



原E

2017年

版权声明

代码K

摘：

使用派  
方法

1.介：

请扫描二维码联系客服

webmaster@csdn.net

400-660-0108

QQ客服 客服论坛

关于 · 招聘 · 广告服务 · 网站地图

©2018 CSDN版权所有 京ICP证09002463号

百度提供搜索支持



CSDN APP

经营性网站备案信息

网络110报警服务

中国互联网举报中心

北京互联网违法和不良信息举报中心

## Geometry and Context for Deep Stereo Regression

<http://blog.csdn.net/lvhao92/article/details/72627386>

<http://www.gc-network.com>

积提取上下文的纹理信息。深度值以差分softargmin操作相规范化，没有额外的预处理或者正则化约束。

立体匹配算法应该看重全局，而非局部。比如车窗，如果以整辆车的视角去看，那么就不怎么会出现问题，而如果以局部车窗去看这个物体，就容易出问题。这说明了如何利用端到端的立体匹配模型，从而理解更加广泛的文本信息。这个模型可以学习语义信息和上下文信息。还有差分softmax可以从视差成本向量中生成像素视差值。

## 2.相关工作:

立体匹配几步走：匹配成本计算，成本支撑聚合，视差计算和最优化，视差refinement。

匹配成本就是对应图像位置上像素不相似的衡量值。

1-D correlation)近似于立体成本向量 (dispNet表示不服), 这个成本向量啊, 连接了单张图片的卷积特征和由一系列的卷积跟着。本文算法不降低特征维数, 卷积来传播信息。

本次工作无后处理，无规范化，纯天然。本文工作以3D卷积传递着数据中的信息，不学习可能的分布，成本函数或者分类结果。直接从立体图像对中regression的亚像素估计。

### 3.学习端到端偏差回归

直接端到端从图像到深度图。直接从原始数据中学习到最佳的成本。该方法旨在降低工程设计复杂性。形成我们的模型通过产生一个代表着传统立体匹配元素的**差分层**，这将使得我们能够学习到整个的端到端模型。

### 3.1 一元特征

使用特征来进行立体匹配成本的计算而非原始像素。这个方法可以对光学外表更加鲁邦且可包含局部信息。

本文的深度表示模型由2D卷积构成。每个卷积层后面是归一化层和矫正非线性。卷积核大小为5，步长为2，下采样，为了简化计算。再附加八个残差块，快具有两个3×3的卷积核。如下图所示，将左右立体匹配图穿过这些层之后组成了一元特征。左右双塔共享参数则可以更加有效的学习相关特征。

### 3.2 成本量






使用深度一元特征来进行立体匹配成本的计算，这里介绍了成本量的部分。虽然原始的方法只是简单的组合左右特征，生成成本量是希望该模型能够保留住的框架。每个立体匹配图片，成本量维度为高 $\times$ 宽 $\times$ （最大偏移+1） $\times$ 特征尺度。得到一个4维向量。

很关键，经过这些操作保存了特征维度，不像先前的工作使用一个点乘然后大面积消除特征维度。这让我们可以学习到特征上所包含的内容。我们发现以连构成的成本量提升了提取特征和使用距离尺的性能。我们的目标是保留特征一元，网络有机会学习绝对表示（因为这不是距离尺）并且将其穿过成本量。这学习语义的能力。相反，使用距离尺限制了网络，该网络只能够学习特征之间的关系而不能将绝对特征表示穿过成本量。

### 3.3 学习内容

给定偏差成本向量，我们现在能够学习一个规范化函数，有能力将这个成本量中的内容进行考虑同时改善偏差估计。一元特征之间的匹配成本不可能非常的至当使用一个深度特征表示的时候。比如说，在单一像素的区域（比如天空），这个成本曲线应该是在任何固定局部内容上任何特征上都是非常扁平的。因希望学习规范项同时提升这个成本量的性能。

