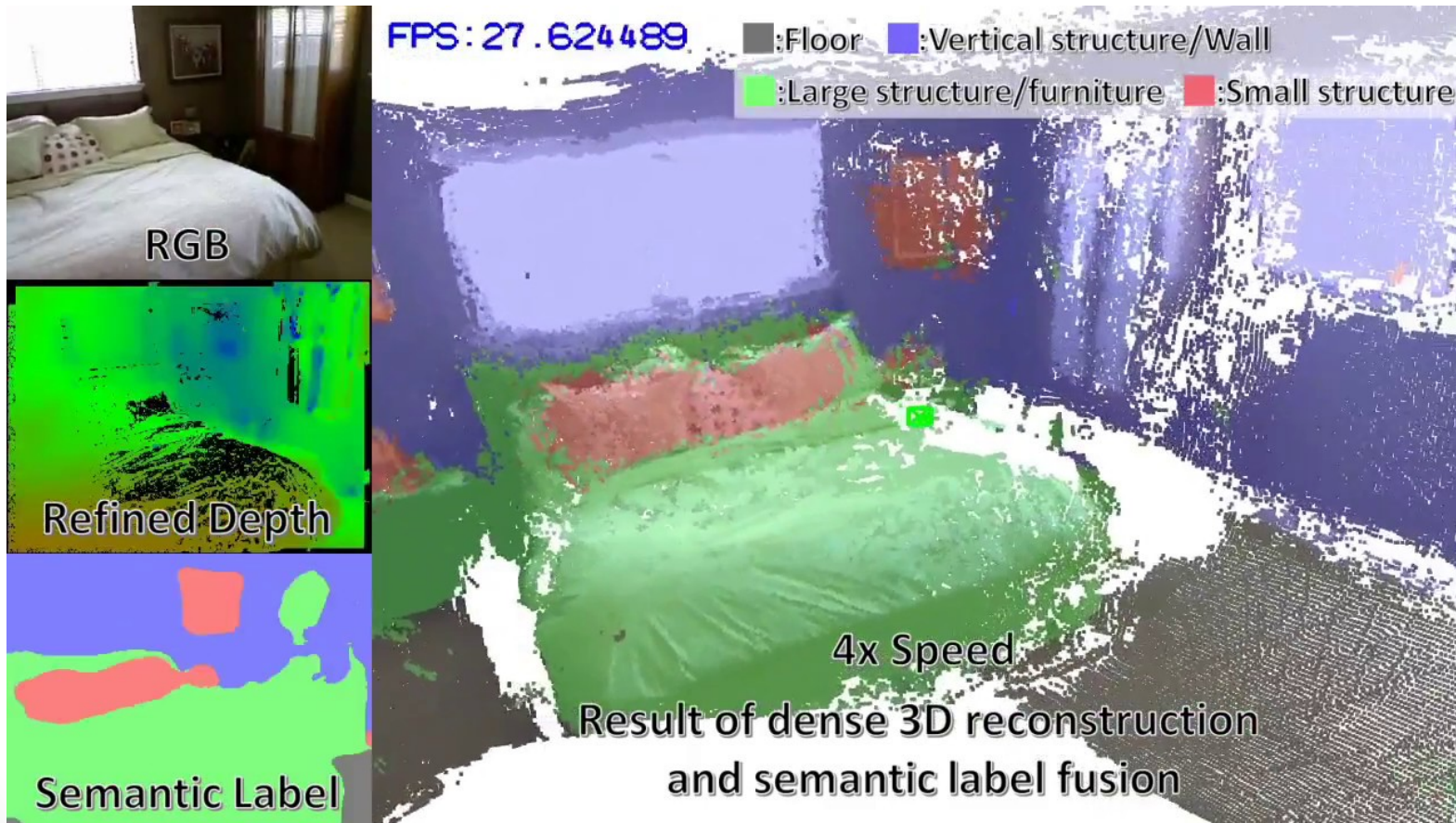
Canonical Model Reconstruction



Warped Model

SLAM

Real-time dense monocular SLAM with learned depth prediction (CVPR 2017)



Aplicações: Detecção de Face

- Face2Face Real-time Face Capture and Reenactment of RGB Videos (CVPR 2016)



Estimação de Ação

- [CVPR 2017] Action-Decision Networks for Visual Tracking with Deep Reinforcement Learning

