

All about **Sequence Matching and Alignment**

分享人 : yongyuan.name

日期 : 20200622

目录

■ 背景

■ 序列匹配相关知识

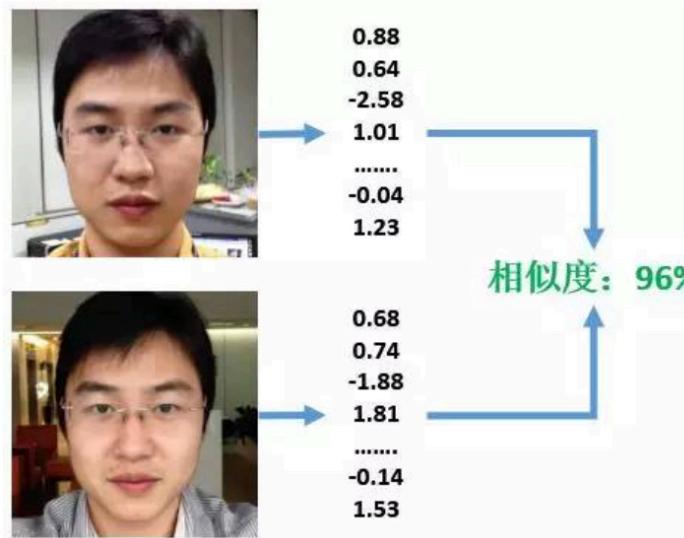
■ SuperGlue

■ 题外话

背景

➤ 单点匹配 , Point matching

应用：被聚合成固定维度的全局特征（ global feature ）



Sentence Vector 1 [0.42,0.84...0.7,0.21]

Sentence Vector 2 [0.43,0.86...0.83,0.62]

背景

➤ 序列匹配 (Sequence matching)

应用：生物序列匹配



指纹匹配

A	G	T	C	C	T	G	A	A	T	G
=	X	X	=	X	=	=	X	=	X	
A	T	C	C	T	T	G	A	G	T	

Mismatch=5

DNA序列匹配

背景

➤ 序列匹配 (Sequence matching)

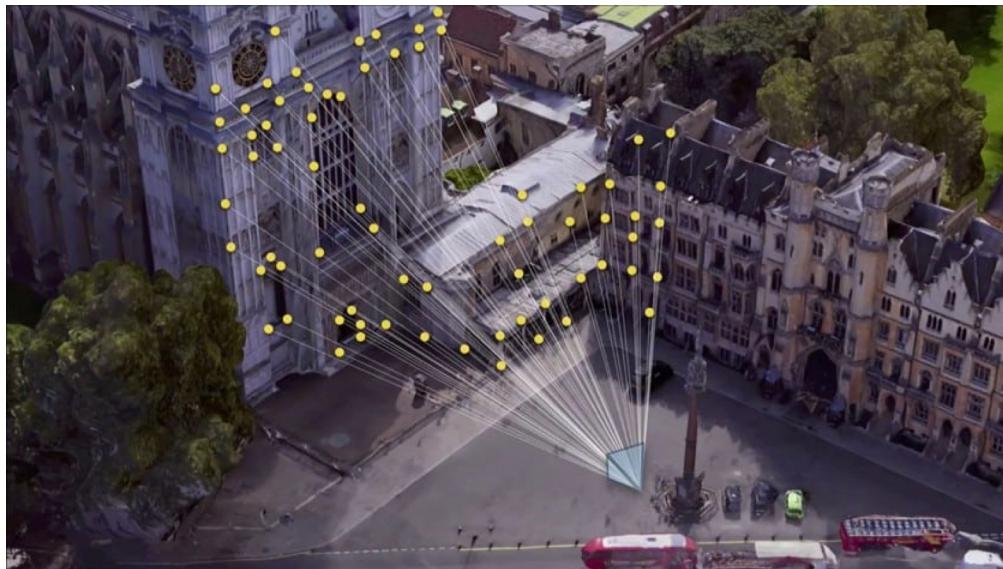
应用：3维重建

SLAM

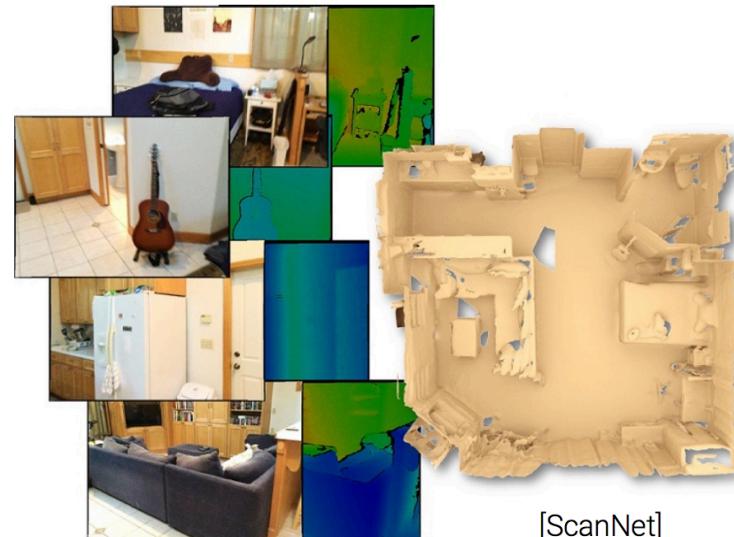
视觉定位系统



[Image Matching Workshop 2020]



Google VPS

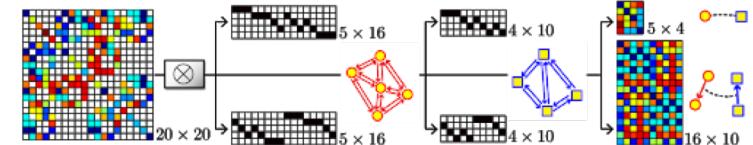
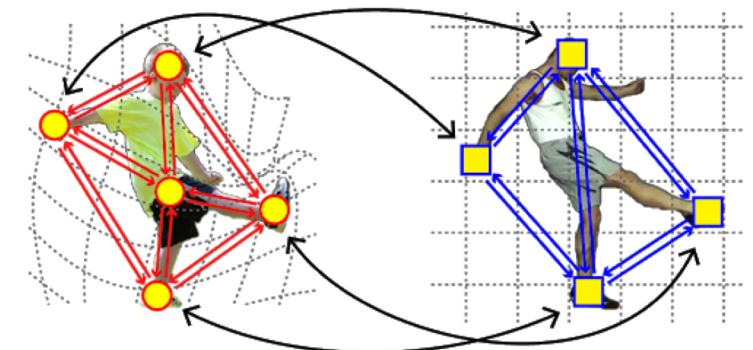
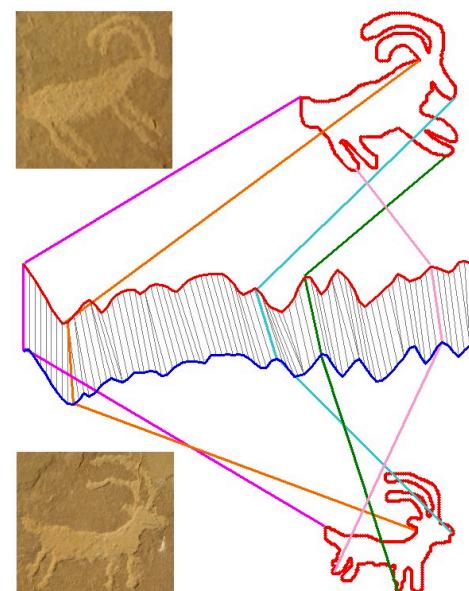
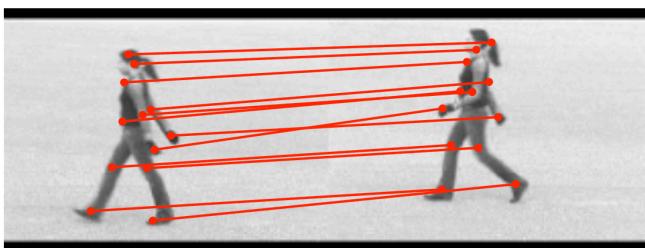
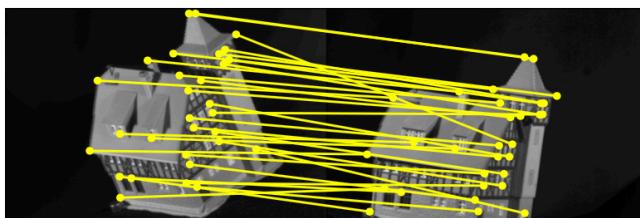


