信息类英语科技文写作

张 林
同济大学软件学院

摘要—为了向全球同行及时有效地传播科学发现和技术创新，研究人员需要撰写英文科技文。因此，英文科技文写作能力的培养是研究生培养过程中的必修环节。全世界每年新增论文的数量以百万计。尽管每篇论文的具体内容各不相同，同一领域的优秀论文之间会存在着很多共性特点。但对于初入此道的研究生们来说，这些共性特点似乎又是摸不着看不到、只能意会不能言传的。正因如此，作者萌生了写作本文的初衷，希望能够把优秀论文的共性规律以文字形式条理清晰地总结呈现给大家。在今后写作过程中，如果你时刻以这些共性规律作为论文的筋骨，你的最终成品离优秀论文就不会有太大的偏差。基于作者自身的理解、感悟与体会，本文从写作的总体原则、论文组成结构及各部分要点、表述的逻辑性、语言的润色四个方面阐述信息类英文科技文的写作要点，希望能够帮助论文写作水平处于初级阶段的研究生们。

索引词—科技文写作，英语写作要点

I. 总体原则

在论文的写作过程中，有一些有益的总体原则需要重视并仔细体会。这些原则需要指导并贯穿论文写作的全过程。

A. 以审稿人的思维和视角不断审视你的手稿

你的论文能够录用与否从形式上和流程上来说，完全决定于将来的 3~5 个审稿人。因此，在撰写过程中，有必要不断跳出来，从未来审稿人的角度全方位审视你自己的论文。需要认识到一个不幸的事实：对于审稿人来说，你是他们的同行。同行是冤家！你发了顶刊或顶会，有可能会挤占审稿人自己的份额。因此，审稿人一般来说是苛刻的而不是宽厚的：**不会因为你的论文有一个闪光点而投赞成票，反而很可能会因为它有一个点不闪光而投了反对票。**因此，在论文投出去之前，要不遗余力地使自己的论文在各方面都做到尽善尽美。牢记一句话：如果你认为你的论文是 shit，审稿人一定也会这么认为！

B. 学会模仿，抓住规律

任何事物都有其发展的客观规律，顶级论文也不例外。他山之石，可以攻玉。好的论文在各个维度都会具有共性特点，我们需要有意识地挖掘和总结。作为初学者，如果不知论文写作如何下手，不妨挑选 3~5 篇你这个领域的顶级牛文（从内容上、组织结构上、语言文字上都可作为典范的论文），打印出来，放在案头。在篇章组织结构、问题引述、前人工作评述、工作动机分析、创新要点总结、实验结果讨论、词句结构使用、逻辑层次构建等方面全方位多角度对模板论文加以仔细揣摩研究，并竭力把这些闪光点嵌入到你的论文之中。久而久之，你的论文写作整体水平也就随之提升了。

我仅从“词句结构使用”这一个维度举个例子，说明一下。下面列举了一些我从别人论文中学习摘抄的好的用词或句子结构，在写论文时会有意识地把它们引入到我自己的论文中：

1) In graph matching, patterns are modeled as graphs and pattern recognition amounts to finding a correspondence between the nodes of different graphs.
2) The main research focus in this theme is about designing efficient algorithms for approximately solving the quadratic assignment problem since it is NP-hard.
3) Palmprint is a combination of two unique features, namely, the palmar friction ridges and the palmar flexion creases.
4) Thus, in order to get rid of the adverse effects of the non-overlapped regions, we need to extract the common areas.
5) Compared with the conventional POC technique, the BLPOC (Band-Limited Phase-Only Correlation) proposed by Ito et al. (2004) is much more effective especially in dealing with biometric images.
6) As ridge frequency is often estimated based on ridge direction, reliable direction estimation is even more important.
7) An effective approach to account for within-class variations is by capturing multiple enrollment impressions of a finger.
8) Within the past two decades, numerous face recognition algorithms have been proposed as reviewed in the literature survey [1].
9) We conduct extensive experiments on publicly available databases to verify the efficacy of the proposed algorithm and corroborate the above claims.
10) An enormous volume of literature has been devoted to investigate various data-dependent feature transformations for projecting the high-dimensional test image into lower dimensional feature spaces.
11) Recent years have witnessed a growing interest in the search for sparse representations of signals.
12) The update of the dictionary columns is done jointly with an update of the sparse representation coefficients related to it, resulting in accelerated convergence.
13) Thus, approximate solutions are considered instead, and in the past decade or so several efficient pursuit algorithms have been proposed.
14) There is a variant of the vector quantization coding method, called gain-shape VQ. (变种这个词要用 variant)