Action-Decision Networks for Visual Tracking with Deep Reinforcement Learning

Sangdoon Yun¹ Jongwon Choi¹ Youngjoon Yoo²
Kimin Yun³ and Jin Young Choi¹

1. ASRI, Dept. of Electrical and Computer Eng., Seoul National University, South Korea
2. Graduate School of Convergence Science and Technology, Seoul National University, South Korea
3. Electronics and Telecommunications Research Institute (ETRI), South Korea

Estimação de Pose

Realtime Multi-Person 2D Human Pose Estimation using Part Affinity Fields,
CVPR 2017

Real-time Multi-Person 2D Pose Estimation Using Part Affinity Fields

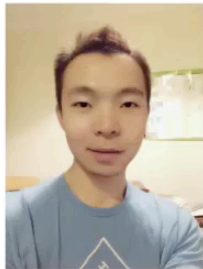
Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh
Carnegie Mellon University

Estimação – Previsão

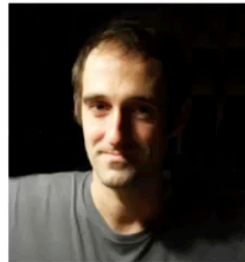
- Estimating body shape under clothing (CVPR 2017)

Detailed, Accurate Human Shape
Estimation from Clothed Scan Sequences