[ScanNet]

背景

➤ 序列匹配 (Sequence matching)

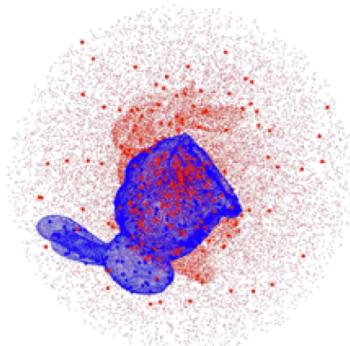
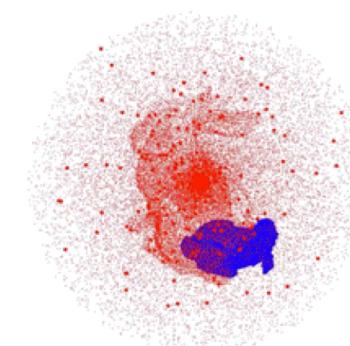
应用：图形、图像、轨迹等匹配



TEASER

RANSAC

Scale Estimation



背景

➤ 序列匹配 (Sequence matching)

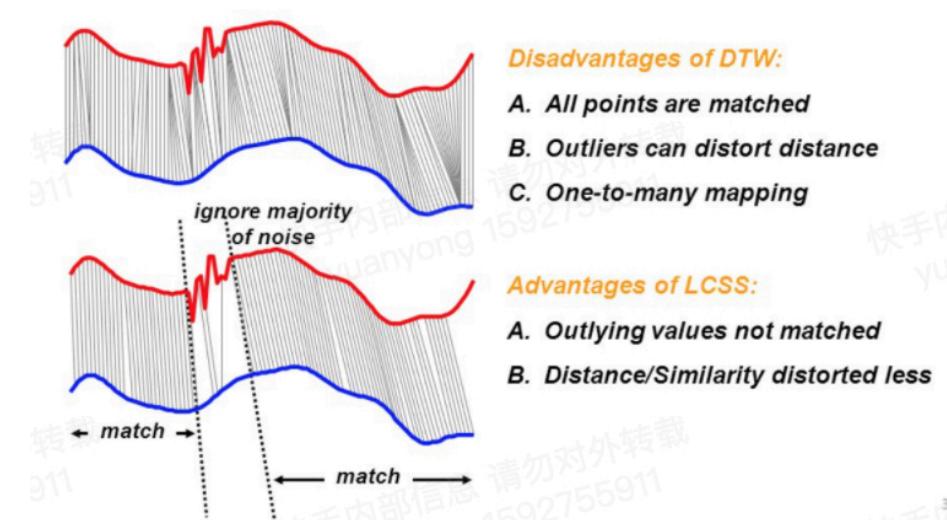
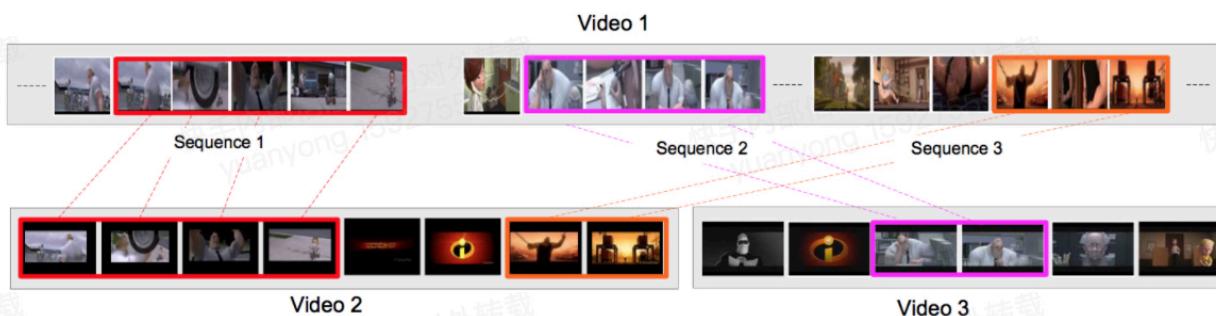
应用：文本序列匹配



背景

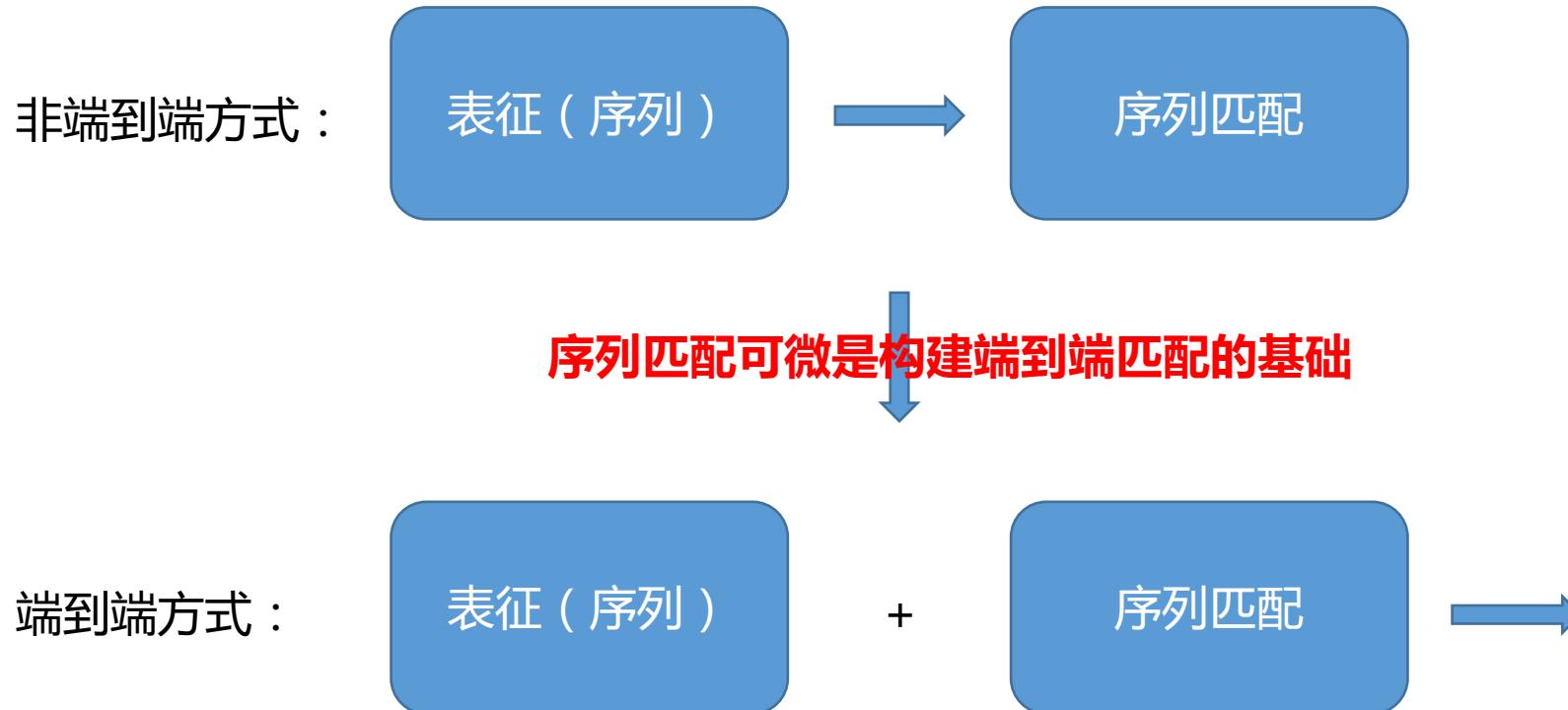
➤ 序列匹配 (Sequence matching)

