

RECONSTRUÇÃO DE OBJETOS 3D BASEADOS EM IMAGEM:

Ivan Lopes <ivanlopes@id.uff.br>
ivanlopes.eng.br
13 de dezembro de 2019

Estado da arte e tendências
na era do
Aprendizado profundo



Reconstrução de objetos 3D

- ▶ problema antigo
- ▶ Estudado por:
 - visão computacional;
 - computação gráfica;
 - comunidades de aprendizado de máquina (2015)



Estrutura

2015 ConvNet - Sistemas neurais Convolucionais;

Trabalho Recentes:

- deep learning;
- estimar objetos 3D usando Imagens RGB (1 ou ∞);



Estrutura

Literatura Revisada:

- representações de forma;
- arquitetura de rede;
- trabalhos recentes com formas e rostos humanos



Estrutura

- ▶ Comparação de desempenho;
- ▶ Pesquisas futuras;



Tabela de conteúdos

- 1 Introdução
- 2 definições
- 3 Método estatístico
- 4 Método cognitivo
- 5 Softwares



Objetivo da reconstrução 3D

Inferir a geometria e a estrutura de objetos e cenas.

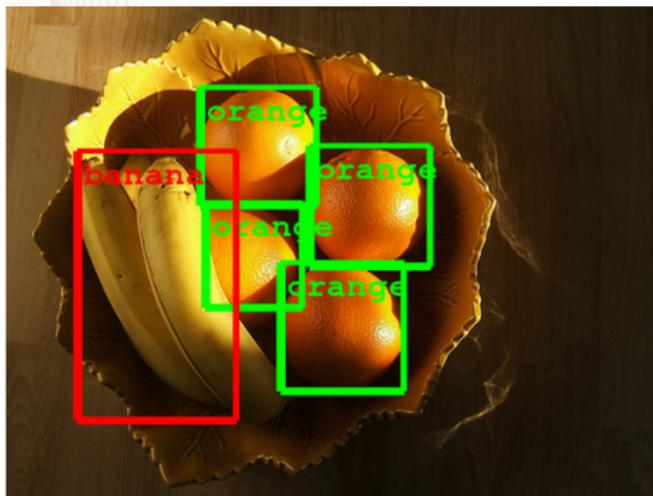
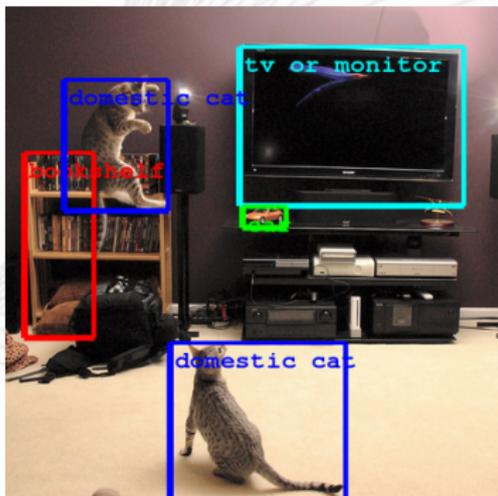


Aplicações

- ▶ como navegação por robô;
- ▶ reconhecimento de objetos e compreensão de cenas;
- ▶ modelagem 3D e animação;
- ▶ controle industrial;
- ▶ diagnóstico médico