<http://buff.is.tue.mpg.de/>



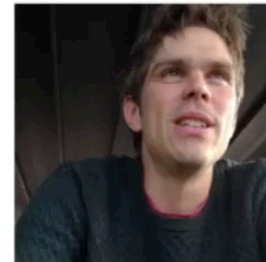
Chao Zhang^{1,2}



Sergi Pujades¹



Michael Black¹



Gerard Pons-Moll¹

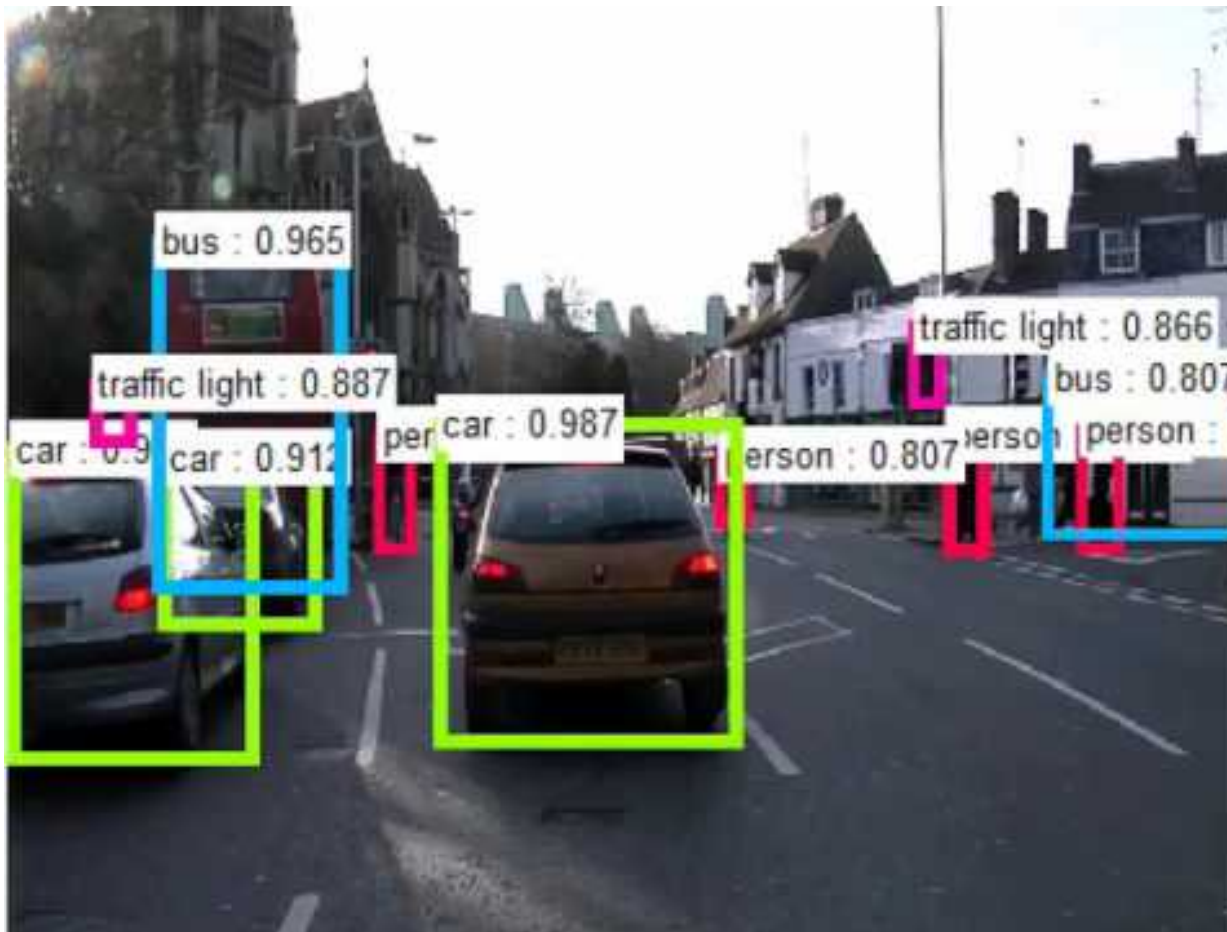
¹ Max Planck for Intelligent Systems

² University of York



Aplicações: Detecção e Reconhecimento Supervisionado

- Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks (Tese – Cornell University)



Reforço

- End-to-end Driving with Deep Reinforcement Learning (CVPR 2017 Workshop)

CVPR 2017 Workshop: Deep Learning for Robotic Vision

End-to-end Driving in a Realistic Racing Game with Deep Reinforcement Learning

Etienne Perot¹ Maximilian Jaritz^{1,2} Marin Toromanoff¹ Raoul de Charette²

¹Valeo

²Inria, RITS Team

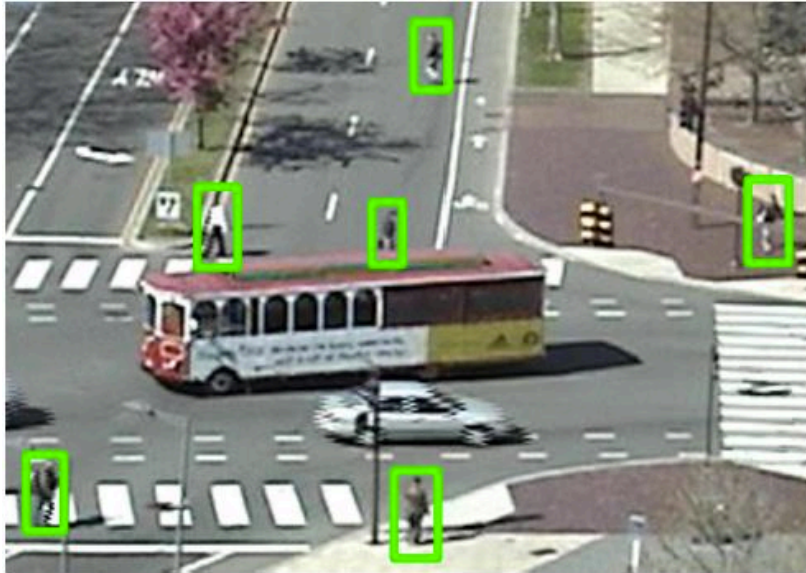


Aplicação: Imitação

- Artistic style transfer for videos (Tese – Cornell University)

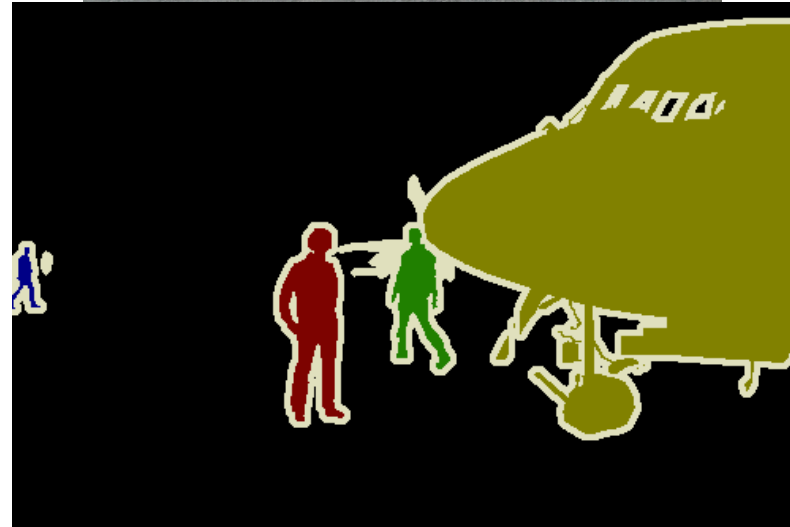


Visão como “IO”