打算使用三维卷积操作来滤波和改善这个表示。3D卷积是有能力学习长宽深维度的特征表示的。因为我们队每个一元特征都计算了成本曲线，我们能够从中直接学习到卷积滤波器。4.1节中，我们展示了3D滤波器在学习内容和提升立体效果上的重要性。

联系我们		Output Tensor Dim.	
		H×W×C	
1 2 3 4-1 18	<div><div></div><div>请扫描二维码联系客服</div><div><div> webmaster@csdn.net</div><div> 400-660-0108</div><div> QQ客服  客服论坛</div></div></div>	3.1)	
		on) : 7	
		3.2)	
		ction 3.3)	
		ide 2	
19 20 21 22 23 24 25 26 27	<div><div></div><div>经营性网站备案信息</div><div>网络110报警服务</div><div>中国互联网举报中心</div><div>北京互联网违法和不良信息举报中心</div></div> <div>CSDN APP</div>	ide 2	
		ide 2	
		ide 2	
		ide 2	
		ide 2	
28	3-D conv, 3×3×3, 64 features	ide 2	
29	3-D conv, 3×3×3, 64 features		
30	From 27: 3-D conv, 3×3×3, 128 features, stride 2		
31	3-D conv, 3×3×3, 128 features		
32	3-D conv, 3×3×3, 128 features		
33	3×3×3, 3-D transposed conv, 64 features, stride 2		
	add layer 33 and 29 features (residual connection)		
34	3×3×3, 3-D transposed conv, 64 features, stride 2		
	add layer 34 and 26 features (residual connection)		
35	3×3×3, 3-D transposed conv, 64 features, stride 2		
	add layer 35 and 23 features (residual connection)		
36	3×3×3, 3-D transposed conv, 32 features, stride 2		
	add layer 36 and 20 features (residual connection)		
37	3×3×3, 3-D trans conv, 1 feature (no ReLu or BN)		
Soft argmin (section 3.4)			
Soft argmin		H×W	

3D卷积的难点是额外的维度造成计算复杂。深度预测中设计的深度编解码任务应付编码过程中的计算负担，通过下采样特征映射时和编码过程和解码过程样。将该想法扩展为三维。以步长为2来下采样输入，因此以八倍速率降低了3D成本向量的尺寸。我们以4个下采样层来组成我们的3D规范化网络。

一元特征已经以2下采样了，特征总共以32的因子进行下采样。这就让我们以一个很宽泛的视野去明确的控制内容。我们采用两个3×3×3卷积串联进行每程。在原始输入分辨率上进行深度预测，我们采用3D转移卷积去对解码器中的向量进行上采样。整个的结构如上图所示。

当简化计算时，下采样在扩大每个特征的接受域时是非常重要的，然而，当分辨率的损失，同样也会降低空间准确性和精确的纹理细节。出于这样的原因，采样之前添加每个更高的分辨率的特征映射。这些残差层有得到更高频率信息的优势，当上采样特征层以一个更宽的视野提供了一个值得注意的特征映射。

最终，我们采用一个单3D卷积，以步长为2和单特征输出。这一层是非常必要的在原始输入维度上进行密度预测，因为特征以因素2进行下采样，最终的结HWD的大小进行规范化。

3.4 差分argmin

通常来说，立体算法从matching cost unaries产生了最终的cost volume。从成本量中可以以argmin操作估算出成本量的深度偏移量。然而，这个操作有两个其一，这是离散的，不能产生亚像素深度偏差估计

