

# CVPR 之感想

田渊栋  
[我的主页](#)

## 章零 引子

结束了，终于结束了，这该死的 CVPR。

不得不承认，写 paper 是一件相当麻烦的事情，没有 deadline 的催促，一般来说是搞不出东西来的，虽然这很不人道。

这篇 paper 严格来说是在一个月之内弄出来的，之前的种种初稿和片断，乃至整体算法，在最后的文章里几乎全部被删去了；所有的东西全部从头来过，以一种更为漂亮优美的手段，重新做。

这种让这篇本来脱胎于项目的，七拼八凑得到的技术报告，变成有希望接收的，可以提交的论文的手段，便是数学。

## 章一 黑夜中，我见到了黎明

整整一个月前，11月6日，周一。

那是一个悠闲的午后，离 deadline 还有一个月，而这篇以我看来不能称其为 paper 的 paper，在老板的催促下，已经改了相当多的遍数了。

说实话，当时我是强忍着一遍一遍地改的，因为整篇文章到处都是直观的述说，及不能整合在一起的目的地各不相同的算法。这些东西直接来自于先前赶工的名为照片聚类系统的项目，没办法，虽然是 MSRA，但是很多项目还是要做的。

我对这篇 paper 并不抱有什么希望，心里也清楚即使是投出去，也一定是白投的；另一方面，对于 Object Recognition 则抱有相当的兴趣，所以从每天的日常工作中抽出了很多时间来进行这方面的初步研究，毕竟 CVPR 接下来就是 ICCV，再不抓住机会的话，自己的硕士兼研究生涯就要悲惨地结束了。

我这样想着，老板来找我了。

老板是个很好的人，编程方面尤为强悍，我们这个组里的很多库都是他写的；不过也许正因为此，对待研究的态度更偏向工程。在我看来这篇纯粹堆砌的 paper，在他看来似乎还行。不过今天，他改变了态度，面色凝重地找我谈话。

因为他把 paper 给大老板看过了。

“Formulation，我们需要 Formulation。”他说道。

我惨然地笑了笑，点点头，“我知道。”

“为什么之前不说？”

“我提过的想法，可是都被你砍掉。所以觉得就这样了，没意思。你为什么不提呢？为什么总是让我改些细枝末节呢？”

老板无言，只是说了句，“快去改吧，时间不多了。下周一要给大老板做报告。”

从那一天，距离 deadline 还有一个月的某一天，我终于开始了在 MSRA 的研究生涯。

大老板，我赞美你。

11 月 10 日，周五。

Formulation 初步完成了。这一周推推公式，写写代码，让我终于有了研究的感觉：自己想方法，自己实现，自己做实验。

实验结果不错，和自己的预期相当符合。我松了口气，将组织起来的一整套模型写完，替换掉原来的垃圾算法，这篇论文开始有可以一投的气象了。

不过时间依然很紧，为了在周一那一天能有拿得出手的数据供大老板观赏，周末的实验相当地匆忙，bug 众多，接连两次跑出的结果都完全作废；还是在周一的早晨，才将统计数据弄完，下午就出阵了。

## 章二 我仍奋斗，你却远走

11 月 15 日，周三。

十一月过半了，北京的气温进一步下降。北方的风刮在手上，皮肤马上就变得粗糙不堪；饭菜又贵又难吃，乌糟糟的色香味俱无。

每天来回吃饭要经过的小路，仍然是那样破旧；卡车轰隆开过，烟尘四起，行人掩鼻。

我的同事要走了，强装的笑颜掩饰不住双眼偶尔流露出的遗憾。同样需要 Formulation，他没有成功，我帮了他三天，却还是没能用清晰而明确的模型解释清楚他文章中究竟是要做什么。老板摇头，不肯再在他的论文上再花费精力，然而他还在解释，还在和老板争辩着可能的出路。