15) The approaches to dictionary design that have been tried so far are very much in line with the two-step process described above.
16) Such sparse feature selection methods are, in a sense, dual to the support vector machine approach.
17) The rapid proliferation of digital imaging technologies has given rise to a growing number of applications which yield images for end-use by humans.
18) To this end, visual fidelity metrics have been developed in an attempt to accurately and efficiently quantify the fidelity of a distorted image relative to the original image in a manner that agrees with subjective judgments made by humans.
19) Ears are a particularly appealing approach to noncontact biometrics because they are relatively constant over a person's life and are unaffected by expressions, unlike faces.
20) A profound challenge in computer vision is the detection of the salient regions in an image.
21) Bolstered by growing general interest in large scale image retrieval systems, holistic image descriptors have become a topic of intense study in the computer vision literature.
22) While there may be a plethora of techniques for deriving fingerprint quality measures, all of them rely on the clarity of ridges and valleys of a fingerprint.
23) A large corpus of research makes use of the notion of intrinsic geometry, an umbrella term referring to geometric structures that remain invariant under non-rigid bendings and other types of transformations.
24) Most focus primarily on either new descriptors for person appearance, or on learning techniques for person re-identification.
25) Existing dehazing approaches can be roughly classified into two categories, i.e., the prior-based ones and the learning-based ones.
26) As a result, it is blessed with stable weak supervision and circumvents the issue of the lack of data that supervised methods suffer from.
27) The BIQA models learned by opinion-aware methods often have weak generalization capability, hereby limiting their usability in practice.
28) Unfortunately, thus far no opinion-unaware BIQA method has shown consistently better quality prediction accuracy.
29) The proposed method does not need any distorted sample images nor subjective quality scores, yet extensive experiments demonstrate its superior quality-prediction performance to the state-of-the-art opinion-aware methods.
30) Magerand et al. present a polynomial projection model for RS cameras and propose the constrained global optimization of its parameters by means of a semidefinite programming problem obtained from the generalized problem of moments method.

如果平时注意总结这些在词汇和结构上的好的用法，在写自己论文的时候有意识地加以体现，论文整体可读性一定会有很大提升。当然，这只是在用词用语这一小方面的做法举例；推而广之，你可以在你关注的论文的任何层面从已有的牛文中挖掘规律、总结经验，并最终为己所用。

C. 不能嫌麻烦

我确信，每一名研究生都有这样的经历（如果你没有这样的经历，只能说明你的导师从来没有认真帮你看过论文，我为你深表遗憾☺）：

你费了九牛二虎之力完成了论文初稿，发给导师，交差了账；之后，你刚想找个电影看看或者想出去打场篮球放飞一下心情，结果你的导师反馈神速，马上把你召至办公室，指着圈圈点点的打印版论文说：写的这是什么垃圾！（当然我本人一般会采用比较委婉的表达方式，但我的内心感受其实就是上面那句话）。接着就开始了一遍遍的修改之旅。之后，只要每次听导师说："你这个论文***地方写得还不行、没表达清楚，还得要好好改一改……"，你就会觉得一阵胸闷。

没办法！必须清醒地认识到，这是修成正果的必经之路。**不经历这样一个痛苦的锤炼过程，你的论文水平（本质上是逻辑思维能力、抽象概括能力和语言表达能力）很难得到提升，或者换一句更直接更本质的话来说：实际上，你连话也不会说！**所以当导师再提出修改意见时，你需要深吸一口气，然后大声对自己说：不嫌麻烦，我一定能行！

D. 清晰的思路比好的英语水平更重要

很多年以前，我在香港理工读博士期间，我的导师张磊教授曾说：**如果你的论文读起来让人感觉逻辑混乱、问题没有讲透彻，本质原因不是你的英文不好，而是你的思路还没有理顺、还比较混乱，即使让你用中文讲一遍，你还是讲不清楚！**

如果导师对你的论文也有类似评价（对于初级水平的研究生来说几乎是必然的），这时候你需要老老实实地坐下来，把你想表达的内容用中文写在纸上，然后反复修改，直至逻辑清楚、表达通顺、没有歧义性和中文语法错误。然后，再把拟好的内容打入电脑里，接下来就是 AI 上场了。当代 AI 技术最成功的应用之一就是机器翻译，现在百度和谷歌的翻译水准和效率，不出意外的话肯定高过你。因此，可以用百度和谷歌对你起草的中文内容对照翻译一下。你再以机器翻译结果为蓝本进行小幅调整即可。如果你觉得机器翻译的英文读不通、有问题，大概率是因为你的中文草稿有歧义性或语法错误，这就需要你再对中文文稿进行修正并迭代进行上述过程。根据我多年的经验，这样一种论文写作过程看似有些效率低下，实则不然，因为它几乎可以保证一遍做对，返工率低。我有很多研究生，论文初稿准备工作并没有认真遵照我建议的这个流程执行；当把论文交给我之后，我不得不陪他（她）再一起退回到这个阶段，重新再来。该扎扎实实一步一步推进的工作，一步也不能偷懒，否则怎么跑过去的还得怎么跑回来，欲速则不达！

E. 论文不是说明书

初学者很容易把论文写成流水账，按部就班的写出第一步干了什么，第二步干了什么，最终得到了什么结果，这样写出来的材料实际上更像是说明书，而不是论文。论文的写作显然要比说明书有更高的要求。