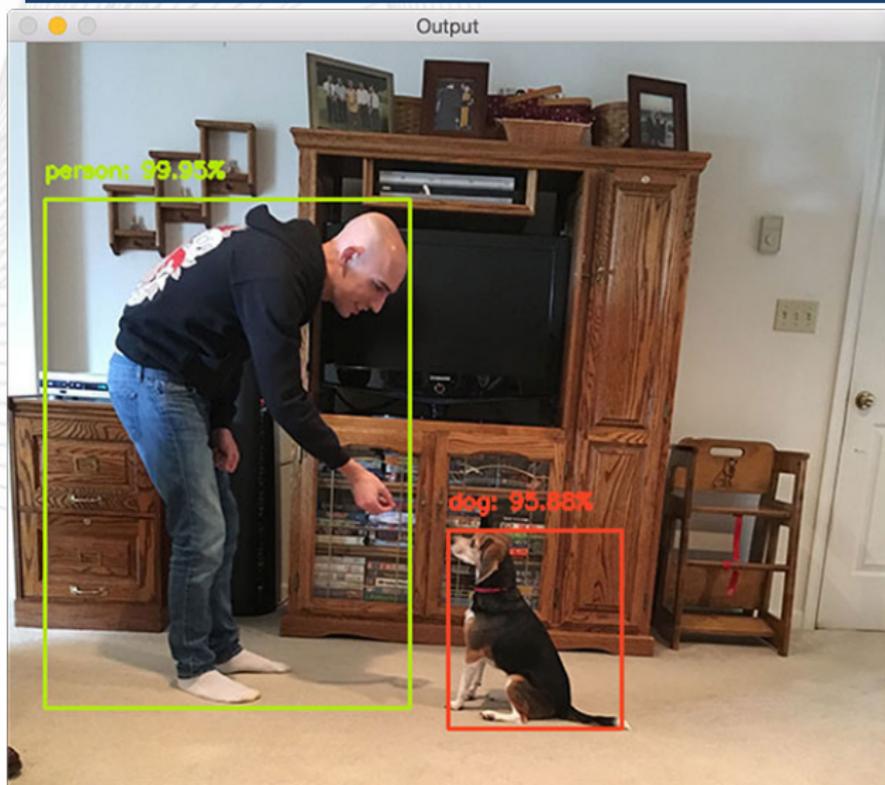


Aplicações





Aplicações





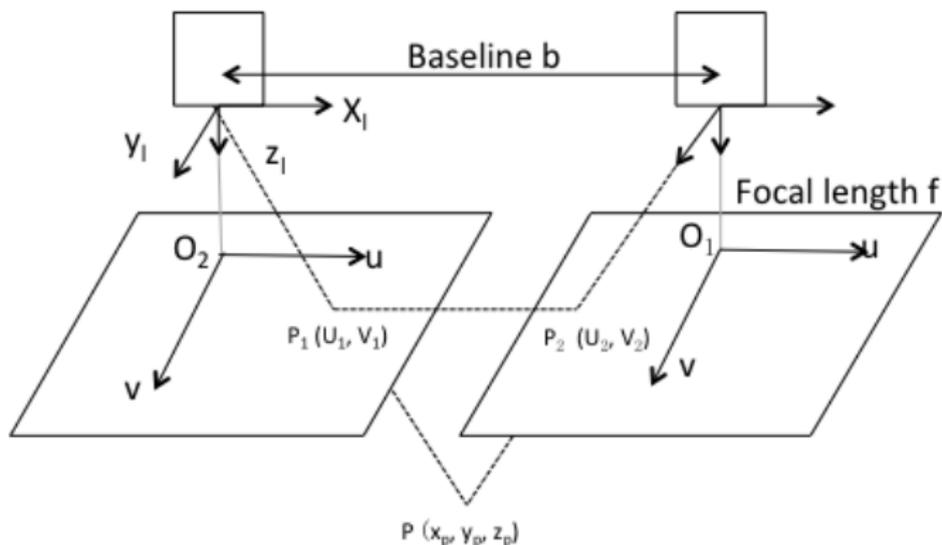
sistema estereoscópico

A recuperação da dimensão perdida a partir de apenas imagens 2D foi o objetivo do estéreo multivista clássico, que foi extensivamente investigado por muitas décadas.

A primeira geração de métodos abordava o problema da perspectiva geométrica; eles focavam na compreensão e formalização, matematicamente, do 3D ao processo de projeção 2D, com o objetivo de elaborar soluções técnicas ou algorítmicas.



sistema estereoscópico





Lei de Gauss

Método determinístico

relacionamos áreas com volume

$$\iiint_V (\nabla \cdot \vec{F}) dV = \oiint_S \vec{F} \cdot \hat{n} dS$$



Lei de Gauss

[da Fontoura Costa, 1999]

Lei de Gauss no processamento e análise de imagens via cálculo numérico de campos vetoriais.



Lei de Gauss

Este artigo descreve uma abordagem numérica para acelerar o cálculo de campos vetoriais em espaços bidimensionais (2D) permitindo a aplicação efetiva da lei de Gauss na análise de imagens e na visão computacional.

A técnica adotada para o cálculo de campos vetoriais consiste na decomposição em dois subcampos escalares, que podem ser obtidos em termos de convoluções da imagem com grandes modelos grandes, efetivamente executada pela rápida transformação de Fourier.



Lei de Gauss

[Homem et al., 2004]

Skeletonization of two-dimensional shapes via fast numerical calculation of vector fields.