可是还是没有成功啊。后来我才意识到，他想要做的两个目标，完全是相互之间背道而驰的；他会说理，但是说来说去难以说清，就像  $C$  是  $A$  与  $B$  的和， $A$  变大  $B$  变小但天知道  $C$  是变大还是变小一样。

无论如何，定性的描述，只能得到这样的结果。文科之所以有百家争鸣，只是因为缺少定量的分析，没有人知道每个因素占多大比重，结果究竟是什么；可是这里不行，绝对不行。

他和我七月份同时来，他却因为学校老板的召唤要先走。在这几个月里，他几乎凌晨五六点回到寝室，然后早上十点又出现在研究院里，紧张地询问老板是否来过；他看过一箱子的文章，却始终得不到自己的见解；他基础不好，但却诚惶诚恐听从老板的指导，在直观与定性说理的路上越走越远。

可惜他几个月的辛苦，只是换来老板一次次冷酷的摇头而已。

这里不是培养人的地方啊，我叹了口气。一切，都只能靠自己努力。

17 日，他回去了。而我只能目送他落寞的背影，消失在茫茫的人群之中。

记得了，救了自己的东西，是数学。

11 月 16 日，周四。

下半部分有关用户交互的建模，从这一天开始了。

从周四到周五，短短的两天，我提出了三个可行的方案，并编程实现了其中的两个。虽然说和以前的用户交互一样也是从直观出发的，不过毕竟我有了用户交互的模型，对于何时记为一次交互有了明确的定义。

大老板提出的要求可以算是部分实现了，然而从直观出发的方案天生不是完美的，总有一天，需要数学的介入，把它 *formulate* 成一个能自圆其说的结构。

不过这些都是后来的总结了，当时出现了令人惊异莫名的事情，即三个方案中的第二个，其实验结果出奇得好，好到我不敢相信自己的眼睛。

兴奋之际，我将结果报告给我的老板，老板在一番思考之后，得出了这个方案是有道理的结论，以结束了这一天的工作。

后来的检查证明，我程序编错了。

研究有时候，只是对于一堆事实的可笑解释罢了。当事实不存在的时候，解释便换作另一条，仍然是如此堂堂正正。而人类的所谓直觉与理性，还是要一个逻辑上完美的框架与相对正确的事实扶起来，才能一步步地往前走。

框架是数学，而事实，则是实验。

### 章三 运气，我赞美你

11月20日，周一。

在相当尴尬地将程序编错这个事实报告给老板之后，老板却也没有责怪的意思，只是嘱咐我要尽早写 *paper*。

“只有两个星期了。”他说道，“先把文章弄好，实验部分可以慢慢弄。”

我知道这也是没有办法的办法，若是实验不好，那么吹牛的负担就要重得多；不过不管怎么样，吹牛总是工程性的，风险较小些。

我于是停止程序上的工作，埋头写了三四天的文章，将目前为止的进展都放进去，主要部分几乎重写。

不过后来证明，这一次的修改，仍然有一半是完全无用的。

11月22日，周三。

CVPR 注册的最后一天。

我仍然不知道注册时间会如此快地到来，仍在悠哉悠哉地改 *paper*。老板和我说是 27 号，我也就点点头表示知道了。

如果不是因为 [Filestorm](#)，如果不是因为周二他在 AI 板上秀了一份他的 *demo* 而我对其大加评论，如果不是因为周三他的名字改得奇奇怪怪以致于我很好奇地问他这是什么，我大概会成为写完 *paper* 却没能注册的不幸的人中的一员吧。想来真是后怕到胆战心惊。这份人情无以为报，看来只有用明年的 ICCV 来抵偿了。

ID 号是 2164，这次命中的概率，实在不容乐观。

11月24日，周五。

还有一周的时间了。我将 *paper* 改完了，交给老板修改。

此时此刻，我仍然对后半部分的建模不够满意。后半部分的建模思想，主要想在那些难的部分中，找一些最难的极少数，让用户标定；然后再次做聚类，以期望得到好的聚类结果，这样用户标起来就会方便些。