应用：视频指纹匹配



背景

- 序列匹配 (Sequence matching) : 表征 + 匹配



序列匹配可微是构建端到端匹配的基础

背景

➤ 序列匹配 (Sequence matching) 的优势

优势：序列匹配能更好的反映全局 (global) 中局部 (local) 匹配的情况，具备更强的**物理解释性**

缺点：噪声干扰比较大，对序列匹配算法剔除误匹配要求比较高

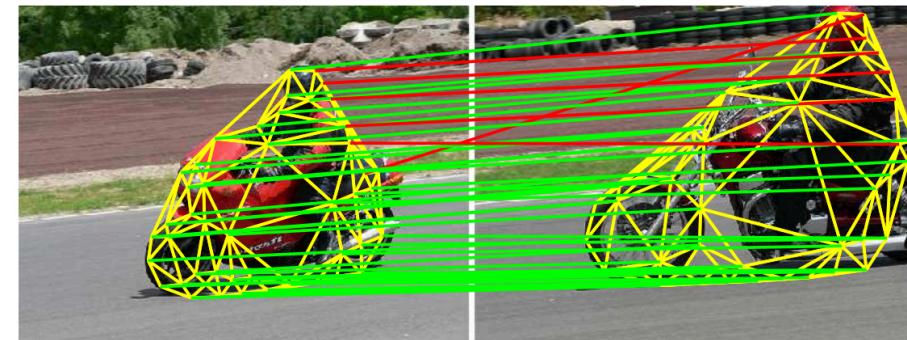
举例：序列匹配，提供局部匹配详情信息，结合先验信息，甚至可以推断出全局的一些特性



背景

➤ 序列匹配 (Sequence matching) 的优势

有意思的发现：在求图像scale的时候，使用graph的边的统计结果，比局部特征自带的scale统计结果，基于graph边的统计结果要更好.



Scale?

背景

➤ Magic Leap

一家神秘、负面消息缠身的公司！



Additional Notes

- We do not intend to release the SuperPoint training or evaluation code, please do not email us to ask for it.
- We do not intend to release the Synthetic Shapes dataset used to bootstrap the SuperPoint training, please do not email us to ask for it.
- We use bi-linear interpolation rather than the bi-cubic interpolation described in the paper to sample the descriptor as it is faster and gave us similar results.

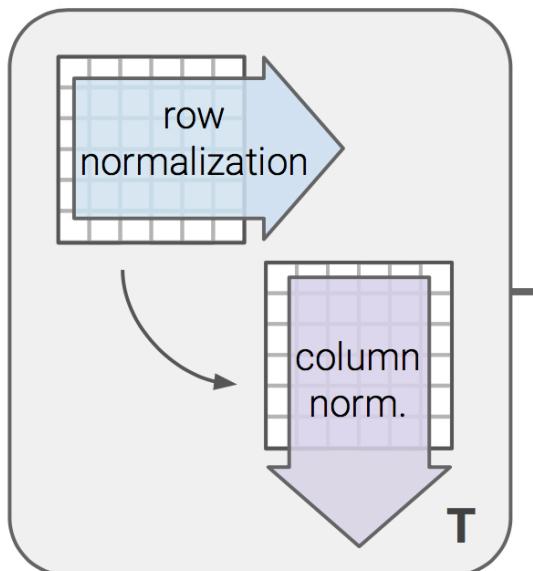
Additional Notes

- For the demo, we found that the keyboard interaction works well with OpenCV 4.1.2.30, older versions were less responsive and the newest version had a [OpenCV bug on Mac](#)
- We generally do not recommend to run SuperPoint+SuperGlue below 160x120 resolution (QQVGA) and above 2000x1500
- We do not intend to release the SuperGlue training code.
- We do not intend to release the SIFT-based SuperGlue models.

相关知识

➤ 背景知识：Optimal Transport最优传输问题

Sinkhorn Algorithm



Sinkhorn算法是一种迭代算法：

- 输入的矩阵交替进行：行归一化、列归一化
- 最终收敛得到一个每行、每列加和均为1的双随机矩阵（doubly stochastic matrix）

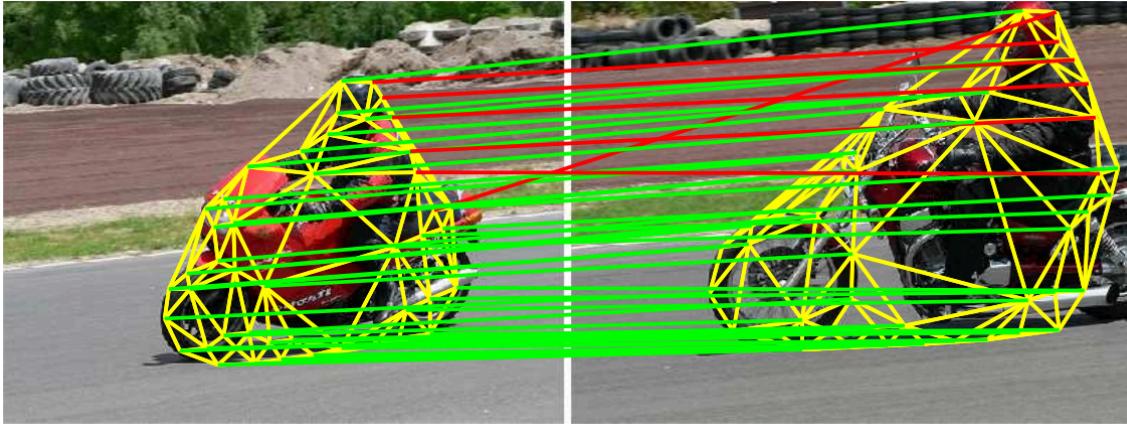
$$\mathbf{P} \in [0, 1]^{M \times N}$$

A small diagram of a 3x3 grid representing a matrix \mathbf{P} . A red arrow points horizontally across the top row with the label "sum ≤ 1". A blue arrow points vertically down the left column with the label "sum ≤ 1". This indicates that each row and each column of the matrix has a sum of 1.

Sinkhorn算法只包含了乘、除操作，**Sinkhorn算法完全可微**，能够被用于端到端的深度学习训练中

相关知识

➤ 背景知识 : graph matching



1. Graph Matching被定义为是一个二次分配问题 (quadratic assignment problems) , NP hard.
2. Caetano在2008将这个问题通过一个浅层模型将该问题简化为一个线性分配最优化问题.

