

# 十年磨砺出钝剑，再十年干将莫邪

山世光 博士，研究员  
中科院计算所智能信息处理重点实验室 副主任

关键词：兴趣，磨砺，积累

《中国计算机学会通讯》专栏编委於志文教授邀请我写一篇短文介绍自己的科研成长经历并分享一些“在路上”的经验和体会，我当时欣然接受，但真正落笔之时却发现自己的经历实在乏善可陈，一些所谓的经验和体会也是支离破碎难成体系，很怕起不到正向的作用反而误人子弟。但最终还是硬着头皮写下来，希望自己的简单经历和点滴体会可以对大家有所启发，如有不当之处还望读者海涵。

## 十年磨砺出钝剑

时光荏苒，从 2002 年留计算所工作至今转眼已过去十年。回顾这十年，我在高文院士、陈熙霖研究员的关心和支持下，带领计算所人脸识别研究组在计算机视觉与模式识别领域做了一些工作，特别是在人脸识别问题上，可谓“十年磨一剑”。我们完成的人脸识别技术于 2004 年和 2006 年分别获得了国内和国际人脸识别竞赛的第一名；2005 年，我们完成的人脸识别相关成果获得国家科技进步二等奖(我是第 3 完成人)；2008 年我与博士生王瑞平合作完成的有关流形到流形距离的论文获得了国际会议 CVPR2008(计算机学会推荐的领域 A 类会议)的 Best Student Poster Award Runner-up 奖。我个人也获得了一些荣誉，如：中科院 2008 年度卢嘉锡青年人才奖；2009 年度北京市科技新星；2009 年“Elsevier Scopus 寻找未来科学之星”银奖(信息领域共 2 名)。此外，今年我还获得了国家基金委“优秀青年科学基金”的资助。

与上述外在的荣誉相比，我更看重的其实还是我们在相关领域最好的国际学术会议和最权威的学术期刊上发表的文章，以及这些文章的影响力。2005 年以来，我们累计在 ICCV/CVPR(计算机学会推荐的领域 A 类国际会议)上发表论文 17 篇，在 IEEE 汇刊上发表论文 16 篇(其中计算机学会推荐的领域 A 类国际刊物 T PAMI 论文 4 篇)。我作为共同作者所发表的论文已被 Google Scholar 引用 3100 余次，SCI 他引累计超过 450 次。这说明我们的工作得到了国内外同行的关注，产生了一些学术影响力。也许正因为如此，去年以来我连续应邀担任了国际计算机视觉大会 ICCV2011(计算机学会推荐的领域 A 类会议)、国际模式识别大会 ICPR2012、亚洲计算机视觉大会 ACCV2012 以及人脸与体势识别大会 FG2013 的领域主席 (Area Chair)，承担了为这些重要国际会议处理数十篇人

脸识别领域投稿的职责。

我同样自豪的是我们团队所研发的人脸识别相关技术在产业界的影响力。我们的技术经过银晨科技等合作企业的产业化，已在公安、金融、网络等领域得到了成功应用，如上海世博会、多省市公安局、海关以及百度寻人等。例如，在XX省出入境管理局，试用期间即已成功比对出百余名多重身份的违法人员。近年来，在该领域与我们合作的企业越来越多，其中既包括百度等多家知名的国内企业，还包括 Omron、Qualcomm、NEC 等多家国际企业。

## 兴趣是最好的老师

要问是什么力量让我在这个略显“古老”的人脸识别领域一钻就是十几年，我觉得最重要的还是兴趣，或者说是对面脸识别以及推而广之的一般视觉问题的好奇心。这要从我最初进入研究领域的经历谈起。在哈尔滨工业大学计算机学院读书期间，为了增加实践经验，我在大三时进入了赵铁军教授负责的机器翻译实验室实习，做了点零零碎碎的小活儿。大四上学期，在获得本硕连读资格后选择研究生导师的关键时刻，赵老师睿智地为我指明了方向，赵老师说：“你应该根据自己的兴趣来选择导师，比如计算机应用领域的典型方向包括声、图、文的智能处理，你对其中哪个感兴趣呢？”。这个问题令我茅塞顿开，我毫不迟疑的选择了“图”，因为我觉得“图”直观，看得见摸得着，好玩有趣。于是赵老师推荐我加入高文教授在哈工大的视觉处理实验室(Vilab)，之后我联系高老师并顺利的加入了该实验室攻读硕士，师从刘岩副教授从事图像识别问题的研究。1997年本科毕业设计时，刘老师给我安排的题目就是人脸识别，我从此与人脸识别结缘。

