

## Accurate, Dense, and Robust Multi-View Stereopsis论文分析与代码实现（一）



# Accurate, Dense, and Robust Multi-View Stereopsis论文分析与代码实现（一）

本文版权属于重庆大学计算机学院刘骥，禁止转载

[Accurate, Dense, and Robust Multi-View Stereopsis论文分析与代码实现（一）](#)

[前言](#)

[论文写作背景](#)

[题外话](#)

[摘要](#)

[引言和相关工作](#)

[文献检索方法](#)

## 前言

本文依据论文[Accurate, Dense, and Robust Multi-View Stereopsis](#)进行写作，分析该论文的各种算法，给出具体的实现方法。Accurate, Dense, and Robust Multi-View Stereopsis这篇论文的第一作者Yasutaka Furukawa（中村雄辉）现在是华盛顿大学（Washington University）的助理教授（Assistant Professor），曾经在Google等IT企业任职。有兴趣的应该看看他的个人主页。总之他是一个牛人。Accurate, Dense, and Robust Multi-View Stereopsis是2007年发表在Computer Vision and Pattern Recognition期刊上的论文，按照时间推测，应该是Yasutaka Furukawa博士期间的工作。这篇论文对应他编写的PMVS2软件。PMVS2软件是少数几个在工业领域成功应用的多视图三维重建软件。

鄙人写作本文的目的主要包括：

1. 向学生介绍（主要是鄙人的研究生）10年前MVS（Multi-view Stereo）方面的重要方法，详细分析算法流程，给出代码实现，帮助他们在学習过程中少走弯路。
2. 通过讲解这篇论文，阐述一些学习和研究的方法。

重庆大学的研究生大体上分层3个层次：看不懂论文、看得懂论文但不知道如何实现、看得懂论文也知道如何实现。鄙人见得最多的是第一个层次。每年研究生招生的时候，问得比较多的一个问题是：“阁下如何规划研究生阶段的学习”。一般的回答都是这样的：“我准备提高自己的动手能力，我不喜欢做理论研究”，稍微愣头青一点的会说：“我不喜欢看论文，我喜欢做项目”（遇到这种学生心都凉半截）。每年这个时节也能在电梯中遇到很多将要毕业的研究生。大体的对话有这些：“论文才写了3000字”、“介绍一个凑字数的办法”、“查重还没有过”。鄙人可是有证据的：



总体来看，尽管研究生是要求写论文的（即便是专硕也是要写毕业论文的），但是论文似乎总是处于研究生学习的对立面。甚至有些学生会认为，如果老师让他们读论文，那是在浪费他们的青春。青春是什么？青春应该是到企业去实习，去做项目。持这种看法的二年级同学特别多。

本想论述一下研究生教育，写了一部分，鄙人就删除了，因为与鄙人的身份和地位不符。鄙人还是谈论现象吧。作为研究生导师，鄙人最喜欢的学生当然是“看得懂论文也知道如何实现”，在社会上发展得比较好的也是“看得懂论文也知道如何实现”的这类学生。我们可以很清楚的发现这样一种现象：**没有学生是懂得实现但看不懂论文**。或者反向推理：**看不懂论文的学生，基本也不懂实现**。

现在的企业招聘搞得很复杂，其实只要问一下研究生是做什么方向的，然后找一篇相关的paper让他去读，看得懂的一般都是技术靠谱的，看不懂的一般技术都有问题，至少后期的发展恐怕很有限。一流企业都很重视研究生在校期间发表的论文水平，也在于此。

Yasutaka Furukawa在校期间就是优秀的研究生，他做的软件有300多MB的源代码，他发表的论文是世界顶级的。阁下说他在搞理论吗？阁下写的程序有300多MB的源代码？阁下做过的项目有300多MB的源代码吗？