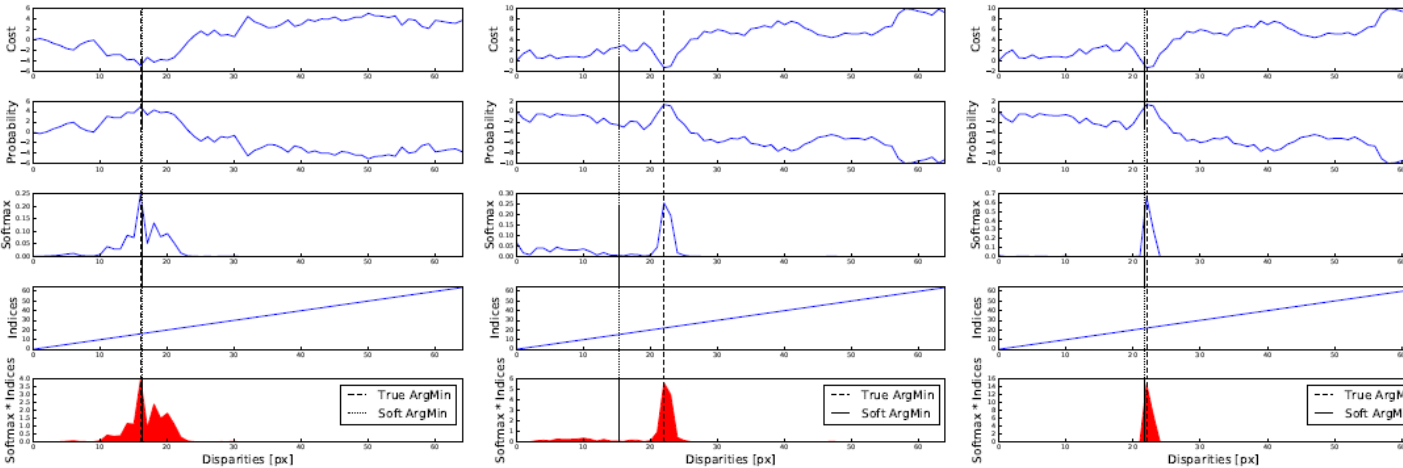
其二，这不是差分的因此不能够通过bp网络来训练（应该是存在梯度，可以后向传播，后向传播不就是传梯度么）。

为了克服这样的限制，定义了soft argmin。这是完全差分的，且有产生平滑深度估计的能力。首先，从cost volumn中提取出每个视差d的预测的成本。接着max操作来对视差维度上可能的向量规范化。以规范化之后的可能性作为权重对每个视差进行加权相加。定义如下：

softargmin: =d× (-Cd)

这个操作完全是差分的，让我们训练同时回归偏差估计。这里，我们展示如何将它应用到立体回归问题上。

然而，与argmiin操作相比，它的输出被所有值影响。这使得它易受多模型分布情况的影响。而不是估算所有模型的权重均值。为了克服这样的限制，我们络的规范性来产生偏差的可能性分布，主要是单峰的。网络也能够预规范匹配成本来控制规范化后softmax可能性的峰值。我们明确的省略最终卷积层的bn从数据中学习它。