Transferring a Generic Pedestrian Detector Towards Specific Scenes

Visão como “Propostas”

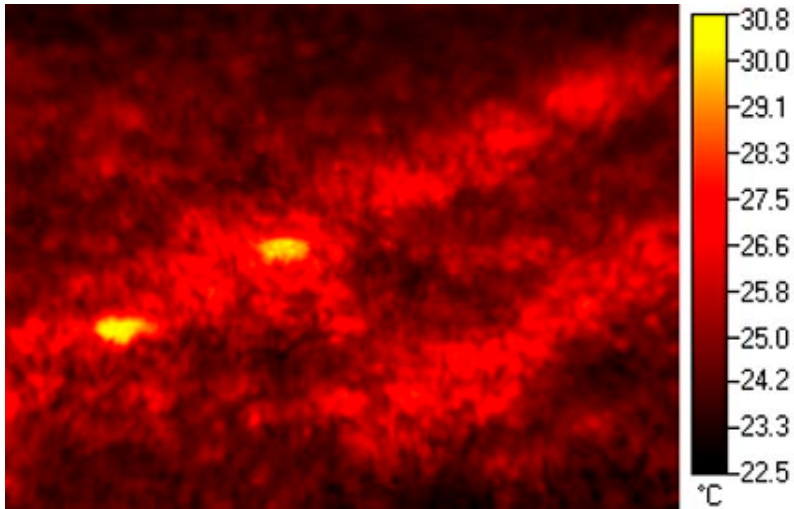


Constrained Parametric Min-Cuts for Automatic Object Segmentation

Visão como “Biometria”

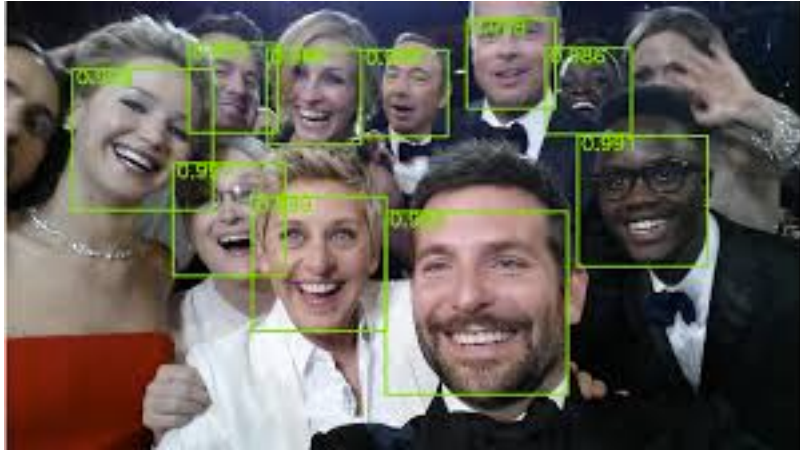


Visão como “Inspeção”