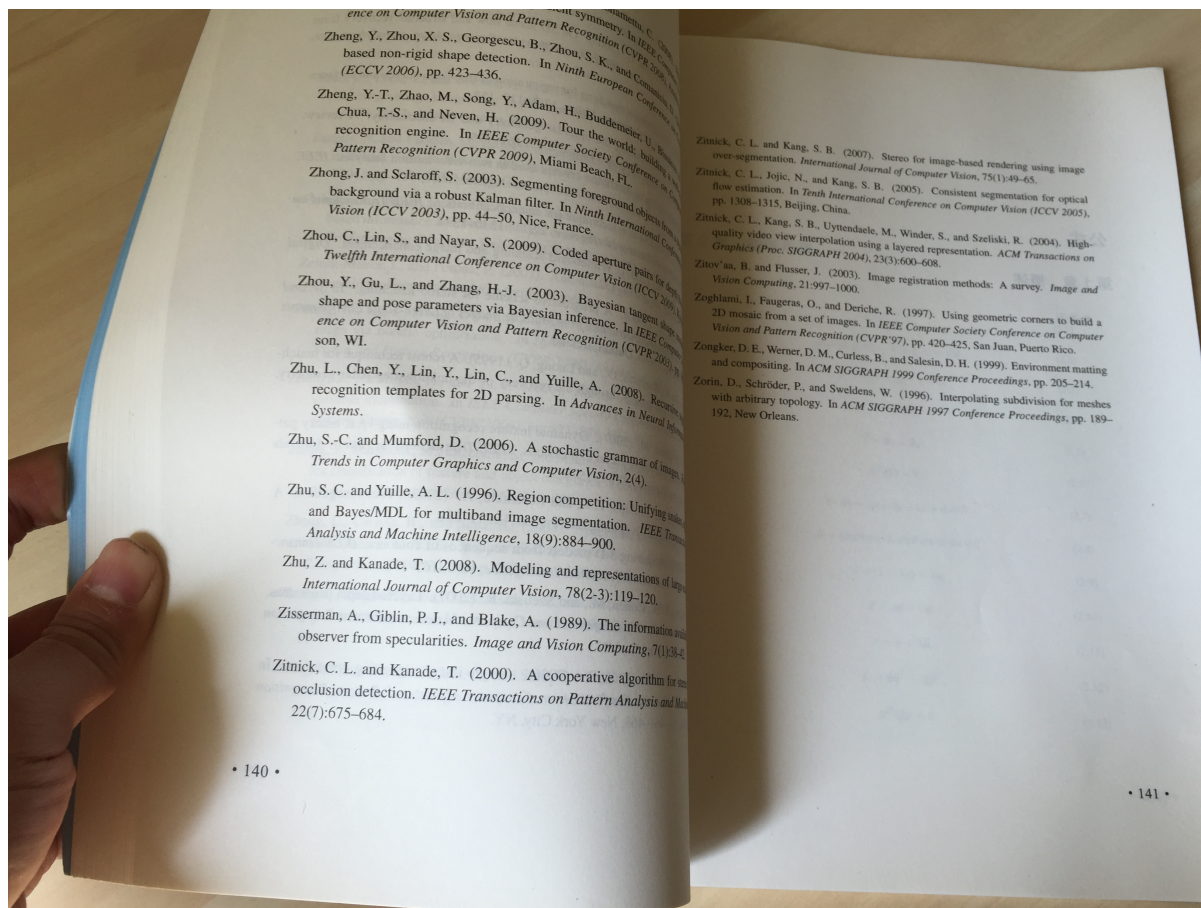
闲话少说，切入正题吧。鄙人这篇文档需要分成好几个部分才能写完，在此期间，请下载[Accurate, Dense, and Robust Multi-View Stereopsis](#)，编译安装PMVS2（很多人直接就会卡在这里吧）。这个文档的写作会花费数周的时间，鄙人会将程序代码发到[coding.net\(https://git.coding.net/liujiboy/Furukawa2006.git\)](https://git.coding.net/liujiboy/Furukawa2006.git)上面。