从最开始被一个简单的“图”字激发起兴趣，到深入了解后对视觉模式分析问题挑战性的深刻认识，我的内心萌发了越来越强烈的探索计算机视觉未知领域的主观愿望，从而引导着我不悔地探索着这个领域。特别是进入人脸识别研究多年以后，我越来越清晰的认识到这一看似“古老”的课题其实还存在很多悬而未决的开放问题。将其推而广之到一般视觉问题，并联想到人类视觉信息处理系统之精巧和神秘，好奇心愈加驱使着我持续探索以尝试突破这些挑战性问题。在我看来，只要问题在那里，无论有多难，就是等着我们去解决的；只要有未知的神秘，就是等待着我们去探索和发现的！

因为这段经历，在我自己做了研究生导师后，每个报考我的研究生都会被我问到这样一个问题：“扪心自问，你到底对什么感兴趣？”。如果那个学生不知道该怎么回答，我也会给他介绍相关领域情况，并希望每一个学生都能够像自己当年一样追随自己的内心做出不悔的选择。我曾在一次给研究生做报告时说：“在选导师或选题的时候，要拿出一种找终身伴侣的精神来”。在我看来这并不是玩笑，而是非常中肯的建议，因为只有当我们对某件事有兴趣时，我们才会愿意去付出，才会愿意舍弃其他，投入激情去做这件事。更何况，真正值得研究的问题

往往都不是那么简单，在这时候，尤其需要浓厚的兴趣和好奇心帮助我们抵抗其他的诱惑，并驱使我们去不断探索，而这其实也正是做研究的乐趣之所在。当我们突然想清楚了某个难题的时候，会感受到一种普通人难以理解的愉悦感，而这正是研究的无穷魅力和乐趣之所在吧。

## 要么不做，要么做到最好！

1999年，我考入计算所读博士，师从高文教授，从那时起到现在，高老师不仅指导了我的论文研究工作，更为我提供了自由宽松的学术氛围和令我受益终身的科研理念。在高老师诸多重要的科研理念中，我印象最深的是高老师曾经教导过的一句话：“要么不做，要么做到最好”。简单的十个字，实际上包含丰富的内涵，也奠定了我自己以及所在团队秉持的一些重要科研理念。

首先，做研究的目标必须要高远，即：要调研清楚自己所在研究领域的国际最高水平(state of the art)并不断超越它。从某种意义上讲，做研究就像参加奥运会，目标就是超越前人的“最好”。其次，做研究要学会坚持！一旦选定一个题目，就要长期坚持下去，切忌频繁更换题目和研究方向，更不能因为一时的挫折而放弃，或者因为别的题目看上去“更热门”或者“更容易出成果”而放弃自己的选题。实际上，奥运会的每个项目夺冠的难度都是差不多的，科研也一样，每个方向做到前沿都需要付出同样大量的努力才行。最后，我特别强调“积累”的重要性。除了幸运地遇上少数的新生方向，可以说大多数研究领域都已经被国内外同行掘地三尺反复研究过数年乃至数十年了，凭什么你一进来就能做到高人一筹？以我从事的人脸识别为例，19世纪60年代国际上就已经开始相关研究，70年代初期国际上就有了这方面的博士论文，90年代初期国际上就曾经掀起过研究热潮，MIT、CMU等几乎每个国际名校都有赫赫有名的教授带领团队从事过相关研究，累计发表了数千篇相关论文，几乎可以说“能想到的别人都想过了”！没有3-5年的积累，没有对前人工作的深刻理解和认识，谈何超越前人的“最好”？