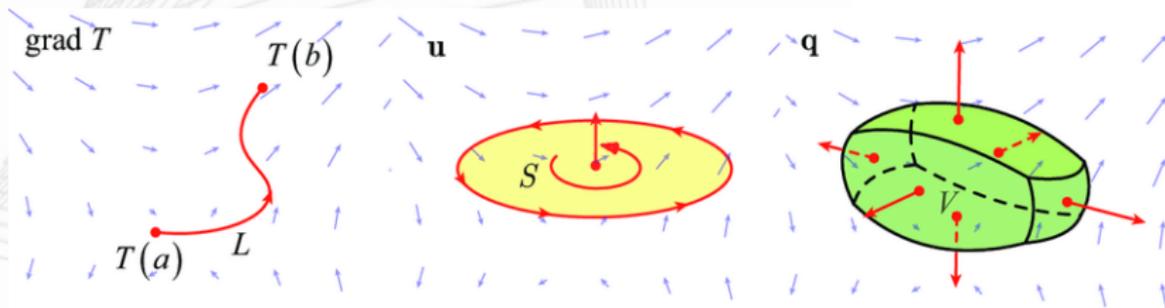


Lei de Gauss

Propomos uma abordagem para a esqueletização bidimensional eficiente de formas binárias através de cálculo numérico de campos vetoriais e curvatura estimativa usando as fórmulas de Weingarten. Além disso, o método proposto demonstrou bom desempenho devido ao fato de que o campo vetorial pode ser fácil e rapidamente calculado usando o Algoritmo Fast Fourier Transform.

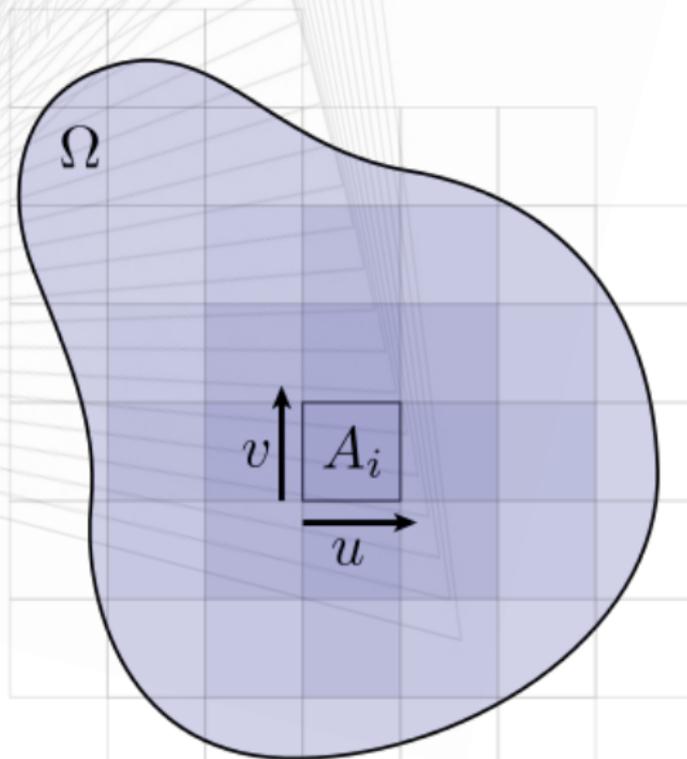


Métodos determinísticos





Métodos determinísticos





Otimização matemática

Seja $I = \{I_k, k = 1, \dots, n\}$ um conjunto de $n \geq 1$ imagens RGB de um ou vários objetos X .



Otimização matemática

A reconstrução 3D pode ser sintetizada como o processo de encontrar o preditor f_{θ} , que pode inferir uma forma X' o mais próximo possível da forma X desconhecida.



Otimização matemática

Em outras palavras a função f_θ é o minimizador da função objetiva de reconstrução $L(I) = d(f_\theta(I), X)$. Onde θ é o conjunto de parametros de f e $d()$ é a medida da distância entre a forma X e a reconstrução da forma $f(I)$

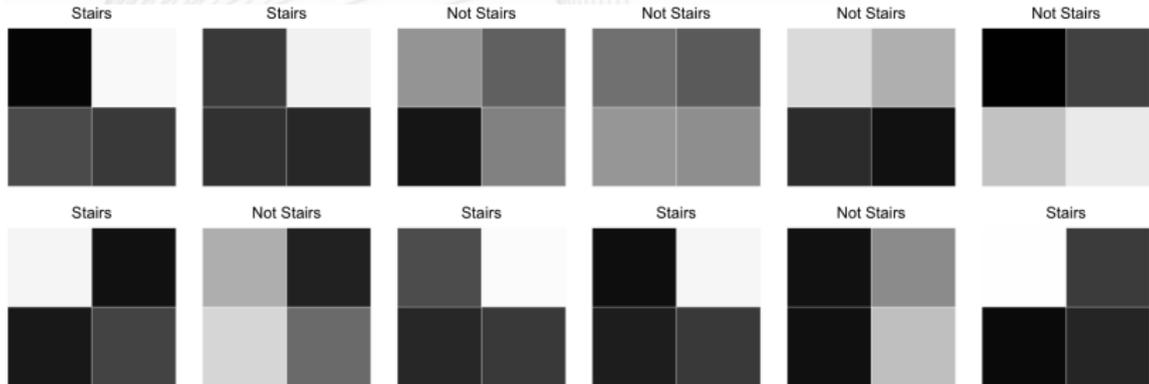


Problema

Temos uma coleção de imagens em escala de cinza 2×2 .
Identificamos cada imagem como tendo uma "escada" como padrão ou não. Aqui está um subconjunto deles.