## 论文写作背景

论文Accurate, Dense, and Robust Multi-View Stereopsis的研究领域属于[三维重建 \(3D Reconstruction\)](#)。三维重建有很多方法，其中一种是[利用多张图像进行三维重建](#)，最近称为MVS (Multi-View Stereo)。这种重建成本低、设备简单、非接触、无伤害，因此在娱乐、CAD、文物保护等领域有广泛的应用。Google地图的街景和Google地球的3D模型都应用了这种技术。介绍一个比较有趣的应用——[Photo tourism](#)，效果非常棒！过去10年，这个领域的发展相当迅速。最近几年随着无人机航拍器的发展，越来越多的人开始将这种技术用在航拍数据的重建。例如香港科技大学权龙教授的团队研发了[三维地图平台](#)。国内的“[第三届全国研究生智慧城市技术与创意设计大赛](#)”中有一个题目就是无人机航拍数据的三维重建。[明德学院](#)提供了一个在线的算法测评，可以在这里查到过去几十年比较有影响力的算法（相关领域最早追溯到上个世纪70年代）。[牛津大学](#)也提供了一些重建的数据。另外还有[斯坦福大学](#)。国内清华、北大、浙大、北航、中科院的团队也在做这方面。当然重庆大学，据鄙人所知，就鄙人一个人在做。纯粹属于个人爱好，也没有什么成果，所以鄙人也属于学习阶段。多扯几句，当初学计算机，就是为了学三维重建，因为很喜欢做游戏。后来莫名其妙就开始做Web和云计算，再后来机缘巧合又开始搞“植物重建”。“植物重建”倒是有些成果（开发了一个完整的重建平台），但是毕竟离主流还有些远（也不通用）。Web倒是做成了精通水平。最近居然机缘巧合将Web和三维重建结合起来，做了一个桥梁健康监测系统。再扯远一点，做这些系统的确是可以赚钱的，但是鄙人更喜欢做PMVS2一样的通用软件。为

什么呢？因为做一套就可以卖很多人（如果不缺钱像PMVS2一样开源也不错）。

拉回现状描述来。Yasutaka Furukawa在2015年写了一个教程——[Multi-View Stereo: a Tutorial](#)是一本不错的入门书籍，可以在他的主页上下载。[Multiple View Geometry in Computer Vision Second Edition](#)是一本关于多视图几何的权威教材（作者之一是牛津大学，剑桥出版社出版），涉及很多基础的数学理论。如果没有看过这本书，读论文是会很头痛的。当然看这本书的过程也是很头痛的。业界超级超级大牛Richard Szeliski(曾经在微软研究院，现在在Facebook，写了著名的[Computer Vision: Algorithms and Applications](#))的主要兴趣就是MVS。因此MVS领域最重要的paper几乎都和Richard Szeliski有关，特别是2006的那篇综述[A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms](#)为2006年之后MVS的发展指明了方向。大家可以想想Richard Szeliski在Facebook会做些什么？游戏？VR？还是无人驾驶？顺便说一下，[Computer Vision: Algorithms and Applications](#)这本书的参考文献也是一本书（有141页之多），Richard Szeliski居然读过这么多论文。现在阁下知道为什么牛人称之为牛人了吧。



Yasutaka Furukawa和Richard Szeliski是合作关系。鄙人不太清楚这个关系是什么时候开始的。但估计2007年的Accurate, Dense, and Robust Multi-View Stereopsis应该是得到了Richard Szeliski的指导。总之，他们应该早就认识啦。