这个东西有两个问题：第一，如何定义“最难”？第二，标最难标几次才做聚类？

到这一天为止，“最难”的定义仍然基于直观上的见解，而没有全局上的数学表达。我首先去做的，就是试图定义什么是“最难”。

该挑哪一张给用户标呢？

这又是一个 Formulation 的过程，上次花了一周，这次若再花一周，那么这篇 paper 就废了。

幸运的是，这次只花了两天。

## 章四 站在数理的肩膀上

11月27日，周一。

当我看着电梯里满满的匆忙的人，带着憔悴的面容走进研究院的时候，我不得不感谢伟大的统计物理。

这个想法，完完全全，就像是做物理的人搞出来的。

在两天的挣扎之后，我终于定义了什么是“最难”。并立即把这些东西写进 paper 里去，而原来写的那些，自然就再一次作废。

不过我的心仍然忐忑不安，因为这个算法的复杂度偏高，如果程序编完了，结果跑不出来怎么办？我的天，只有一个星期了。

幸运的是，实验的结果不错，速度也可以忍受，用户交互次数下降了。更为让人觉得不错的是，将现在的这个系统和网上的另一个知名系统比较，现在的这个效果相当好。

这些东西，让我头一次看到了被录取的希望。

我深呼了一口气，开始到处征用机器以进行实验。在我的心里，Formulation 已经算是大功告成，虽然第二个问题仍然悬而未决，剩下的这个挑多少次最难然后聚类这个参数明摆在那里给审稿人当活靶子，但我已决定放弃了。

11月30日，周四。

一个值得纪念的日子。

在一整个白天的改 paper 及调参数之后，晚上，老板来找我了。

“这里会被审稿人说的。”他指着 paper 的一处说道。我看了看，正是我心知肚明的那个弱点。

“是啊，确实。”我承认，“不过现在暂时想不出更好的办法来 Formulate 它，只能这样放着。我们的时间不太多了。”

老板仍然摇头，他显然觉得这里是不可忍的。我一直觉得奇怪，为什么一个月前他面对这样一篇不入流的 paper 还能忍，现在对这样一个相比之下好得多的地方却不能容下。

当然，对我而言，前者是无需再忍，干脆眼不见为净算了；而后者仍有挽回的余地，只是时间不够了而已。

不过他的话，激起了我还想试一试的冲动。

一番讨论下来，他提出的东西相比现在使用的方法，仍然没有很大的差别，参数没有消除，而是变多了。我看着直摇头，说出坚决的反对意见：“数学上的美被破坏得一干二净了。”

“那你说应该如何做？”他问我。

讨论果然是好东西，我突然想到了另一种方法，可以把选最难的，把聚类这些都放到一个框架里面去。我简略地将自己的想法提出，他却说：“你就不要想了，就用我的吧，时间来不及了。”

我摇头：“我刚才想到的办法很好的，应该要试一试。给我一个晚上。”

他显然有些生气，对我这种在离 deadline 只有几天的时候还要大费周章地去想的作法感到不可理喻，“我觉得没必要再花时间去做这个，现在没几天了。照我的方法去做吧。”

“可是那还不如不改……”我争辩道，“你那样参数怎么调？我这个办法是没有参数的。”