《说文》：论者，议也。论文写作重要的是要落在“论”字上面，也就是说，论文不单单要阐明“新方案（方法）是如何设计的？具体运行步骤（模块）包括哪些？实验结果是什么？”等客观事实，更要阐明“方案设计的动机（motivation）是什么？这么做的合理性（理论支撑）何在？实验结果说明了什么问题？对于科学界将来会有什么意义或者启示？”等主观论述分析。**论述分析实际上要体现在论文中的每一个环节中，包括对研究现状及研究问题的分析、对所提新方案的设计合理性（设计动机）的分析、对文献中已有的同类型工作的分析、对实验结果的分析等等。**

F. 细节决定成败

经过近代一百多年的发展，英文科技文的写作在宏观上和微观上都形成了一套基本规范。宏观规范容易掌握和理解，但微观规范往往存在于每个审稿人的心中，很少有人把它们完整地总结出来。若想提升论文的命中率，必需要注意一些细节，因为它们很可能是审稿人所关注的。我总结了如下一些在微观层面上论文写作需要注意的一些问题，它们虽小，但有时却是致命的。

1) 一个句子不能以参考文献引用开头，比如，[1] is a typical approach belonging to this category，这是不可以的；可以改成，The approach proposed in [1] is a typical one belonging to this category。
2) 在摘要中，不能出现参考文献引用。按照科技文写作规范，摘要是可以独立于正文而存在的。
3) 少用主动语态。科技文是技术化的、严谨的，因此要尽量保持文风的客观性。鉴于此，科技文的语态要以被动语态为主，少用主动语态（但不是说完全禁用），尤其是在阐述客观规律、分析实验结果的时候。比如，我们可以说 it can be seen that，尽量不用 we can see that。
4) 如果参考文献的作者多余两个人，在文中提及作者的时候应该是：first_author_surname *et al.*；如果文献的作者只有两个人，在文中提及作者的时候应该是：Zhang and Wang。
5) 插图必须要是矢量图格式。BMP、JPG 这类图片格式都是不行的。EMF、EPS、PDF 格式一般是可以的，具体要看你用什么制作工具了。总而言之一个要求：**最终论文的 PDF 版任意放大后，论文中的插图不能糊！**
6) 插图制作要精益求精，内容要丰富，文字元素和图示化元素组合合理，颜色搭配要考究，让人一看要赏心悦目。可以参考图 1、图 3。
7) 插图下面的说明文字（figure caption）不能过于简略。一般要求说明文字要具有自解释性，也就是说，读者可以不依赖于正文而基本理解插图所表达的涵义，参考图 1、图 3。
8) 全文数学符号的格式（字体、字号、是否粗体斜体等）要统一，不管这个符号是出现在单独编号的公式中、正文中、插图中、还是表格中。
9) 插图或表格与正文的间距要适当。很多情况下，Latex 的默认排版会使得图表与正文的间距太大，这时可以用 \vspace 命令把间距缩小一些。但要注意，\vspace 命令一般只用于图表的后面，不能过度滥用。
10) 注意冠词的使用。可数名词，要么是复数形式，要么前面必须有冠词；不可数名词一般不加冠词，但如要表示“强调、特指”时，需要加定冠词。
11) 要确保各种不同类型参考文献格式的一致性。期刊论文、会议论文、书、书的章节、博士论文、在线材料等不同类型的参考文献有不同的格式写法，但同一篇论文里面同一个类型的参考文献格式务必要保持高度一致。
12) Latex 里面，双引号”的输入法是’’；双引号“的输入法是``。
13) 单独编号的公式，如果恰巧是句子的结尾时，不要忘记公式后面要加个句号；如果不是句子结尾，则不用加句号。
14) 当一个会议对投稿论文的页数有上限要求时，要尽量顶着上限写。

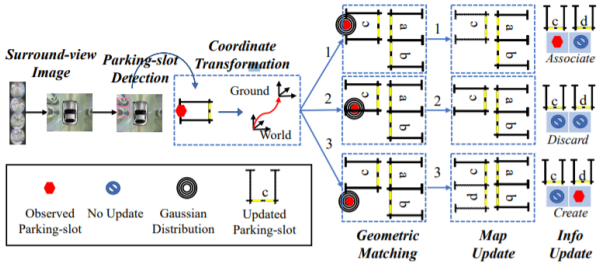

Figure 4: Parking-slot association. Parking-slot association is based on geometric distances between parking-slots in the map and the observed one. According to the distances, the parking-slot observation will be (1) associated with one in the map, (2) discarded as abnormal observation, or (3) regarded as a new one.