# 题外话

上面就是鄙人所做的文献综述，非常不严谨。一方面本文不是严格意义上的学术文献，另一方面就在于鄙人对现状的了解仍然有局限性（鄙人见过几位比较有名学者的照片，其他有名学者的照片没有见过，名字也记不住）。如果是哪位大牛（或者稍微有点水平的）看到，一定会感慨：写的什么玩意啊！所以如果各位要做研究，首先要了解现状。现状了解首先要**知道知名研究者的姓名**，其次是**知道联系方式**。如果研究过程中，阁下甚至都没有给这些牛人写过一封信，那是称不上在做研究的。很难想象真正的粉丝不给偶像写信，不知道偶像的生日。当然最重要的阁下需要知道牛人们都做了些什么，这样做的意义在于：

1. 避免“开黄腔”（重庆话，不懂的百度），特别是如果阁下的论文不小心被某知名学者审稿，而阁下恰好开了一个不小的黄腔。例如阁下说XX学者的XX方法存在XX问题，而恰好审稿的就是这个学者——玩笑就开大啦（前提阁下正好说错了）。
2. 避免重新发明车轮。有些学者喜欢把代码公开，如果阁下不知道，那是会重复很多无用的工作。
3. 开阔眼界，不要做井底之蛙。“我想做项目，不想搞理论”就是这种井底之蛙的外在表现吧。真正的好项目，哪有没有技术含量的？山寨都需要技术含量。很多学生想做数据库系统（算是应用最广泛的项目类型），会做数据库肯定不缺工作。但是不懂数据库原理，做出来的系统，真是不好评价。数据库系统一点不简单（鄙人做了很多数据库系统），如果阁下觉得简单，那是因为阁下做得很烂。不信？阁下想想支付宝或者银行的系统吧，想想可怜的12306吧。阁下以为懂个SQL语句就能够搞定高并发吗？双11和春运的并发问题，也算是世界水平的难度。另外还有安全性、健壮性、扩展性等等问题。
4. 发现问题，发现真正有价值的问题。发现问题是很多研究生的弱项，当然也包括鄙人。**我们的研究生总是擅长解决问题，而不擅长发现问题**。这恐怕和我们多年受到的训练有关，没办法中国教育都有这个毛病。千万不要以为所有的中国人都有这个毛病。哪怕只有0.001%的中国人没有这个毛病，也是很多人。这些人往往就是各行各业的精英。**只有发现问题，才有可能解决问题，创新总是从发现问题开始**。我举一个典型的例子，哥德巴赫猜想是一位中学教师哥德巴赫发现的，他将这个问题寄给了欧拉。哥德巴赫对这个问题的解决没有任何贡献，但这个问题却叫做“哥德巴赫猜想”。当然能够发现问题，又能够解决问题的人——可能就是牛顿或者爱因斯坦，要不然也是爱迪生。
5. 发现解决问题的思路。如果已经发现了问题，那么就要解决问题。解决问题需要思路，可以从广泛的文献中查找。爱因斯坦的相对论是在黎曼几何的基础上建立的。但据说爱因斯坦的数学很差（大学老是翘课），于是找到大学的数学老师，老师告诉他要用黎曼几何。黎曼几何不会怎么办？继续找同学帮忙。当然多年以后，爱因斯坦已经能够独立的使用黎曼几何。鄙人相信，国产手机抄苹果也是基于这种原理。

全面了解现状需要查阅大量的资料，论文阅读是最简单易行的方法。除此之外进行市场调研、阅读专利文档都是常见的方法。现在由于互联网的普及，阅读论文已经是中学生、小学生都可以做的事情啦。几年前美国中学生[Jack Andraka](#)就是靠着互联网上的论文资料，发明了检测胰腺癌的新方法（鄙人看到的资料是2年时间阅读了200多篇）。鄙人看到很多研究生都热衷于创业，可是创业之前，又花了多少时间在调研上面？阅读了多少文献资料呢？如果连文献综述都写不出来，那是很有问题的啊。几年前，某个医科大学的博士研究生联系鄙人，他想用一个医疗App创业。鄙人直接告诉他，他对App的了解不够，而做一个好医生也是“钱途无量”的。