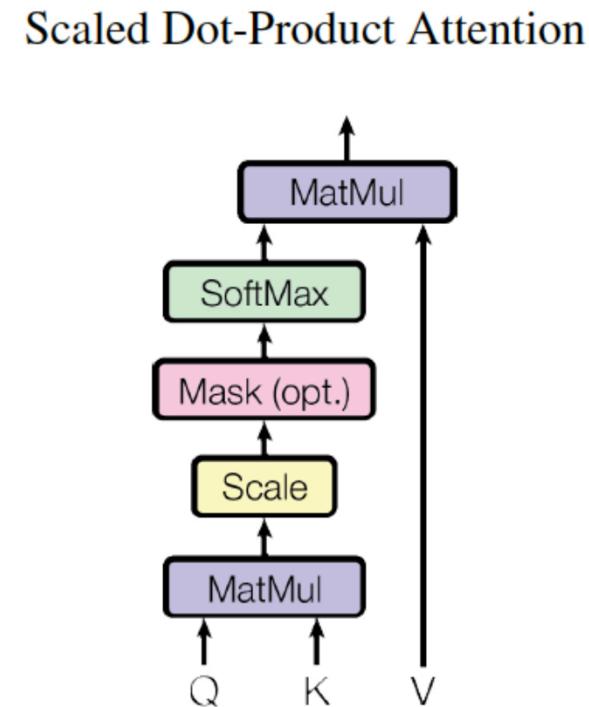
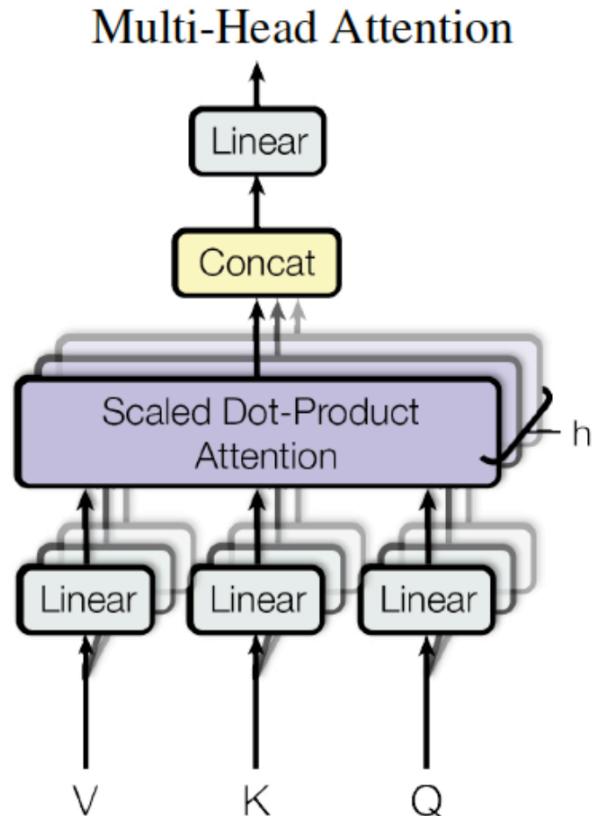
1. https://en.wikipedia.org/wiki/Quadratic_assignment_problem

2. <http://www.sc.ehu.es/acwbecae/ikerkuntza/theses/Ch2.pdf>

3. Learning graph matching. *IEEE TPAMI*, 2009

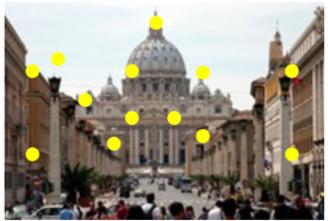
相关知识

➤ 背景知识：multi-head Attention



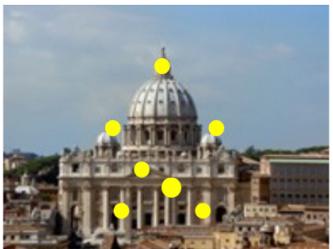
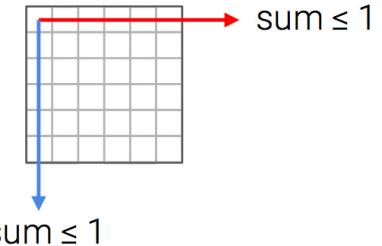
SuperGlue

➤ 问题定义



图像A : keypoints : $(x, y, confidence)_i$,
visual descriptors : d_i

$$\mathbf{P} \in [0, 1]^{M \times N}$$

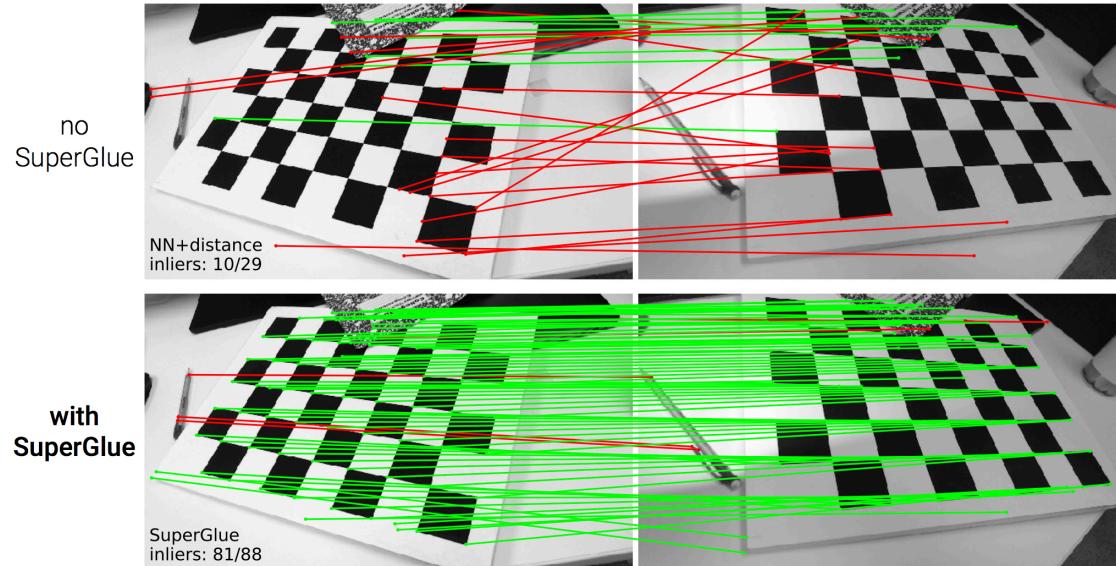


图像B : keypoints : $(x, y, confidence)_j$,
visual descriptors : d_j

soft partial assignment: Single a match per
keypoint + occlusion and noise

SuperGlue

- 先验知识：图像空间的上下文信息（邻近顶点（patch）之间的上下文信息）



人类通过来回浏览两个图像试探性筛选匹配关键点，并进行来回检查（如果不是匹配的特征，**观察一下周围有没有匹配的更好的点**，直到找到匹配点/或没有匹配）。上述过程人们通过主动寻找**上下文来增加特征点特异性**，这样可以排除一些具有奇异性的匹配。