图 1. 论文中的插图示例。

II. 论文组成结构及各部分要点

英文科技文的章节布局有着基本固定的结构，不能自己随意发挥。在这一节中，介绍一下论文的组成部分和各部分的写作要点及注意事项。

A. 摘要

采用“narrow down”的写法，先用一句话说一下大的背景领域，再逐步过度到你的小的研究领域，也可提一下它的研究意义，最后过渡到 research gap。针对这个 research gap，阐述你论文的主要创新点（以及它们的合理性）、方案和贡献。最后，介绍论文所取得的效果和成果。如果你的代码及数据是可以公开的，可以把公开网址链接写上，增加论文工作的可信度。下面给出 3 个范例，供体会参考。

范例 1：

In the automobile industry, recent years have witnessed a growing interest in developing self-parking systems. For such systems, how to accurately and efficiently detect and localize the parking slots defined by regular line segments near the vehicle is a key and still unresolved issue. In fact, kinds of unfavorable factors, such as the diversity of ground materials, changes in illumination conditions, and unpredictable shadows

caused by nearby trees, make the vision-based parking-slot detection much harder than it looks. In this paper, we attempt to solve this issue to some extent and our contributions are twofold. First, we propose a novel deep convolutional neural network (DCNN)-based parking-slot detection approach, namely, DeepPS, which takes the surround-view image as the input. There are two key steps in DeepPS, identifying all the marking points on the input image and classifying local image patterns formed by pairs of marking points. We formulate both of them as learning problems, which can be solved naturally by modern DCNN models. Second, to facilitate the study of vision-based parking-slot detection, a large-scale labeled dataset is established. This dataset is the largest in this field, comprising 12165 surround-view images collected from typical indoor and outdoor parking sites. For each image, the marking points and parking slots are carefully labeled. The efficacy and efficiency of DeepPS have been corroborated on our collected dataset. To make our results fully reproducible, all the relevant source codes and the dataset have been made publicly available at https://cslinzhang.github.io/deepps/.

范例 2：

Biometric authentication has been found to be an effective method for recognizing a person's identity with a high confidence. In this field, the use of palmprint represents a recent trend. To make the palmprint-based recognition systems more user-friendly and sanitary, researchers have been investigating how to design such systems in a contactless manner. Though substantial effort has been devoted to this area, it is still not quite clear about the discriminant power of the contactless palmprint, mainly owing to lack of a public, large-scale, and high-quality benchmark dataset collected using a well-designed device. As an attempt to fill this gap, we have at first developed a highly user-friendly device for capturing high quality contactless palmprint images. Then, with the developed device, a large-scale palmprint image dataset is established, comprising 12,000 images collected from 600 different palms in two separate sessions. To the best of our knowledge, it is the largest contactless palmprint image benchmark dataset ever collected. Besides, for the first time, the quality of collected images is analyzed using modern image quality assessment metrics. Furthermore, for contactless palmprint identification, we have proposed a novel approach, namely CR_CompCode, which can achieve high recognition accuracy while having an extremely low computational complexity. To make the results fully reproducible, the collected dataset and the related source codes are publicly available at http://sse.tongji.edu.cn/linzhang/contactlesspalm/index.htm.

范例 3：

The semantic SLAM (simultaneous localization and mapping) system is an indispensable module for autonomous indoor parking. Monocular and binocular visual cameras constitute the basic configuration to build such a system. Features used in existing SLAM systems are often dynamically movable, blurred and repetitively textured. By contrast, semantic features on the ground are more stable and consistent in the indoor parking environment. Due to their inabilities to perceive salient features on the ground, existing SLAM systems are prone to tracking loss during navigation. Therefore, a surround-view camera system capturing images from a top-down viewpoint is necessarily called for. To this end, this paper proposes a novel tightly-coupled semantic SLAM system by integrating Visual, Inertial, and Surround-view sensors, VISSLAM for short, for autonomous indoor parking. In VISSLAM, apart from lowlevel visual features and IMU (inertial measurement unit) motion data, parking-slots in surround-view images are also detected and geometrically associated, forming semantic constraints. Specifically, each parking-slot can impose a surround-view constraint that can be split into an adjacency term and a registration term. The former pre-defines the position of each individual parking-slot subject to whether it has an adjacent neighbor. The latter further constrains by registering between each observed parking-slot and its position in the world coordinate system. To validate the effectiveness and efficiency of VISSLAM, a large-scale dataset composed of synchronous multi-sensor data collected from typical indoor parking sites is established, which is the first of its kind. The collected dataset has been made publicly available at https://cslinzhang.github.io/VISSLAM/.

B. 引言

引言部分也是采用由大到小的写法，从大的应用（领域）背景逐步过渡到你这篇论文所在的小领域。然后，逐步明确现有研究工作的不足或空白之处，也就是阐明“research gap”。

接下来，针对 research gap，论述你当前这篇论文的主要想法（idea）是什么、大致是如何做的、为什么这么做、新方法（理论、系统等）潜在的好处和优点等；这部分基本上是以定性描述为主，**要注意论述分析，而不是简单的阐述你是如何做的**。下面给出一个范例，来体会一下论文分析大致应该如何做。

范例：

Interestingly, due to the characteristics of the two categories, prior-based methods are relatively better for restoring visibility whereas learning-based methods are preferable for improving the result's realness. Fig. 1 provides dehazing results of (a) the learning-based AODNet [5] and (b) the prior-based DCP [8] to illustrate this phenomenon. As we can see, DCP's result has less haze but more artifacts, whereas the result of AODNet is in high realness but with more haze. In Appendix A, we provide some theoretical explanations for the preference of prior-based and learning-based methods. To further improve the dehazing results, it is a natural idea to exploit the advantages of both categories, but surprisingly such a simple idea has seldom been explored in the literature. In this work, based on the above-mentioned findings, we propose a two-stage weakly supervised dehazing framework, RefineDNet (Refinement Dehazing Network), to merge the merits of the two categories.