他终于不再说什么了，也许他想到了，我的 paper，终究是我的吧。

为什么在那天晚上我会如此固执，我也不太清楚；若是弄不出来，又会怎么样，我也没有想过。

不管怎么样，在第二天早晨，以没有睡好为代价，我弄出来了。

伟大的概率论，伟大的统计物理。

每想到这，我不禁热泪盈眶。

12月1日 周五

一早在与老板讨论完了之后，我终于以自己的模型把他说服，并且在中午之前，写完了代码。

效果更好了！现在程序能动态地选出一组大小不定的脸，让用户来标定；并且奇妙的是，几乎选出来的都是同一人的脸，这样用户对每一组，只要标定一次了。

对于这个理论上和实验上都很漂亮的结果，我感到无比的兴奋；只是 paper 还有实验，不可避免地又要重写重做了。

Deadline 是西雅图时间的 12 月 3 日 23 时 59 分，对应北京时间 12 月 4 日下午 4 点。

感谢这个旋转不息的地球，感谢我们的祖先把居住地选在了东方，它让我们多了 16 个小时。

## 章五 决战，惊心动魄

12月2日 周六

阳光明媚的早晨，回暖的天气。

我脚步急匆地踏进研究院的大门。

感谢我的同事，他在走的时候嘱咐过要让我把他用的电脑还回去，可是我忘记了。于是就没有人收，就那样闲置在那里，3.0G 的 CPU，1G 的内存，没有用来玩游戏，没有用来上网聊

天，用来救我的命了。

我的实验啊。三台机器不分昼夜地跑着，结果存成一个个的文本文件。我小心地把它们收集起来，时刻备份，不敢有丝毫大意。

丢了一个，就没有翻盘的机会了。

我这样边做实验，边将自己新写进文章里的 Formulation 不停地修改。文章里的其它部分，如引论及结论，则交给我老板修改，毕竟这些对审稿者的印象有至关重要影响的部分，还是要让有经验的人反复地去打磨的。

如此一直工作到晚上，我已哈欠连天，成强弩之末了。休息的欲望占据了身心，但是想来想去，还是看一眼实验结果为好。Matlab 就开着，我将读数据并画图的指令写在小程序里，以便以后可以运行一次小程序就出结果，改起图来也比较方便。

当时不觉得这是个好习惯，但它在最后一天的搏杀中，起到了至关重要的作用。

晚上 9 点多，图跳了出来。我看着屏幕上的曲线，冷汗直冒，发现了问题。

天啊，这个……有一个数据集的结果是反的！

我几乎要绝望了。为了实验更有说服力，老板又找来两个新的数据集看看效果，但是在这两个数据集上我并没有好好地调过参数。虽然说我的第二部分的用户交互是不需要参数的，但是第一部分作为第二部分的初始化，不得不有参数。