(a) Soft ArgMin

(b) Multi-modal distribution

(c) Multi-modal distribution with prescaling

argmin

联系我们

请到真

x可能

3.5 LO

端到端

预测偏



请扫描二维码联系客服

 webmaster@csdn.net

 400-660-0108

 QQ客服  客服论坛

关于 · 招聘 · 广告服务 · 网站地图

©2018 CSDN版权所有 京ICP证09002463号

 百度提供搜索支持



经营性网站备案信息

网络110报警服务

中国互联网举报中心

北京互联网违法和不良信息举报中心

CSDN APP

Input Stereo Images

2D Convolution

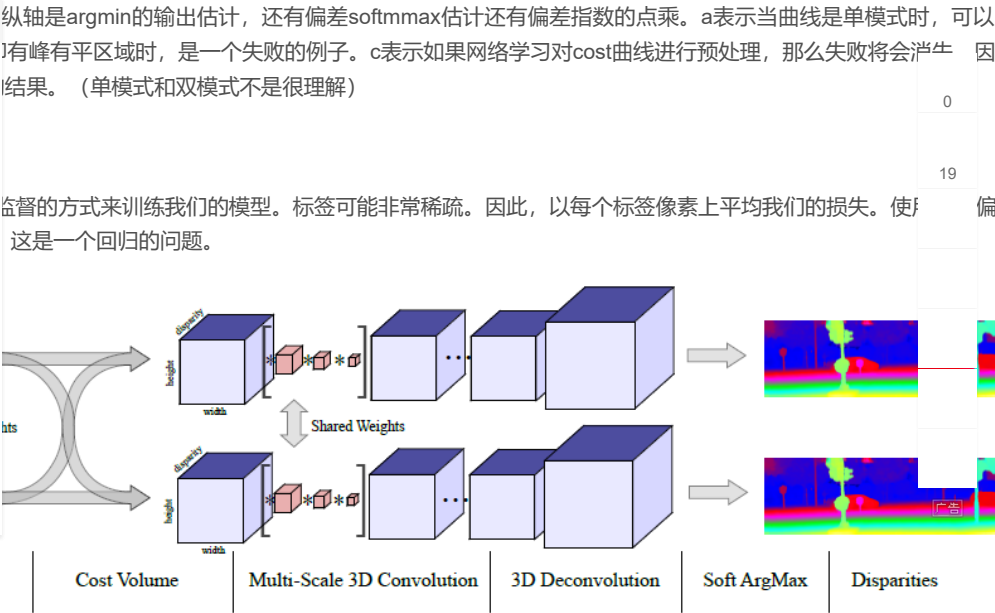
Cost Volume

Multi-Scale 3D Convolution

3D Deconvolution

Soft ArgMax

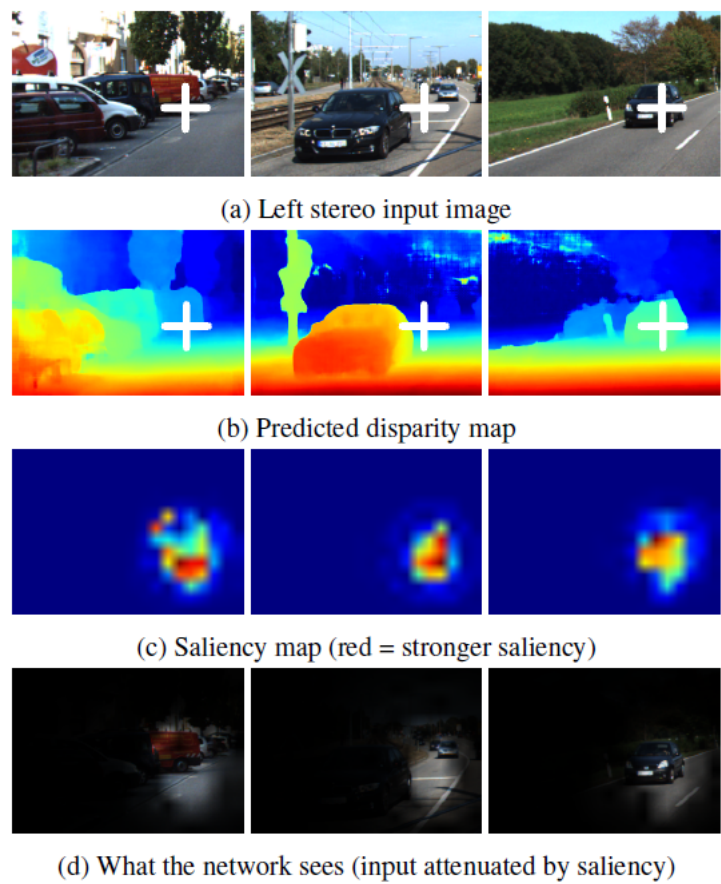
Disparities



### 4.3 model saliency

本文模型能够使用更加宽泛的纹理信息来更好的表示图像局部几何信息。下图，saliency map表示的是针对每个输入像素的输出敏感度。51中的方法暗示了图是occlude the input images的一种函数。通过点的视差来抵消每个图片中的occlusion。