Specifically, in the first stage, RefineDNet restores the visibility of the input hazy image by producing preliminary results with DCP. We embed DCP dehazing in our framework to enable end-to-end training and evaluations. In the second stage, RefineDNet improves the realness of the preliminary dehazed image and the quality of the transmission map by refining them using two refiner networks. During training, we

update the refiner networks via adversarial learning with a discriminator on unpaired images. This weak supervision with unpaired data is beneficial to dehazing because it is possible to collect a large amount of unpaired images from the real world to train our model. In this way, RefineDNet suits to process real-world foggy images better than supervised methods that are trained on simulated images and may overfit those data.

Beside the refined dehazed image, RefineDNet reconstructs another dehazed image using the hazy input and refined transmission. Since the refined and the reconstructed dehazed images are generated in different ways, they are unlikely to perform the same in all regions. It is highly possible that either of them may outperform the other in some regions. Thus, fusing better regions in either of them can boost performance. To this end, we propose a perceptual fusion strategy to fuse the refined and the reconstructed dehazed images. In this strategy, greater weights are assigned to regions that are closer to natural images. To obtain such weights, we exploit powerful features in the field of image quality assessment (IQA).

With the two-stage dehazing strategy, RefineDNet divides the dehazing task into two less intractable subtasks, namely, visibility restoration and realness improvement, and leverages priors and learning to handle the two subtasks, respectively. Since priors and learning are used in separate stages of RefineDNet, they are unlike to affect each other. Thus, RefineDNet merits the advantages of both prior-based and learning-based methods. Besides, RefineDNet only needs to remove artifacts in the refinement stage, and thus, its learning encounters less ambiguity of dehazing. As a result, it is blessed with stable weak supervision and circumvents the issue of the lack of data that supervised methods suffer from. To support our claims, we show that even implemented with basic backbone networks, RefineDNet is able to outperform state-of-the art supervised methods on both indoor and outdoor datasets. Moreover, since there is no off-the-shelf outdoor training set for RefineDNet, we built an unpaired outdoor training dataset, RESIDE-unpaired, using images from RESIDE [11].

上述范例中，红色标记的地方都是作者自己对他新提出方法的分析讨论的地方。对创新点的论述分析需要讲清楚创新点的 rationale 在哪里、它为什么会产生预期的好结果、它的特性是什么等等。

引言的最后需要总结这篇论文的主要贡献和创新点。引言中的主要贡献和创新点的总结是整篇论文的灵魂，必须要认真推敲。一般来说，主要贡献和创新点可以分条目总结，3~4 条即可。对于每一条，可以按照“应对什么问题？我们提出或做了什么？具体如何做的？技术特点或特性是什么？在理论和实际应用上达到了什么效果？有什么重大意义？”这个模式来展开。当然，根据论文内容的实际情况，这几个要点并不需要都顾及到，但一般来说，**要有一句定性的语句来阐述你这一条贡献的理论（技术）特性或者给研究领域带来的价值**。关于引言部分中的主要贡献和创新点总结，下面给出 2 个范例，供大家参考体会，其中红色标记部分为对该条贡献的定性分析或总结。

范例 1：

Taking aforementioned analysis into considerations, we attempt to build a semantic SLAM system, namely VISSLAM (a SLAM system integrating visual, inertial, and surround-view sensors), specially for the task of autonomous indoor parking. Our contributions can be summarized as follows:

(1) Specially designed for navigation in the indoor parking site, VISSLAM is the first tightly-coupled semantic SLAM system that fully explores parking-slots detected in surround-views in its optimization framework. Surround-view images are synthesized online from bird’s-eye views generated from a surround-view camera system, comprising four fisheye cameras. Since VISSLAM can construct maps with semantic parking-slot information, it can be naturally integrated into a high-level self-parking system.

(2) In order to improve the localization accuracy and to construct the semantic map, in VISSLAM parking-slots in surround-view images are leveraged for optimization in which they are modeled as an adjacency term subject to the existence of adjacent neighbors and a registration term constraining by registering between the observed parking-slots and their positions in the world coordinate system. Experiments demonstrate that the semantic constraints induced by parking-slots can significantly improve the performance of VISSLAM.

(3) At present, there is no publicly available dataset for SLAM research that contains visual information of surround-views. To fill this gap to some extent, in this work we established a large-scale dataset comprising synchronous multi-sensor data from typical indoor parking sites. It will benefit further SLAM studies, especially conducted for autonomous indoor parking.