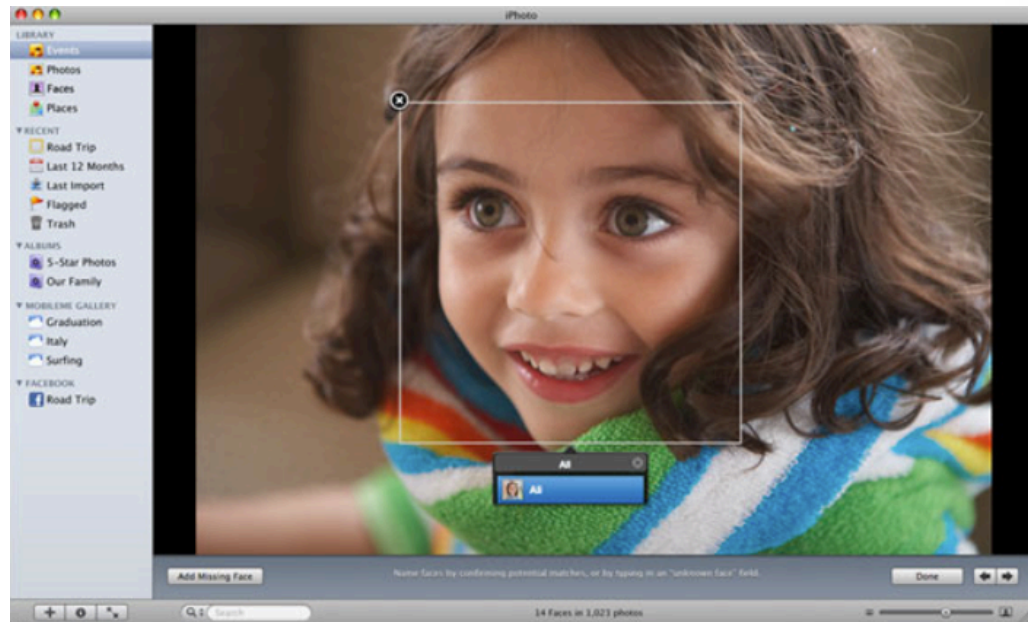


Detecção de Minas Terrestres

Visão como “detecção”



Face Detection

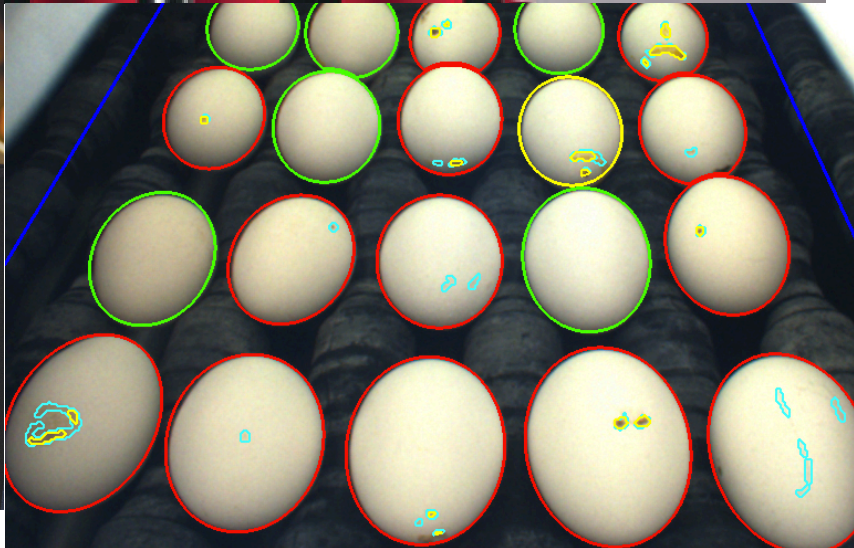


Face Recognition

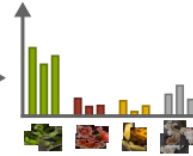
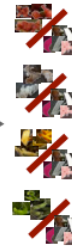
Visão como “medida”