题外话结束，最后隔空给鄙人的研究生喊一下话：**无论阁下今后想做什么，查阅文献（论文只算其中一种）的能力是必备的，特别是那个谁谁谁准备当产品经理的。**

行文至此，已经写了接近5000字符，鄙人保证全部原创，查重重复率应该接近于0。写一篇万多字的硕士大论文也不是很难的事情。

## 摘要

翻开论文首先看到的是论文摘要。Accurate, Dense, and Robust Multi-View Stereopsis一文的论文摘要，全文摘抄如下：

This paper proposes a novel algorithm for calibrated multi-view stereopsis that outputs a (quasi) dense set of rectangular patches covering the surfaces visible in the input images. This algorithm does not require any initialization in the form of a bounding volume, and it detects and discards automatically outliers and obstacles. It does not perform any smoothing across nearby features, yet is currently the top performer in terms of both coverage and accuracy for four of the six benchmark datasets presented in [20]. The keys to its performance are effective techniques for enforcing local photometric consistency and global visibility constraints. Stereopsis is implemented as a match, expand, and filter procedure, starting from a sparse set of matched keypoints, and repeatedly expanding these to nearby pixel correspondences before using visibility constraints to filter away false matches. A simple but effective method for turning the resulting patch model into a mesh appropriate for image-based modeling is also presented. The proposed approach is demonstrated on various datasets including objects with fine surface details, deep concavities, and thin structures, outdoor scenes observed from a restricted set of viewpoints, and “crowded” scenes where moving obstacles appear in different places in multiple images of a static structure of interest.

**摘要的阅读是文献阅读的重点，也是论文写作的重点。**好的摘要能够开门见山的告诉阁下，论文的目的、意义、研究内容、技术路线和实验结果。所以读完摘要，有经验的读者（或者

背景知识娴熟的)能够大体确认是否有必要详细阅读全文。没错,地球上有很多学术论文,有的论文堪称垃圾(包括鄙人的论文),这些论文是不值得阁下浪费时间的。如何确定哪些论文是精品,哪些论文会垃圾呢?简单的方法就是阅读摘要。当然要读懂摘要,必然需要相关的背景知识,因为摘要精炼,不会对术语和背景进行任何解释(正文中也不怎么解释,但要稍好一些)。那么这不是形成了一个死循环吗?要想阅读摘要,就必须有背景知识,要想获得背景知识,要靠阅读文献,而决定文献是否有价值阅读,又靠阅读摘要。要解决这个问题,建议优先阅读一些综述性的文献,例如[A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms](#)或者[Computer Vision: Algorithms and Applications](#))。总有一些文献是你必须花时间的, [Computer Vision: Algorithms and Applications](#))这样的文献就是如此(这本书的阅读难度相当大,能够顺畅的读完,已经是业界大牛)。

我们来一句句的分析这篇文献的摘要说了什么。

(1) This paper proposes a novel algorithm for calibrated multi-view stereopsis that outputs a (quasi) dense set of rectangular patches covering the surfaces visible in the input images.

本文提出了一个算法(novel algorithm, novel就不翻译吧,反正是作者自夸的)。这个算法是干吗的? for后面是一个长句,我不是英语老师,英语不过关就只有靠多读来解决了。简单来说,这个算法处理的对象是经过标定的多视图图片(calibrated multi-view stereopsis)。算法的输出是稠密的(dense)或者说准稠密的(quasi)矩形块(rectangular patches)。对于有三维重建背景的,都知道dense和rectangular patches是怎么回事。最后一个定语, covering the surfaces visible in the input images,是说输出的块覆盖了输入图像的可见表面,换句话说,看不到的部分不会重建(这不是废话吗?)。

(2) This algorithm does not require any initialization in the form of a bounding volume, and it detects and discards automatically outliers and obstacles.