种种迹象表明，给定点的视差预测取决于上下文区域的视野宽度。比如，车辆前面的视差取决于车辆的输入像素和它下方的路面。这表明本文模型有能力理文本信息，而不是像之前深度学习方法只预测9×9局部块信息。



这个图表明对于所选的输出像素（白叉）有效的接受区域，这表明本文算法有能力以一个更加宽泛的视野范围和更加重要有效的图像上下文信息去回归拟合视差，这远比局部几何和外形信息要好。比如，我们观察到本文模型就会考虑车辆的周边信息及路面信息，一起来估计视差。

总结：因为需要端到端网络，因此成本量阶段不光获取两特征之间的关系，还需要将特征传下去，因此维度比较大。四维，为了降低计算负担，采用的是编解码操作。上采样的时候到编码时候的特征层，就跟flownet的时候一样。由于成本量是四维的，所以卷积的时候用的是三维卷积，我去，三维卷积。

与flownet相比，同样是端到端结构，可人家是三维的。因为flownet通过concatenate这个二图的相关信息和特征信息，而没有额外为了特征信息而多出来一维。最后的翻卷积，t时候也是不一样的。flownet每层都进行了深度图的预测。而该文就没有，只是利用编码时候产生的特征层还有解码时候的前一层。效果谁优谁劣，不得而知。这个模型应该是比杂点的。

与dispnet相比，文中解释，dispnet首先是在光流上面预训练而成的端到端回归网络，本文算法比dispnet好在notable margin for all test pixels. dispnet用一relation层近似得到立体匹配的cost volume。相反，本文算法使用3D卷积和soft argmin层得到一个更加完备的cost volume从而可以充分利用图像几何信息。spnet要优秀的。

个网络的视野范围确实宽泛。也许是会比dispnet要好，可惜就是没有代码。

登录

注册

×

联系我们



u



E




属



h

关于 · 招聘 · 广告服务 · 网站地图

©2018 CSDN版权所有 京ICP证09002463号

 百度提供搜索支持



经营性网站备案信息

网络110报警服务

中国互联网举报中心


北京互联网违法和不良信息举报中心

CSDN APP

请扫描二维码联系客服

 webmaster@csdn.net

 400-660-0108

 QQ客服  客服论坛

想对作者说点什么？

我来说一句

post volume 里面不都是特征图吗 (07-19 12:51 #6楼)

的维度是1/2D\*1/2H\*1/2W\*2F 这里的1/2D应该怎么理解呢？ (05-16 17:23 #5楼) [查看回复\(5\)](#)


ure map是怎么打包成4D的吗？ 最大偏移+1指的是什么？ (04-24 17:59 #4楼) [查看回复\(1\)](#)

eature map是怎么打包成4D的吗？ 最大偏移+1指的是什么？ (01-22 04:01 #3楼) [查看回复\(2\)](#)

[查看 19 条热评](#)



Geometry and Context for ...   306

，并采用回归的方法去预测视差值，不再使用传统的成...

Geometry and Context for De...   639

端到端学习几何和背景的深度立体回归 我们提出了一个新颖的深度学习体系结构，用于求解矫正后的立体图像...

深度学习+几何结构：1+1>2?

  1465

转载： [https://mp.weixin.qq.com/s?\\_\\_biz=MzIxOTczOTM4NA==&mid=2247484275&idx=1&sn=ab9208a8ec01...](https://mp.weixin.qq.com/s?__biz=MzIxOTczOTM4NA==&mid=2247484275&idx=1&sn=ab9208a8ec01...)

EdgeStereo: A Context Integrated Residual Pyramid Network for S...