Sch 01 OK PLine
1.015
Region 1st
A C 01:00
79100 89 91 89 86100

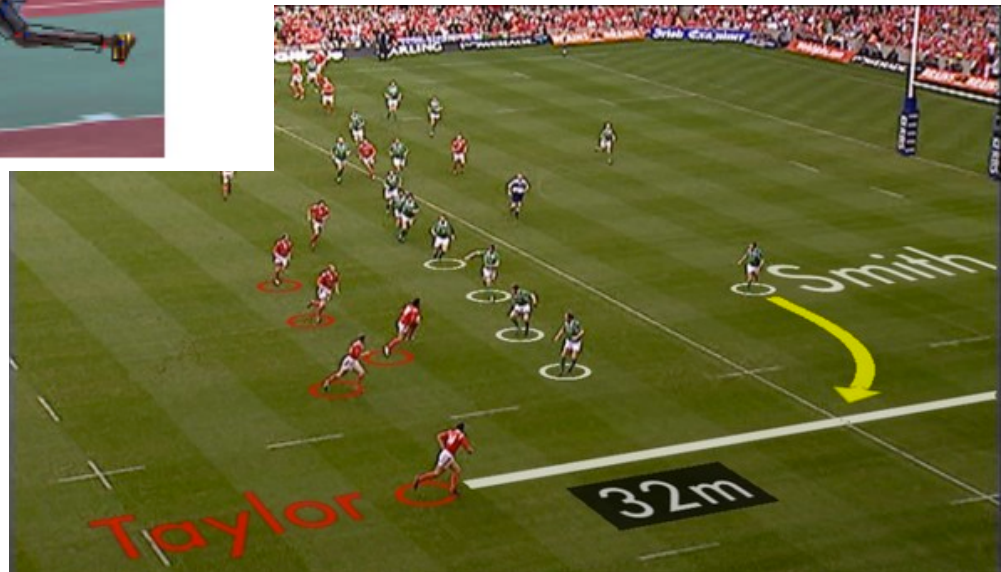
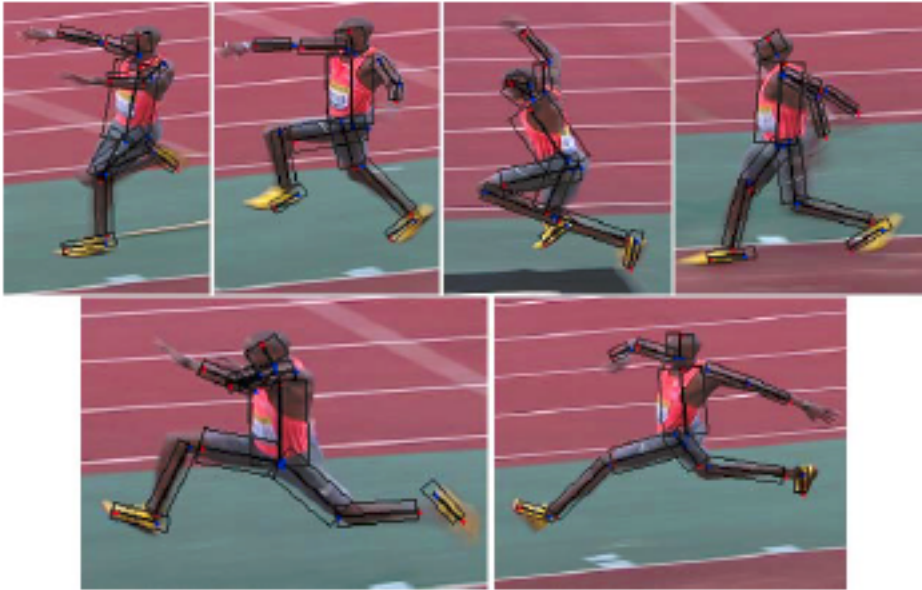


Visão como “saúde”