太可怕了，太可怕了……这些该死的至关重要的东西啊。我发誓在以后的模型里，再也不能有可调的参数了。

但是再发誓也没用，时间如冰冷的剑，正一寸一寸逼近。研究院里静得可怕，只有利落的键盘声，还有我的心跳。

现在再没有时间调它们了，只能凭感觉先选定一个，明天看结果！

我大致估计了一下，选了一个，配置完，走人。

天黑，风寒，心摇。我的面色一定白得可怕，只是在这个清冷的周末，没有人看见而已。

12 月 3 日 周日

我一早过来，第一件事就是查一下昨天的结果。

幸好这个反常的数据集并不大，一个晚上就跑完了。Matlab 的窗口仍然开着，我把画图的小程序打开，迟疑了一下，终于还是让它运行了。

图跳出来，搞定了！

我瘫坐在椅子上，默默地感谢上苍。

与此同时，老板那边也把东西改得差不多了，他听到我的消息，也是松了口气。

希望就在眼前，忽明忽暗地闪动着。它已经不再是幻影了。

我开始写实验部分，一直到晚上，两人把文章的版本统一了一下，今晚就能给大老板看了。

月亮是圆的，月光还是很皎洁的。

我大概这时在傻笑。

晚上 10 点我准备回寝室休息，老板说他要熬夜改，我带些歉意地笑了笑。

谢谢了。

一切都似乎尘埃落定，风平浪静了。

但是……我睡不着。

12 月 4 日 周一

还有 8 个小时。

我看着手机，上面的数字准确地显示上午 8 点。

“竟然会这样，我……”我抓着头抱怨着。凌晨零点半的时候，我老板发来消息，说我的实验有大问题，而那时我手机的电池板已经取了下来，正在充电！

老天啊，开什么国际玩笑！

我冲上楼，打开电子邮件，里面只有一封信，是早上 6 点 45 分寄来的。

还好还好。看完信，我喘了口气。他只是让我删掉几个实验，多做另几个实验而已；而要多做的那些实验，数据都是现成的，所有的工作量只是用 Matlab 编写程序。

那还不简单么？！

我三啃两啃地将早饭消灭，开始工作。但是由于积累了一周极高的工作强度，昨晚又没有睡好，头晕目眩，效率下降至极低点：编了两个小时，才勉强将一个 50 行的小程序写完。

快了，快了，快到头了。最后的搏斗啊！

小程序 bug 众多，好不容易修好它们，将图表画出，却发现结果不很理想。我正抓头之际，老板出现在我的面前，我和他说这些图表是如何画出来的，他皱着眉头连连摇头。

“不是不是，我的意思不是这样……你把你的第一阶段都固定，然后比第二阶段！”

“和谁比？”

“和简单的聚类方法比。”

“你是说在第一阶段我们的算法跑完之后，把分离开的难的东西再聚类？再逐个逐个标？”

“不错。”

“可是……”我站了起来，“可是这些没做过实验啊，没有数据！”

“现在就做吧。快一点。”老板说。

我呆住了，瞥了一眼屏幕右下角的时钟。

十一点了。还有 5 个小时。

“好的。”我说了一句，“我尽力。”

老板点点头，离开了。

我看着屏幕。这 5 个小时，要编程，要调 bug，要跑出结果，要画图，要写分析；但现在连程序也没有！

不行了，只能硬着头皮上了。

打开 VS2005，快速浏览编过的片断，想着抓取可用的部分。我东拼西凑，终于在一点半时将程序写完调试完，开始跑了。

其间老板竟送来楼下的日餐，令人好生感动啊。

程序算跑得很快了，非常快了，但是还是要等一个多小时，还是要在近三点的时候，才能出来结果。这该死的 EM 迭代啊，要运行两百次以消除无聊的随机初始值的影响，每次还有  $O(n^3)$  的特征值分解……

求求你了，快一点……

快三点的时候，程序终于跑完了。我把数据收集起来，小心翼翼地打开 Matlab，开始画图。我的手在发抖，几乎要祈祷了。虽然说我在正式运行之前检查过一遍，但程序的 bug 总是不可避免的；如果是那样，就彻底玩完了……

Matlab 闪了几下，图没有出来。

我手忙脚乱地查看 Matlab 的画图程序，终于发现一个隐藏很深的问题，把它修好。

三点十分，图终于再次出现了，但是横轴有误.....

怒了，怎么会在这种时候频频犯低级错误！我再次修改，图又出现了，效果很好！

是三点二十五分。

**It works!**

程序没错，结果是对的，是啊，没错的，我居然搞出来了.....我想说一千个赞字，但是终究还是没有说。把图打包，发给老板，并在每幅图下面配以标题和分析，全部弄完，上传，正是三点四十五。

我看着惨白的天花板，手里捏着 paper 的打印稿。

啊，几乎要吐血了。

## 章六 余音

好坏是出来了。

这是我所能想象的最好的结果，之前有任何一步走错，都会来不及的。

如果老板晚一些将 paper 给大老板过目，如果大老板没有提到关键点上，如果 Filestorm 没有告诉我 Deadline 的日期，如果老板没有斤斤计较文章中的问题，如果我没有在一个晚上把东西弄出来，如果我把我同事的机器给退还了，如果最后一个周末实验时及画图时的程序拙劣些，如果.....

呵呵，我只能安慰自己说，已经没有这些如果了；不过，经历了这一切，我大概真的会相信命运的存在。

[tydsh](#) 于大战之后的黎明