正是秉持这样的理念，十多年来，无论在经费紧张的情况下，还是在遇到研究瓶颈的时候，我都没有想过要放弃人脸识别的研究。相反，我和高老师、陈老师共同带领的团队一方面深入研究以人脸表示和判别特征提取为核心的人脸识别的基础理论和方法，另一方面则面向实用人脸识别系统的开发研究其中涉及的人脸检测、特征定位、人脸分类等关键技术以及姿态、光照、老化等问题的解决方案，并积极收集整理越来越丰富的人脸图像数据资源用于验证上述理论、方法和技术。这些不懈的努力使得我们对该问题的认识越来越深刻，从而在FERET、FRGC、Multi-PIE以及CAS-PEAL等公开人脸识别评测集上取得了国际领先的结果，还赢得了前述两次人脸识别竞赛的第一名。同时我们的技术也得到了产业合作伙伴的认可，从而在众多实际系统中得到了应用。

## “美”的才是好的！

理念决定态度，态度决定行为。上述“做到最好”的理念显然是正向的、积极的，但怎么算做的“好”却不是一个简单的问题。

在我踏上学术道路之初，国内学术评价体系还停留在“数文章数”阶段，而且多数文章发表在国内刊物或会议上，其产生的国际影响力非常小。当时还处于学术童年期的我同样对该做什么样的研究，或者更具体的对该发表什么样的学术文章懵懂无知，我还记得当年得知自己投稿 ICIP2000 的文章被接收时的欣喜以及大家的祝贺（现在看来这个很一般的会议，在当时似乎并不简单，据统计当年 ICIP 来自大陆的文章仅十几篇）。对我这一个标准的“土鳖”来说，走出上面的误区并不容易。幸运的是，我有高老师高屋建瓴的指导，更有从 CMU 归国的陈熙霖老师的言传身教。我们这个团队在国内较早的明确了冲击领域顶级国际会议和国际期刊的学术目标，即从单纯的“数数”升级为有选择的“数数”。再后来我们更加明确的认识到，发表所谓顶级会议/期刊文章本身并没有实质意义，其实质在于顶级会议文章普遍具有的更大的潜在影响力。因此，学术评价的准则应该回归其本质，即实质影响力。所谓实质影响力，我个人认为更多体现在“两用”上，即“引用”和“应用”：前者强调的是学术界的认可，对基础研究尤其重要；而后者则体现的是工业界的认可，这一点在工业界日益重视前瞻研究和创新的今天显得更加重要。

需要进一步诠释的是，我认为上面所说的“两用”还只是现象或结果，并没有说明“好”的学术工作应该具有怎样的内涵。几年前，我们投稿的一篇论文被拒稿，审稿人之一认为我们的工作思路上有创新，实验也比较充分，但认为我们的工作不够“elegant”。这一事件或者说 elegant 这个词对我可谓醍醐灌顶，促使我再次认真考虑学术评价问题，并真正深刻意识到好的科研工作应该是美的，是优雅的。优雅和美更多是一种主观感受，但简单却是一种可度量的美之准则。科学并不排斥复杂，但一定会排斥无谓的复杂，在同等条件下优先选择简单。奥卡姆剃刀原则为此做了完美的诠释，即“如无必要，勿增实体”。实际上，上述的“两用”评价准则就是科学共同体对学术工作“美”与“不美”的投票：以人脸识别为例，过去数十年来学术界逐渐沉淀下来的高引用的工作，如 Eigenface, Fisherfaces, ASM, AAM, Gabor, AdaBoost, LBP 等等，从某个视角看无不具有简单之美。

## 再十年干将莫邪

尽管我们团队过去做了一些工作，我更愿意说的是“十年磨砺出钝剑”，与国际最高水平相比，我们还有很大的差距。尽管很难，但我始终期待着自己和团队能够在未来 10 年内，不仅仅把目前已然成形的钝剑磨砺的更加锋利，而更要做

出突破性的创新性工作，所谓“再十年干将莫邪”。在我的理想中，“干将”应该是能够突破人脸识别现实应用中诸多瓶颈的关键技术，这些技术能够让公共场合的犯罪分子无可遁形；而“莫邪”则是关键技术背后的创新理论和方法，它们应该简单而优美，得到国内外同行广泛的关注，从而成为经典之作。

最后，我想以我经常和团队成员说的一段话结尾，以飨读者：“在这个世界上，有很多同行在做着与我们相同的问题，有很多还来自国际名校，他们的智商不比我们低，甚至坦率的说，他们的智商恐怕要比我们高。而据我所知，他们的勤奋程度往往超过我们，请问我们有什么理由超越他们做出一流的学术成果？难道仅仅靠运气吗？！显然，我们唯一能做的就是比他们更加勤奋刻苦”。