范例 2：

The main contributions of our work are as follows:

(1) We propose a two-stage weakly supervised framework RefineDNet which first adopts prior-based DCP to restore visibility and then employs GANs to improve realness. It is demonstrated that RefineDNet integrates the advantages of both prior-based and learning-based dehazing methods and generates visually pleasing results with high visibility. Moreover, due to the two-stage dehazing strategy, RefineDNet boasts effective weak supervision with unpaired foggy and clear images, which avoids the issue of the lack of paired data for supervised methods.

(2) We propose a novel perceptual fusion strategy to blend different dehazing results. Our experimental results demonstrate that this strategy is effective with performance gain in various datasets.

(3) We also construct a necessary unpaired dataset with 6,480 outdoor images to facilitate the relevant studies of weakly supervised dehazing approaches.

Table 1
Summary of existing contactless palmprint acquisition devices in the literature.

Refs.	Year	Main features	Shortcomings
Chen et al. [26]	2007	Two 3U 23-watt lights were used to provide illumination and were arranged in appropriate positions; a normal digital camera was used to capture the image.	Hardware parts were not integrated into a complete usable system; images can only be manually captured; it can only be used in lab.
Kumar [27]	2008	Hardware parts are enclosed in a semi-closed box; the palmprint image can be captured against a dark background.	The free space left for the hand in the box is quite limited and thus it is highly likely that the user's fingers may touch the walls of the box. So, it may raise users' hygienic concerns.
Hao et al. [28]	2008	It is a multispectral system; when sample images are being captured, 6 groups of LEDs will be switched on sequentially, and thus 6 images can be collected.	Palmprint acquisition using such a device is quite time-consuming; the light may hurt the user's eyes.
Michael et al. [25]	2008	A 1.3 mega pixel web camera and a 9W white light bulb are mounted inside a semi-closed box.	There is no top cover; the quality of collected images may vary much; when it works, its intense light may directly enter into the user's eyes.
Poinsot et al. [30]	2009	It collects palmprint images using a Logitech QuickCam Pro-9000 webcam against a green background under natural illumination.	Neither a housing nor additional lights are used in their setup; it cannot be deployed in practical applications.
Ferrer et al. [31]	2011	It could capture palm images under visible and IR lighting conditions simultaneously.	It has no housing so it has similar potential problems as Michael et al.'s [25].
Morales et al. [32]	2012	Two plates delimit the area through which the user passes his/her hand with a vertical movement.	The moving speed of the user's hand cannot be controlled so the quality of collected images may vary greatly.
Aykut et al. [33]	2015	Their capturing device is a low-cost CCD camera with a DC-auto iris lens.	The distance between the user's hand and the camera cannot be restricted, so captured images may be out of focus; the LED light may hurt the user's eyes.

图 2. 可以用表格的形式对相关工作的特点和特性进行总结。

范例 3：

To this end, our paper investigates this practical problem and the contributions are summarized as follows:

(1) We propose a new error model namely "Bi-Camera error", which can measure the photometric discrepancy between two corresponding pixels $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$ on two images, where $\boldsymbol{p}_i$ and $\boldsymbol{p}_j$ are the projections from the same physical point $\boldsymbol{p}_G$ on the ground plane. Cameras' extrinsics are embedded in the projection relationships between $\boldsymbol{p}_G$ and $\boldsymbol{p}_i$ (or $\boldsymbol{p}_j$). Thus, by minimizing the Bi-Camera errors, optimal extrinsics can be worked out.

(2) Based on Bi-Camera error model, an online extrinsics correction algorithm for the surround-view system, namely (Online Extrinsic Correction for the Surroundview system) (OECS for short), is proposed. In OECS, cameras' optimal extrinsics are figured out by iteratively minimizing the system's overall Bi-Camera error. It needs to be noted that OECS follows a sparse direct framework, implying that it does not depend on visual feature points. Hence, OECS requires less on its working conditions and is quite robust.

(3) Within the sparse direct framework, a novel pixel selection strategy is proposed. Using such a selection strategy based on color matching and gradient screening, noise and unmatched objects between images captured by adjacent cameras can be eliminated effectively. BiCamera errors are then only computed on the selected positions. Such a pixel selection scheme can effectively improve OECS's speed and robustness.

C. 相关工作总结

相关工作主要指的是和你的论文解决相同或相近问题的论文。另外，如果你的工作主要使用了某项理论或基于某些前期工作，而这些理论或前期工作并不是广为人知的，也可以（且需要）在相关工作这一节中加以介绍阐述。

对文献中相关论文的总结与评述，一般需要按照某种准则分类进行，这样会使得条理更加清晰，切忌把材料毫无章法的堆砌在一起。在大多数情况下，需要总结出相关文献的不足之处，可以对几个同类型方法的共性不足之处集中陈述。另外，也建议在对一大类相关工作评述完之后，简要说一下本文与他们的不同之处，以体现出你论文的先进性。下面给出一个对相关工作进行总结的范例，供大家参考。其中，红色标记的地方是对某一个类型方法特性（或不足之处）的总结，或者是把这些方法同本文提出方法的某些特性进行了对比分析。

范例：

The restoration of underexposed images has been a longstanding problem with a great progress made over the past decade. Here we divide them into two categories, plain ones and data-driven ones, based on whether or not supervised learning is used.