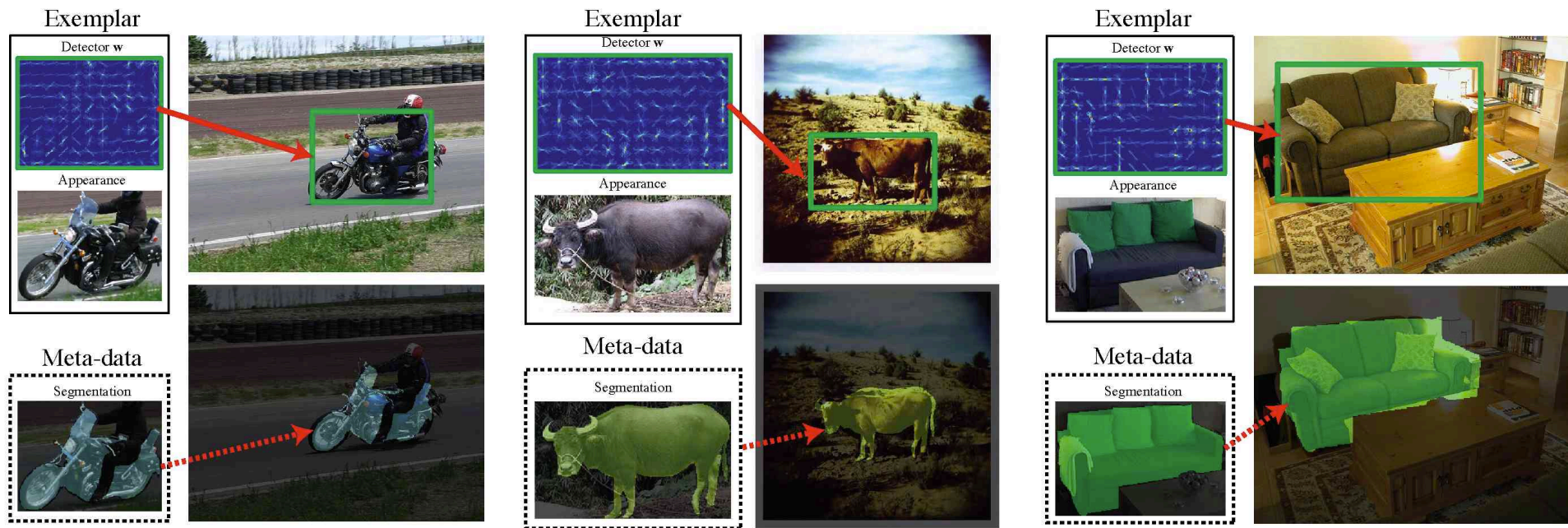


Food-101 – Mining Discriminative Components with Random Forests

Visão como “melhoria” ou “tática”



Visão como “correlação”



Ensemble of Exemplar-SVMs for Object Detection and Beyond

Visão como “modelagem” e “visualização 3D”

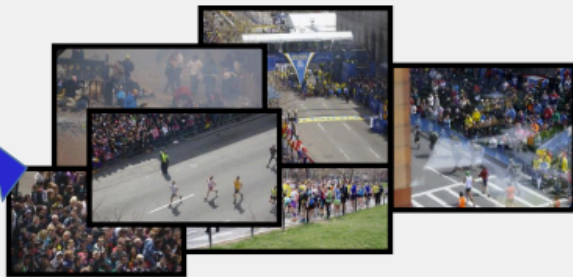


<http://photosynth.net>

Visão como “semântica”



Suspect #1: Male, sunglasses, black and white hat, blue shirt



Person search

[Kumar et al. 2008, Feris et al. 2013]



How we teach computers to understand pictures

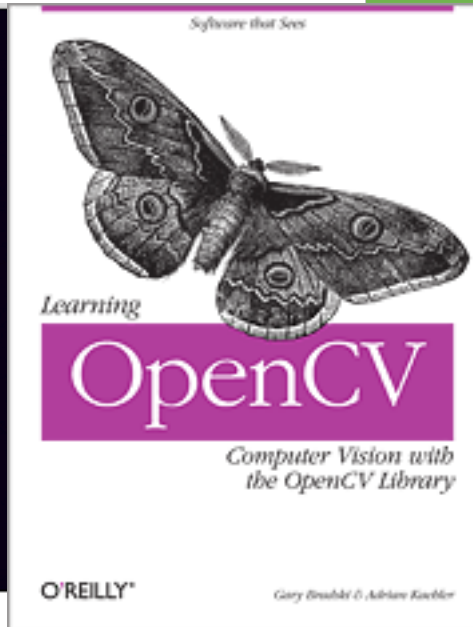
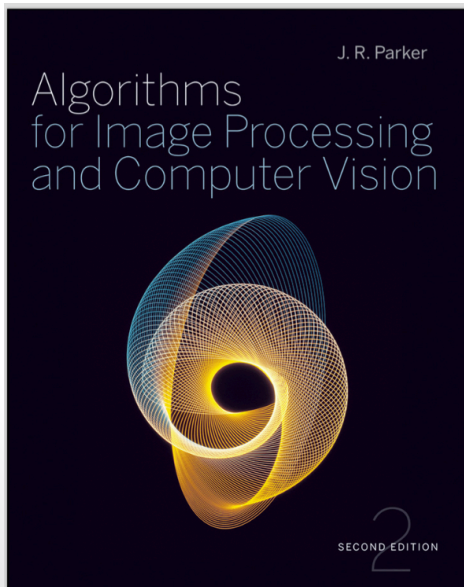
- Fei Fei Li – Ted Talks



Objetivo Máximo

Detecção + interpretação → em qualquer
cena!!!

Algumas bibliografias



Todos disponíveis na internet