这个算法不需要任何bounding volume形式的初始化,并且可以自动检测和丢弃离群值(outliers)和障碍物(obstacles)。bounding volume也是一个常规术语,三维重建领域有很多方法都是基于bounding volume的。outlier翻译为外点、离群点、离群值或者异常值,异常值就懂了。因为是在现实环境中进行重建,异常和障碍都是需要处理的。

(3) It does not perform any smoothing across nearby features, yet is currently the top performer in terms of both coverage and accuracy for four of the six benchmark datasets presented in [20].

算法没有根据周围的特征进行任何平滑,并且这个算法是目前精度最好的:对文献[20]提供


的6个测试数据中的4个进行重建，在覆盖性和准确性上都是目前最好的。文献[20]就是A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms，coverage和accuracy都是文献[20]提出的测评指标。如果阁下看过文献[20]，那么一定知道本文的结果会显示在<http://vision.middlebury.edu/mview/eval/>。如果是2007年读到此文，一定会到这里去看看，现在去也不晚。并且读完这句，基本可以确定文献是有很高的阅读价值的。为什么？就凭 top performer就足以说明啦。

Depth Fusion		0.53	99.5	0.72	96.8		0.46	99.5	0.42	97.8
ECCV2014_1338				0.62	95.2				0.37	99.2
ECCV2016_104	0.41	99.6	0.49	99.5	0.57	98.1	0.44	99.8	0.46	99.7
ECCV2016_624	0.37	98.9	0.49	97.6	1.27	39.2	0.26	97.8	0.31	99.5
Fuhrmann-SG14	0.39	99.4							0.28	98.1
Furukawa	0.65	98.7	0.58	98.5	0.82	94.3	0.52	99.2	0.42	98.8
Furukawa 2	0.54	99.3	0.55	99.1	0.62	99.2	0.32	99.9	0.33	99.6
Furukawa 3	0.49	99.6	0.47	99.6	0.63	99.3	0.33	99.8	0.28	99.8
Galliani	0.39	99.2	0.48	99.1	0.53	97.0	0.31	99.9	0.3	99.4
Gargallo			0.88	84.3	1.05	81.9			0.6	92.9
Generalized-SSD	0.53	99.4	0.81	95.8			0.55	98.1	0.6	96.0
Geodesic GC					0.88	94.7			0.71	96.8
Goesele	0.42	98.0	0.61	86.2	0.87	56.6	0.56	80.0	0.46	57.8
Goesele 2007	0.42	98.2					0.46	96.7		
Guillemaut	0.43	99.0	0.71	97.6	0.86	96.2	0.35	100	0.58	99.5
Habbecke	0.66	98.0					0.43	99.7		
Hernandez	0.36	99.7	0.52	99.5	0.75	95.3	0.49	99.6	0.45	97.9
Hongxing	0.83	95.7	0.79	96.3	0.97	93.9	0.62	96.3	0.5	99.1
Hornung	0.58	98.7					0.79	95.1		
ICCV2015_1020	0.45	99.2	0.56	99.2			0.35	99.5	1.05	99.2
ICCV2015_293	0.52	99.2					0.62	99.0		
IRAY					1.2	91.9				0.63
Jancosek-3DIM99	0.65	85.8	0.7	78.9	0.59	74.9	0.91	73.8	0.71	76.6
									0.66	74.9

Reference: Y. Furukawa and J. Ponce. Accurate, dense, and robust multi-view stereopsis. CVPR 2007.

Normalized Time (H:M:S): 3:31:01

Furukawa 2 and Ground Truth



(4)The keys to its performance are effective techniques for enforcing local photometric consistency and global visibility constraints.