Plain methods. Conventional image enhancement methods such as histogram-based methods [7, 8, 9, 10] can be explored to enhance underexposed images, but in most cases, their efficacy is quite limited. Yuan and Sun [11] proposed an automatic exposure correction method using S-curve tone mapping. Zhang *et al.* [12] designed an unsupervised scheme to estimate the best-fitting S-curve of the input. The parameterized S-curve adopted in these methods may compress the mid-tones and the output images look too flat and unnatural. Early attempts [13, 14] based on Retinex theory remove the illumination and directly extract the reflectance as the enhanced results. Later work in this branch mainly focused on the estimation [3, 4] and adjustment [2] of illumination. These Retinex-based methods assume that the input images are noise-free and amplify the latent noise in dark regions. Fu *et al.* [15] estimated illumination and reflectance simultaneously. Li *et al.* [5] further introduced the noise term to the classic Retinex decomposition. These two methods suppress noise via imposing a constraint on the reflectance or noise. Differently, we applied the illumination guidance to noise estimation, thus

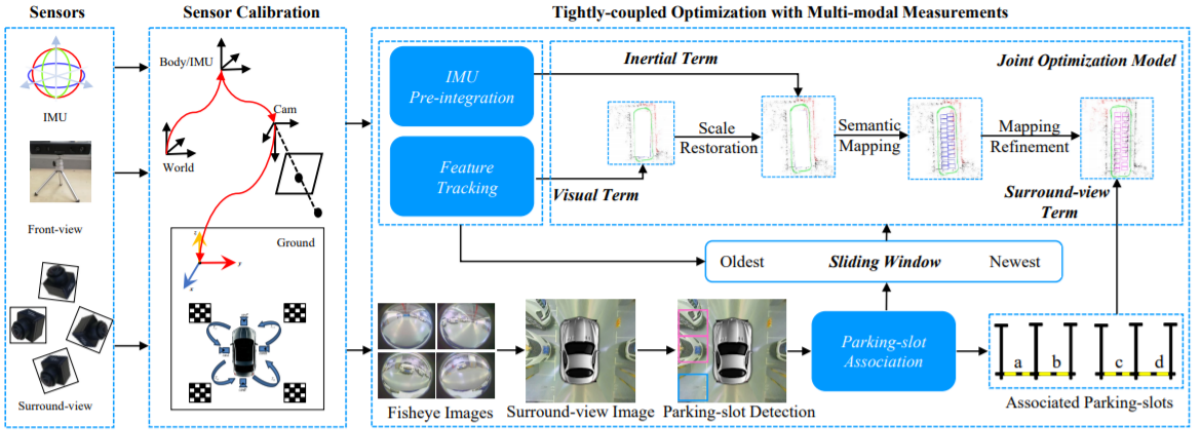

Figure 1: The overall processing pipeline of VIS$_{SLAM}$. Multi-modal sensors are first spatially registered with one another. Visual features are detected and tracked to construct a 3D map of an indoor parking site with no scale. By aligning pre-integrated IMU measurement with the visual features in the front-view image, a map with metric scale can be obtained. In order to build a semantic map suitable for autonomous indoor parking, parking-slots in each surround-view image are detected and geometrically associated to constitute a surround-view constraint. The visual term, IMU term as well as the surround-view term are integrated into VIS$_{SLAM}$ during optimization. Joint optimization is performed in a sliding window, giving a trade-off between the speed and the flexibility.

图 3. 总体方案流程图或系统结构图示例。

performing more targeted denoising in the dark.

Data-driven methods. Black box models [16, 17, 18, 19] roughly follow such a pipeline: first collect or synthesize a dataset containing input-output pairs, and then find the mapping relationship or train an exposure correction model based on the dataset. Based on Retinex theory, Shen *et al*. [20] proposed MSR-net based on multi-scale Retinex theory and trained it on synthesized pairwise images. Wang *et al.* [21] trained an illumination mapping estimation network on the new dataset they built, including underexposed images and expert-retouched references. Wei *et al.* [22] and Zhang *et al.* [23] trained decomposition networks on a dataset containing low/normal light image pairs. The performance of these supervised-learning-based methods highly depends on the training dataset despite the fact that building such a dataset including various types of illumination and contents is a challenging task itself.