SuperGlue

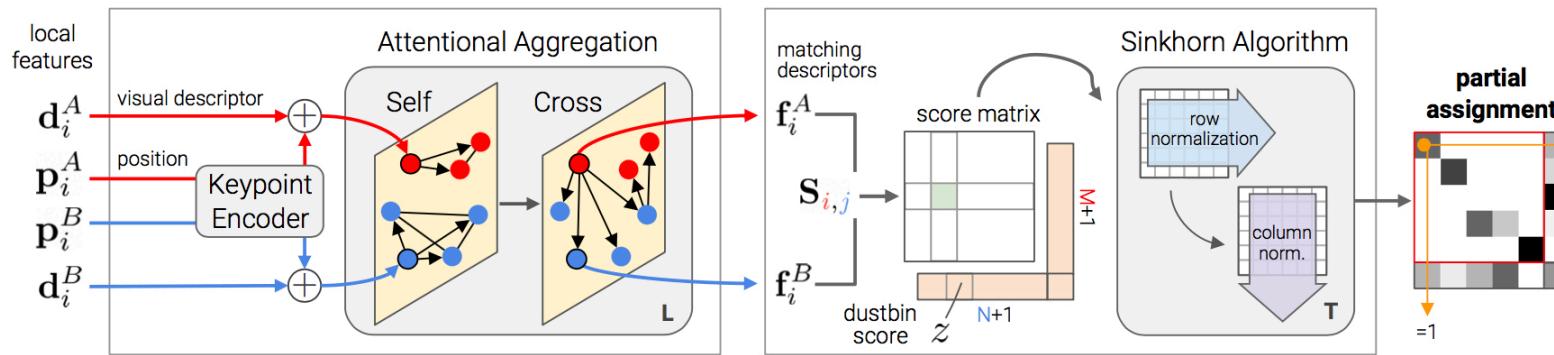
- 先验知识：特征点的位置以及视觉外观能够提高其特异性

$$^{(0)}\mathbf{x}_i = \mathbf{d}_i + \text{MLP}_{\text{enc}}(\mathbf{p}_i)$$

when combined with attention, and is an instance of the “**positional encoder**” popular in language processing.

SuperGlue

➤ 具体方法介绍



A Graph Neural Network
with attention

Encodes **contextual cues** & priors

Reasons about the 3D scene

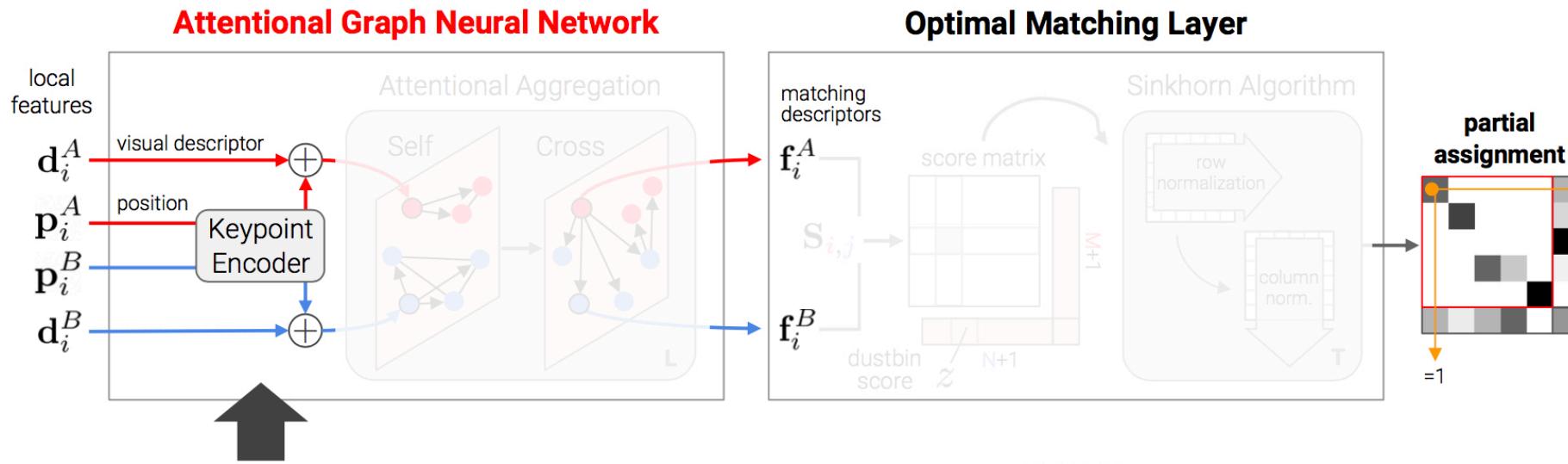
Solving a partial
assignment problem

Differentiable **solver**

Enforces the assignment constraints
= **domain knowledge**

SuperGlue

➤ 具体方法介绍：AGNN



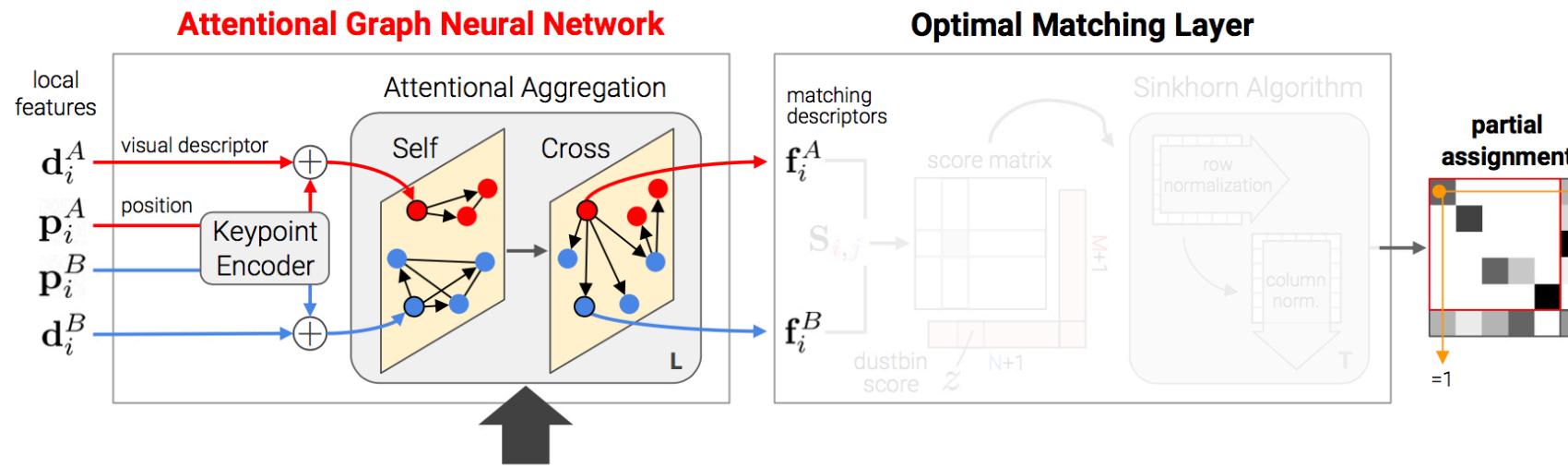
- Initial representation for each keypoints i : ${}^{(0)}\mathbf{x}_i$
- Combines visual appearance and position with an MLP:

$${}^{(0)}\mathbf{x}_i = \mathbf{d}_i + \text{MLP}(\mathbf{p}_i)$$

Multi-Layer Perceptron

SuperGlue

➤ 具体方法介绍：AGNN



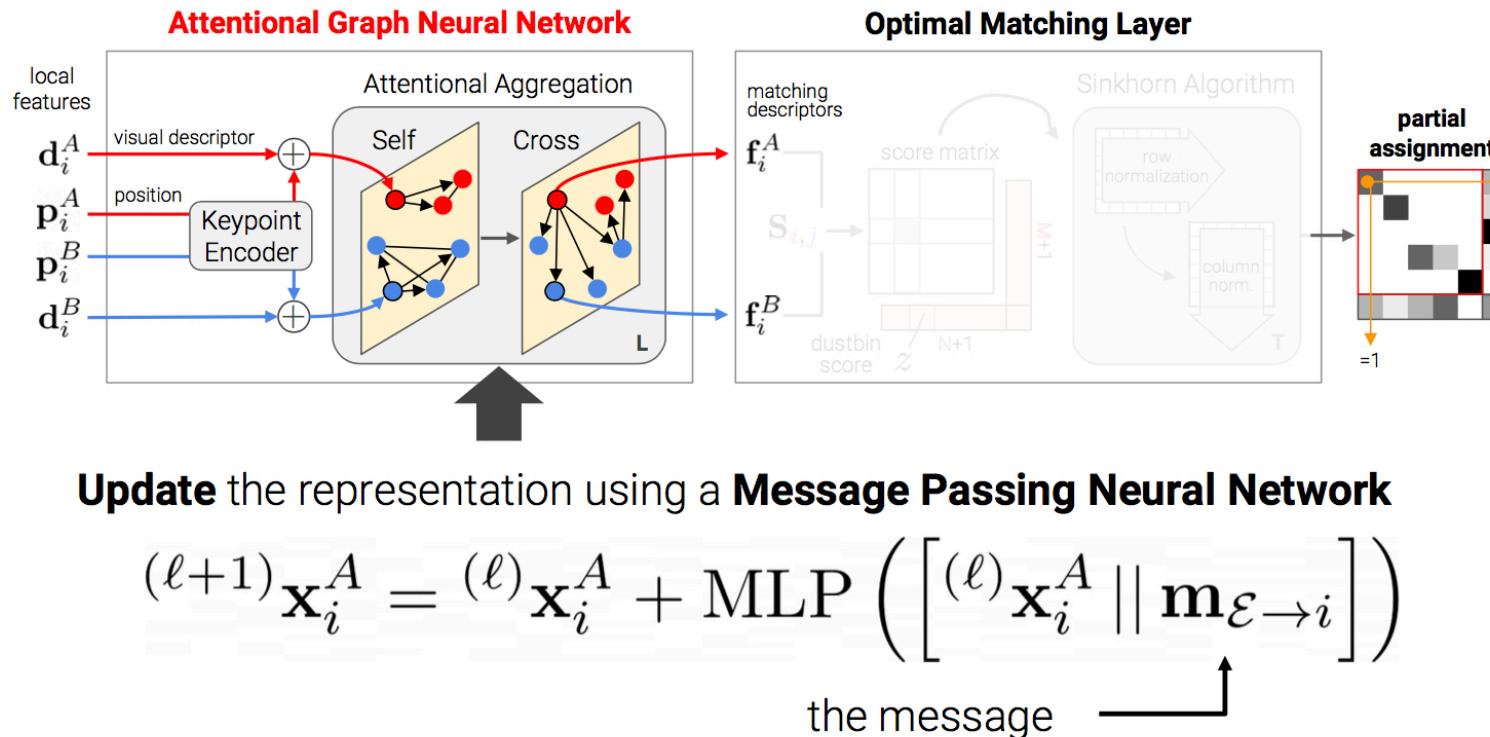
Update the representation based on other keypoints:

- in the same image: “**self**” edges
- in the other image: “**cross**” edges

→ A complete **graph** with two types of edges

SuperGlue

➤ 具体方法介绍：MPNN



$[xx \parallel yy]$, 表示两个结果concat起来

SuperGlue

➤ 具体方法介绍：Attentional Aggregation聚合

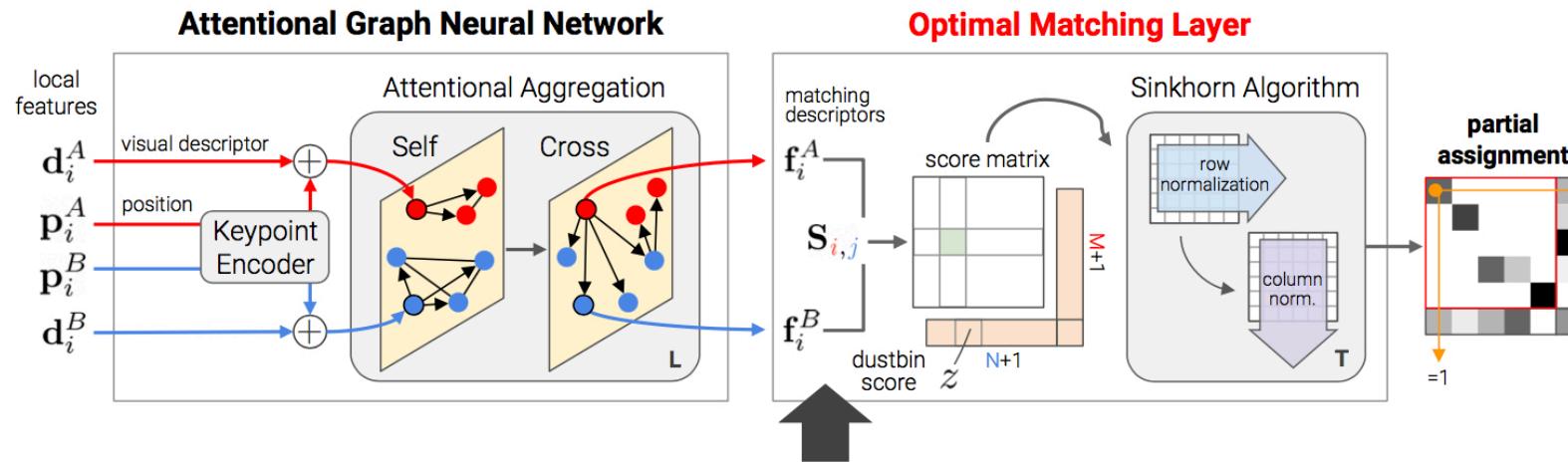
- Compute the **message** $\mathbf{m}_{\mathcal{E} \rightarrow i}$ using **self** and **cross attention**
- Soft database retrieval: query \mathbf{q}_i , key \mathbf{k}_j , and value \mathbf{v}_j

$$\mathbf{m}_{\mathcal{E} \rightarrow i} = \sum_{j:(i,j) \in \mathcal{E}} \alpha_{ij} \mathbf{v}_j \quad \left| \quad \begin{array}{l} \mathbf{q}_i = \mathbf{W}_1^{(\ell)} \mathbf{x}_i + \mathbf{b}_1 \\ \begin{bmatrix} \mathbf{k}_j \\ \mathbf{v}_j \end{bmatrix} = \begin{bmatrix} \mathbf{W}_2 \\ \mathbf{W}_3 \end{bmatrix}^{(\ell)} \mathbf{x}_j + \begin{bmatrix} \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix} \end{array} \right.$$



SuperGlue

➤ 具体方法介绍：OML层



Compute a **score matrix** $\mathbf{S} \in \mathbb{R}^{M \times N}$ for all matches:

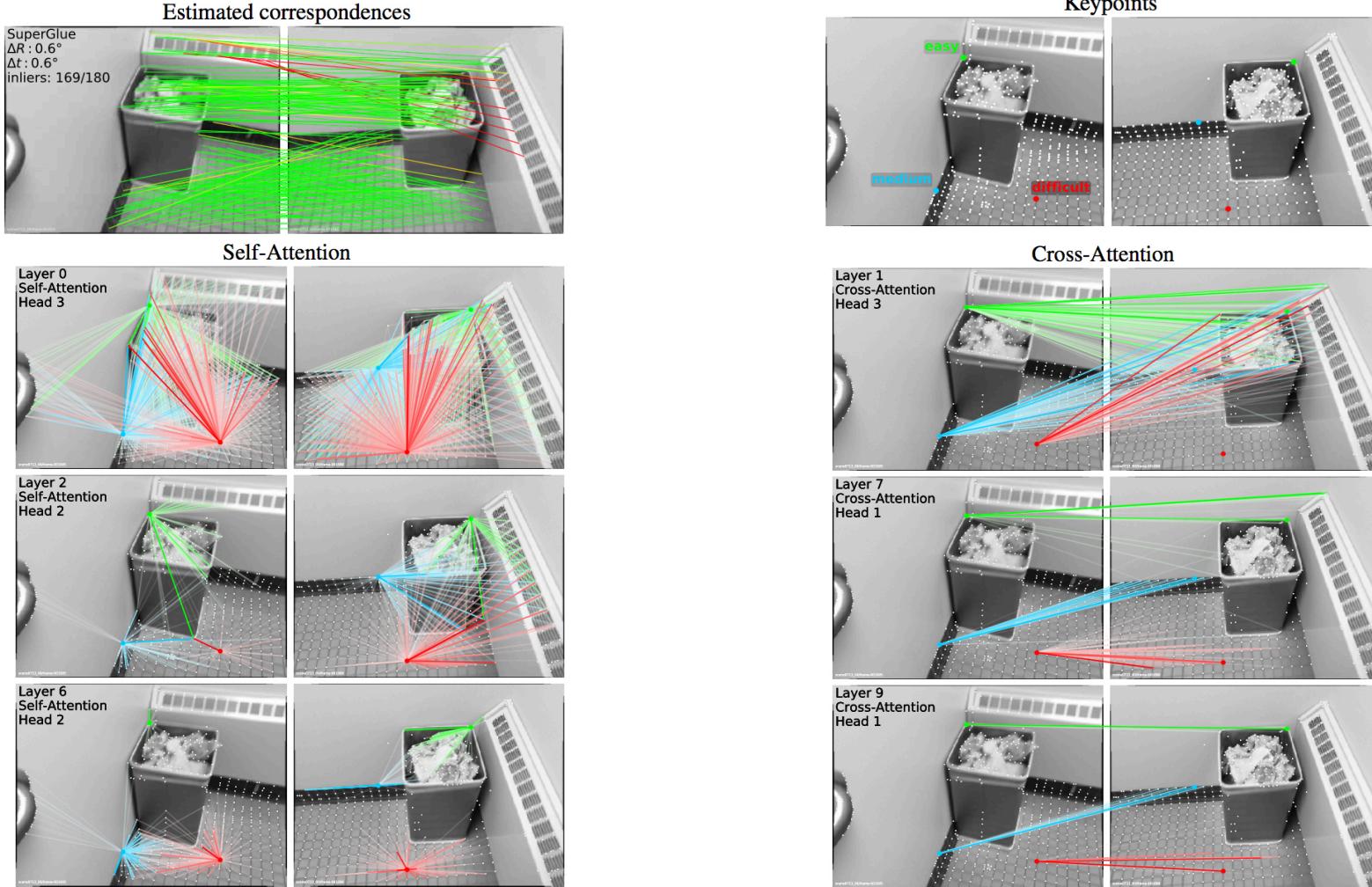
$$\mathbf{f}_i^A = \mathbf{W} \cdot (L) \mathbf{x}_i^A + \mathbf{b}$$

$$S_{i,j} = \langle \mathbf{f}_i^A, \mathbf{f}_j^B \rangle$$

为啥要增加Dustbin score ?