算法有如此性能的关键在于：用有效的技术加强了局部（local）photometric consistency和全局的可见性约束（visibility constraints）。photometric consistency一直没有找到合适的翻译，姑且就不翻译。photometric consistency和visibility constraints两个术语，显然也不是本文作者的发明，而是有些年头了。作者只是采用了effective techniques，至于effective techniques 是什么，摘要看不出来。

(5)Stereopsis is implemented as a match, expand, and filter procedure, starting from a sparse set of matched keypoints, and repeatedly expanding these to nearby pixel correspondences before using visibility constraints to filter away false matches.

Stereopsis找不到合适的翻译（尽管可以翻译为立体影像，但是并不贴切），姑且就不翻译。这句在讲重建的过程。重建过程是匹配（match）、扩张（expand）和过滤（filter）的过程。开始于匹配特征点的稀疏集合（matched keypoints），然后从这些集合扩张至相邻的像素对（pixel correspondences），该过程会重复多次，最后用可见性约束过滤掉错误的匹配。这一句出现了很多术语，例如matched keypoints、sparse set、pixel correspondences、visibility constraints、false matches。即便可以对这些术语进行翻译，如果缺乏背景知识，也不能很好的理解。例如keypoint（特征点）就是一个比较复杂的概



念，[Computer Vision: Algorithms and Applications](#)讲了一个章节。sparse（稀疏）又涉及到MVS算法的前序过程Structure from Motion，[Multiple View Geometry in Computer Vision Second Edition](#)用了一本书。论文阅读中遇到的术语，不是简单翻译可以解决的。就好比阅读物理论文的时候，看到“引力波”这个术语，你以为不学习相对论能搞懂这个概念？好吧，一旦要学相对论，还得先把初等物理全部学完。大多数研究生无法理解论文，核心原因之一就是无法理解术语。这里给大家推荐[维基百科](#)。大多数术语可以在这里查到。理解术语的过程，也是一个漫长的学习过程。

(6) A simple but effective method for turning the resulting patch model into a mesh appropriate for image-based modeling is also presented.

提出了一种简单有效的方法实现块模型（patch model）转变为网格面（mesh），这种网格面适合（mesh appropriate for）于图像的建模（image-based modeling）。mesh和image-based modeling也是三维重建方面的术语。

(7) The proposed approach is demonstrated on various datasets including objects with fine surface details, deep concavities, and thin structures, outdoor scenes observed from a restricted set of viewpoints, and “crowded” scenes where moving obstacles appear in different places in multiple images of a static structure of interest.

这句话很长，翻译起来很累，但是意思很简单：本文提出的方法经过测试，效果很棒！实验的测试环境包括：（1）objects with fine surface details；（2）objects with deep concavities；（3）objects with thin structures；（4）outdoor scenes observed from a restricted set of viewpoints；（5）“crowded” scenes where moving obstacles appear in different places in multiple images of a static structure of interest。

实验是一篇论文很重要的环节，没有好实验，没有好论文。研究生不会做实验，却想做项目，这不是在做梦吗？这篇优秀论文中可以看出，作者为了证明方法的有效性，进行了大量实验。换句话说，写了很多代码。“**写论文是不需要写代码的**”，这也是很多研究生的误解。现在很多学生喜欢读学硕，问他们为什么，不少人就回答——**不喜欢写代码**。殊不知，好的学术研究都要写好多代码。不少学生到中科院或者微软研究院去实习，也就是去帮别人敲敲代码。不会写代码，谁敢让你做科研？真以为计算机是搞数学吗？计算机只是用数学。

当然现实中，做科研涉及到：查阅资料、分析问题、提出问题、解决问题、实验、论文发表。如果不是违反学术道德，实验和论文发表请人来做也未尝不可，俗称外包。但是外包通常是需要钱的，如果解决的问题不是特别有价值，这个钱恐怕花得也不值。无论如何查阅资料、分析问题、提出问题、解决问题最好是自己来做。考虑到今后的就业，实验也尽量不要外包。至于论文写作外包，真心不好找人。