  501



3 Approach 先介绍原始的CP-RPN，接着介绍端到端多任务的EdgeStereo，包括HED网络以及CP-RPN与HED...

Cross-Scale Cost Aggregation for Stereo Matching

  457


摘： 人类根据不同的尺度来处理立体匹配等问题。然而，这种灵感被现如今的成本聚合方法当中所忽视了。本...

Classification and evaluation of cost aggregation methods for ster...

  347


2.成本聚合阶段的识别 尽管一些包含可变支撑的方法限制了成本聚合步骤，只以固定的权重在固定的正方形窗口...

【深度学习论文】Dispnet

  685


【深度学习论文】Dispnet

论文阅读《ActiveStereoNet:End-to-End Self-Supervised Learning fo...

  96



本文出自谷歌与普林斯顿大学研究人员之手并发表于计算机视觉顶会ECCV2018。本文首次提出了应用于主动双...

论文笔记《End-to-End Deep Learning for Person Search》

  2657

这篇文章是香港中文大学的郑晓刚于2016提出，本文是对文中的重要工作所做的一份梳理。 论文链接 <http://ww...>

论文笔记《End-to-End Training of Hybrid CNN-CRF Models for Ster...

  515

论文作者提供的源码 <https://github.com/VLOGroup> 摘要： 1. 介绍 2. 相关工作 3. CNN-CRF 模型 3.1 Unary CN...

CVPR2017论文



CVPR2017论文 <http://openaccess.thecvf.com/menu.py> Computer Vision Foundation open access These re...

立体匹配之（二）：[MC-CNN] 2015CVPR: Stereo Matching by Training a Convolutional Neural Netw




1 摘要 project主页：<https://github.com/jzbontar/mc-cnn> 基于patch的提取与比较，学习其相似性得到一个matching cost，并将正确匹配的patch...

立体匹配综述阅读心得之Classification and evaluation of cost aggregation methods for stereo correspondence



学习笔记之基于代价聚合算法的分类，主要针对cost aggregation分类，2008 CVPR 1. Introduction 经典的全局算法有：BP[31...

个人资料



lvhhh

关注

原创58

粉丝61

喜欢22

评论90

等级： 博客 4

积分： 1217

访问： 5万+

排名： 4万+

恒

文章

ActiveStereoNet: End-to-End Self-Supervised Learning for Active Stereo Systems

登录

注册

×

Learning for Sem

Runtime st match

ngleton

Pyramid

Adam那(1)

联系我们

请扫描二维码联系客服  
✉ webmaster@csdn.net  
☎ 400-660-0108  
🗨 QQ客服 🗨 客服论坛

关于 · 招聘 · 广告服务 · 网站地图  
©2018 CSDN版权所有 京ICP证09002463号  
🔍 百度提供搜索支持

CSDN APP

经营性网站备案信息  
网络110报警服务  
中国互联网举报中心  
北京互联网违法和不良信息举报中心

个人分类

机器学习

数据结构

python

笔试题

深度学习

3篇

展开

归档	
2018年8月	1篇
2018年4月	1篇
2018年3月	6篇
2018年1月	2篇
2017年12月	1篇
展开	

热门文章

阿里天池比赛快速入门

阅读量：3959

FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks

阅读量：3128

支持向量机-SVM

阅读量：2722

争取几句话描述一下爬山法，模拟退火，遗传算法

阅读量：2297

集成学习1-Boosting

阅读量：2137

最新评论

Pyramid Stereo Ma...

u010277444: 您好，没明白怎样将cost volume回归成视差

End-to-End Learni...

u010277444: 从cost volume 怎样回归成视差，cost volume 里面不都是特征图吗

RuntimeError: The...

weixin\_41236528: [reply]miluzhiyu[/reply] 请问您这个问题解决了么

RuntimeError: The...

weixin\_41236528: [reply]qq\_40715394[/reply] 请问您这个问题解决了么？

RuntimeError: The...

weixin\_41236528: 请问您这个问题解决了么？

0

19

↑

🛡

登录

注册

×