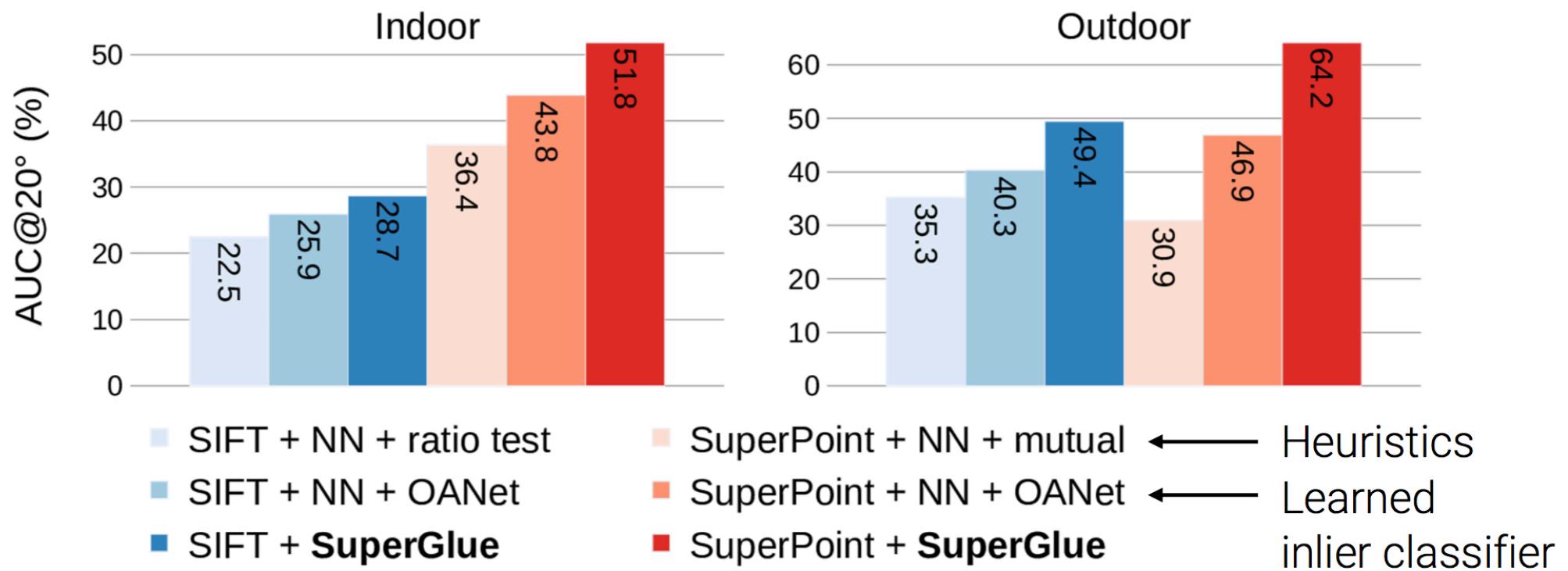
SuperGlue

- Self-edge attention, Cross-edge attention可视化



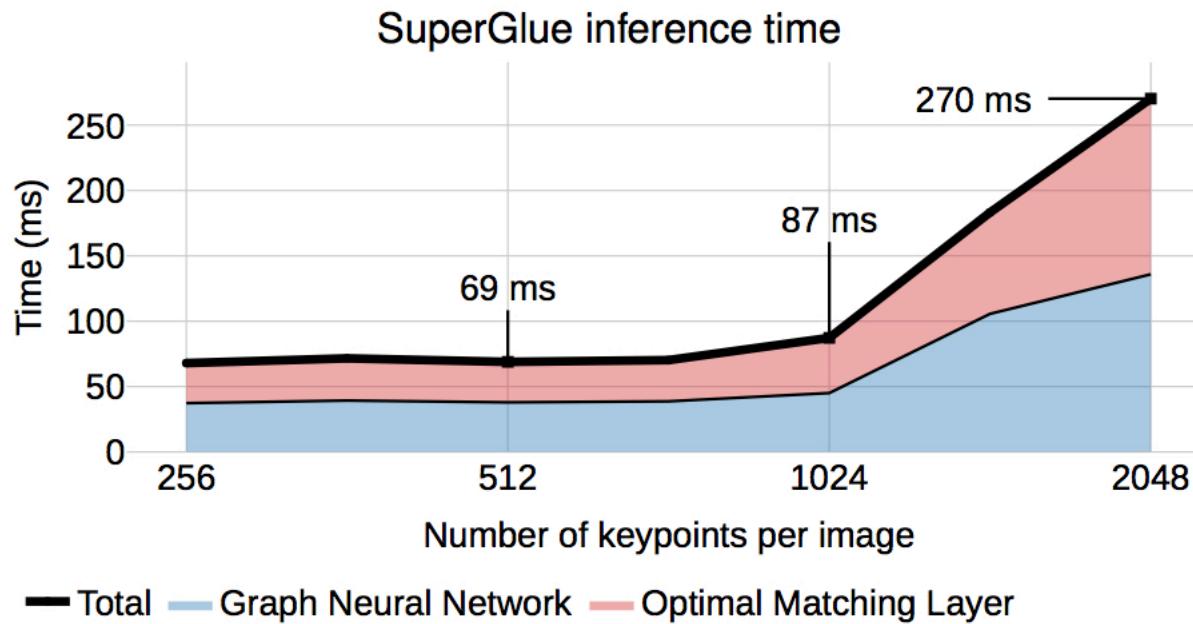
SuperGlue

➤ 实验指标



SuperGlue

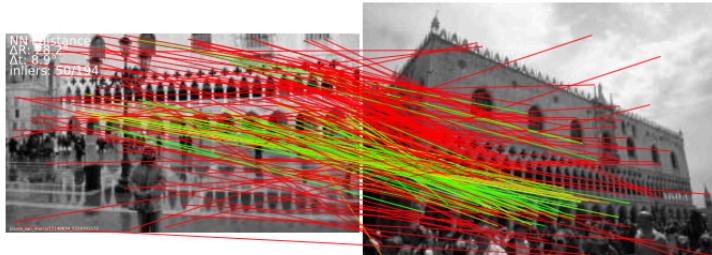
- Inference time 指标



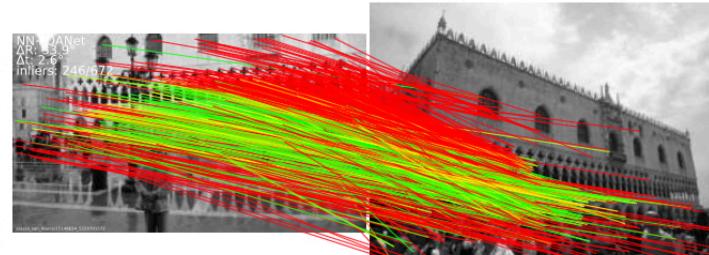
SuperGlue

➤ Feature + Matching

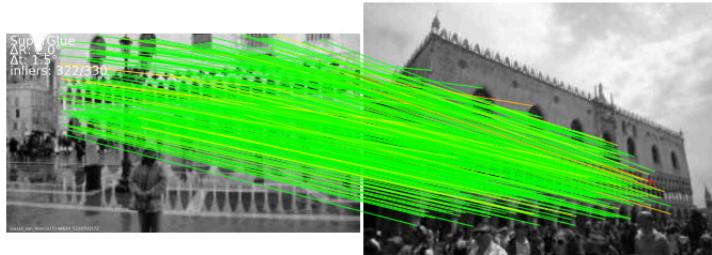
SuperPoint + NN + distance



SuperPoint + NN + OANet



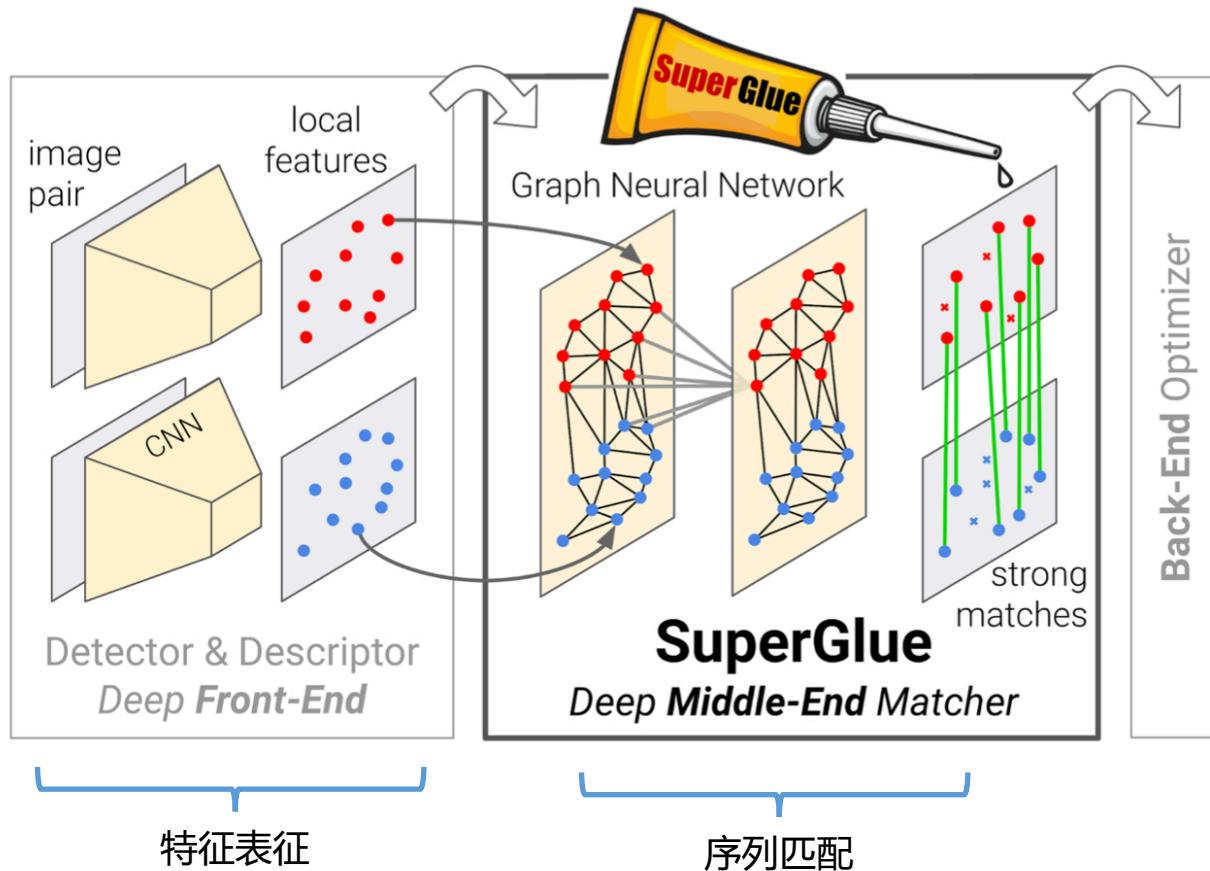
SuperPoint + SuperGlue



题外话

- 端到端特征表征与序列匹配

We believe that, when combined with a deep front-end, SuperGlue is a major milestone towards end-to-end deep SLAM.



END