# 引言和相关工作

学术论文是典型的八股文，格式固定，内容固定，连标题都是差不多的。引言

(Introduction) 往往是摘要的扩大版，会比摘要有更多的细节。如果读完摘要，对论文的价值尚有疑问，可以继续读读引言。一般引言读完，是否继续读下去，自己心里就有谱了。引言的内容主要包括：（1）研究背景（2）研究目的（3）研究方法概述（4）实验效果概述（5）本文方法与其他方法的异同。当然做一些文学性的抒情也是可以的，但如果考虑到论文版面费，还是节俭一点为好。

一般引言之后就是相关工作（Related Work）。相关工作就是对论文相关的研究进行概述。Accurate, Dense, and Robust Multi-View Stereopsis这篇文章在2006年的版本不是标准的写法（实际上还有一个2010年的版本），相关工作是混杂在引言中来写的。无论放在什么地方写，它的基本内容不会变。要写好相关工作，就必须对相关工作有充分的了解。所谓充分的了解，仅仅局限于“知道”是不够的，至少需要达到“理解”这个层次，否则就很难用自己的语言表述。

“相关工作”体现的不是作者写作的功底，而是作者知识的积累。没有大量文献的阅读，没有对技术深刻的认识和理解，相关工作就只能抄。反过来说，认真阅读“相关工作”是非常有意义的，特别是才开始学习的研究生。一方面“相关工作”是前人知识的总结，指明了文献查阅的方向。另一方面阅读“相关工作”可以检验自己对知识的理解，如果有“英雄所见略同”的感受，那多半是已经入门了。

Accurate, Dense, and Robust Multi-View Stereopsis引言部分就留给大家阅读。这里只摘录一段，关于论文方法概述的：

It implements multi-view stereopsis as a simple match, expand, and filter procedure (Fig. 1): (1) Matching: features found by Harris and Difference-of-Gaussians operators are matched across multiple pictures, yielding a sparse set of patches associated with salient image regions. Given these initial matches, the following two steps are repeated  $n$  times ( $n = 3$  in all our experiments): (2) Expansion: a technique similar to [16, 2, 11, 13] is used to spread the initial matches to nearby pixels and obtain a dense set of patches. (3) Filtering: visibility constraints are used to eliminate incorrect matches lying either in front or behind the observed surface.

这段的内容和摘要有些相似，但是提供了更多细节：

（1）Matching：先用Difference-of-Gaussians对图像进行处理——也就是提取了边界明显的地方。再采用Harris进行特征（feature）提取，Harris是一种角点（corner）提取算法。特征点提取之后进行匹配，形成稀疏的匹配集合。论文写作的年代，Harris算法还不那么普及，所以后文用一个脚注来介绍这部分的实现细节。今天可以用OpenCV直接做这个事情，

并且SIFT通常是比Harris更好的选择（后文鄙人会使用SIFT而不是原文中的Harris）。

（2）Matching之后进行 Expansion和 Filtering，共重复n次（文章为3次）。Expansion采用的方法与文献 [16, 2, 11, 13]类似，所以如果已经读过 [16, 2, 11, 13]，本文的阅读难度就会下降很多。Filtering是基于visibility constraints的，目的是为了移除错误的匹配。

引言和相关工作之后，就开始了论文的主体部分。下一个文档再讲。

## 文献检索方法

总结一下文献检索的方法：

1. 阅读一些综述性的文献，获取基本的背景知识。
2. 根据关键词（通常为术语）进行检索，得到一些候选文献。
3. 阅读文献摘要，确定文献价值，价值高的优先读。
4. 记下论文作者的姓名、工作单位、联系方式，检索作者的个人主页，优先阅读作者的相关研究。
5. 根据引言和相关工作章节列出的文献进一步检索。
6. 根据参考文献进行检索。
7. 仍然按照摘要确定文献阅读的优先持续。
8. 需要一款文献管理软件，Mac电脑推荐**bibdesk**，windows电脑推荐excel（不喜欢endnote，因为付费）。