Problema





Descrição da rede

Nosso problema é de classificação binária. Isso significa que nossa rede pode ter um único nó de saída que prevê a probabilidade de que uma imagem recebida represente escadas.



Descrição da rede

No entanto, optaremos por interpretar o problema como um problema de classificação de várias classes - aquele em que nossa camada de saída possui dois nós que representam “probabilidade de escada” e “probabilidade de outra coisa”.

Isso é desnecessário, mas nos dará uma ideia de como poderíamos estender a tarefa para mais classes. No futuro, podemos querer classificar {“padrão de escadas”, “padrão de piso”, “padrão de teto” ou “outra coisa”}.



Descrição da rede

Nossa medida de sucesso pode ser algo como taxa de precisão, mas para implementar a retropropagação (o procedimento de ajuste), precisamos escolher uma função de perda conveniente e diferenciável, como entropia cruzada.



entropia cruzada

A “entropia cruzada” entre duas distribuições de probabilidade p e q sobre o mesmo conjunto subjacente de eventos mede o número médio de bits necessários para identificar um evento extraído do conjunto se um esquema de codificação usado para o conjunto for otimizado para uma distribuição de probabilidade estimada q , em lugar da distribuição verdadeira p .



Problema

Nosso conjunto de dados de treinamento consiste em imagens em escala de cinza. Cada imagem tem 2 pixels de largura por 2 pixels de altura, cada pixel representando uma intensidade entre 0 (branco) e 255 (preto). Se rotularmos cada intensidade de pixel como p_1 , p_2 , p_3 , p_4 , podemos representar cada imagem como um vetor numérico que pode alimentar nossa rede neural.



fluxo de vídeo

Quando a entrada é um fluxo de vídeo, pode-se explorar a correlação temporal para facilitar a Reconstrução 3D, garantindo que a reconstrução seja suave e consistente em todos os quadros do vídeo corrente.

Além disso, a entrada pode representar um ou vários objetos 3D. objetos pertencentes a categorias de formas conhecidas ou desconhecidas.



Representação

A representação é crucial para a escolha da arquitetura de rede. Ele também afeta a computação eficiência e qualidade da reconstrução.



arquitetura de backbone

Uma variedade de arquiteturas de rede tem sido utilizada para melhorar implementar o preditor f . Uma arquitetura de backbone, que pode ser diferente durante o treinamento e o teste, é composto por um codificador h seguido de um decodificador g , ou seja, $f = g \circ h$.



arquitetura de backbone

O codificador mapeia a entrada em uma variável latente x , referida como um vetor de característica ou um código, usando uma sequência de operações de pooling, seguidas de operações totalmente conectadas a supervisão,



Gerador

O decodificador, também chamado de gerador, decodifica o vetor de característica na saída desejada usando camadas totalmente conectadas ou uma rede de desconvolução (um seqüência de operações de convolução e upsampling, também parâmetros podem ser conhecidos ou desconhecidos)



Instant Meshes

Ferramenta de retopologia 3D



Instant Meshes





Instant Meshes

Apresenta uma nova abordagem para refazer uma superfície em uma malha triangular isotrópica ou quad-dominante usando um operador de suavização local unificado que otimiza as orientações da borda e as posições dos vértices na malha de saída.

O algoritmo produz malhas com alta isotropia, alinhando e ajustando naturalmente as arestas às características nítidas.



OpenMVG (Multiple View Geometry)

OpenMVG (Multiple View Geometry) "open Multiple View Geometry" é uma biblioteca de visão computacional especialmente direcionada à comunidade da Multiple View Geometry. Projetada para fornecer acesso fácil aos algoritmos clássicos de geometria computacional e auxiliar na solução de problemas com precisão.



da Fontoura Costa, L. (1999).

Gauss' law in image processing and analysis via fast numerical calculation of vector fields.

Real-Time Imaging, 5(4):243–251.



Homem, M. R., Costa, L. d. F., and Mascarenhas, N. D. (2004).

Skeletonization of two-dimensional shapes via fast numerical calculation of vector fields.

In Proceedings. 17th Brazilian Symposium on Computer Graphics and Image Processing, pages 106–112. IEEE.