如果篇幅允许，你也可以用表格的形式对相关文献的主要特性加以归纳总结，这往往会使得评审人认为你的调研工作做得非常认真细致，从而有效提升了你论文的印象分。例如，图 2 所示的表格总结了文献中现有非接触式掌纹采集设备的特性。

D. 工作主体介绍

在介绍工作主体内容时，章节安排可以比较灵活。根据内容性质，可以写在一个 Section 里，也可以分在几个 Section 里。一般来说，这部分内容是你自己做的，故而会比较容易把握一些，只要条理清晰的把工作内容组织阐述清楚即可。有几个有益的建议供大家参考。

- 总体流程图或系统结构框图

系统或方案类的论文必须要有一个总体结构图或流程图。这个图是工作介绍部分的核心，对内容介绍起到了提纲挈领的作用，需要全面体现出方案的处理流程或系统主要模块之间的逻辑关系。这个图的重要性不言而喻，它对于全篇来说往往起到了画龙点睛的作用，因此必须要做的精益求精、内容丰富而不凌乱。图 3 给出了一个系统总体方案结构图示例。

- 算法伪码

当描述算法或较为复杂的方案时，最后可以用伪代码的形式对处理流程进行总结，给读者一个宏观、完整、形式化的介绍，让他们更加清晰无歧义地理解你的工作。图 4 给出了一个通过算法伪码形式总结主体方案流程的示例。

E. 实验与讨论

实验与讨论这一节是要验证你所提出的方案或系统的性能，尤其是需要展示出你本文的工作和其他已有的同类型工作相比，你的优势在哪里。

一般来说，实验部分上来要有一小节概括式的说一下实验的设计与设置问题，一般要包括这几个方面：1）你的算法（方案、系统）的一些重要实现细节，包括实现语言、关键超参数设置、运行平台系统等；2）如果后面要比较计算时间的话，需要给出实验的测试软硬件环境；3）所用的实验数据来源；4）要进行对比对照的已有同类型方案；5）性能指标，也就是在哪几个方面对方案进行评测，评测指标是什么？（当然，如果针对每个测试维度都有不同的实验数据、对照方案、评测指标的话，相关设置信息需要分配到与某个测试维度有关的小节中去。）

然后，就要从几个不同的方面（维度）来对你的新方法进行评测。一般情况下，要对照比较已有同类型方法。比较是多维度的，同时，除了定量的比较之外，也可以包括定性的分析。

实验部分除了汇报定量的实验结果之外，更重要的是要进行讨论和分析，也就是要讨论清楚从实验结果之中我们

能够观察到什么、能够得到什么结论或启示；非常忌讳的是，只给出了具体数字，而没有对结果的讨论分析。实验结果分析的写法一般可以有两种模式：一是针对每一个实验维度，可以在汇报完实验结果后，直接进行讨论分析；也可以在汇报完针对所有维度的实验数据之后，再额外增加一个讨论小节，对所有实验结果进行综合讨论分析。一般来说，前者用的多一些（尤其是针对期刊长文来说）。下面列举了 3 个范例，供大家来揣摩体会实验结果的分析讨论部分应该如何展开。

Algorithm 1 Algorithm for $\mathbb{T}$'s construction

Input:
The priori map, $\mathcal{M}$
Output:
The location-related ANN search tree, $\mathbb{T}$
Gridding $\mathcal{M}$ according to a certain resolution, get all passible grids $\mathbb{G}$
$\mathbb{T} \leftarrow \{\}$, set $\mathbb{T}$ as a new empty tree
Set the signature histogram's dimension as $\tilde{d}$
Set the sensor's maximum range as $\tilde{r}$
$\tilde{r}' \leftarrow \tilde{r} + \tilde{\epsilon}$
$\overline{r} \leftarrow \tilde{r}' \div \tilde{d}$
for all $\mathbf{g} \in \mathbb{G}$ **do**
  Scanning at $\mathbf{g}$ from $\mathcal{M}$ to get a virtual scan $\mathbb{S}_v$
  Assign $\tilde{r}'$ to all points that exceed $\tilde{r}$ in $\mathbb{S}_v$
  $\boldsymbol{s} \leftarrow [0, ..., 0]$, set siganture $\boldsymbol{s}$ as an array with size of $\tilde{d}$, each element with value of 0
  for all $\mathbf{p} \in \mathbb{S}_v$ **do**
    Get $\mathbf{p}$'s scan distance d
    $i \leftarrow \lfloor d \div \overline{r} \rfloor$
    $\boldsymbol{s}[i] \leftarrow \boldsymbol{s}[i] + 1$
  end for
  Normalize $\boldsymbol{s}$ to $\overline{\boldsymbol{s}}$
  Insert $\overline{\boldsymbol{s}}$ into $\mathbb{T}$
end for
return $\mathbb{T}$

图 4. 以伪码形式来总结算法或方案的总体流程示例。

范例 1：

The results in Table II lead us to the following conclusions. First, the prediction performance of opinion-aware methods tends to drop with decreases in proportion of the training subset. However, it is very interesting to observe that the partition ratio has little effect on the performance of the opinion-unaware methods. Indeed, when the partition ratio is low (e.g., 10%), the opinion-unaware methods, especially IL-NIQE, actually perform better than all the opinion-aware methods. Second, IL-NIQE performs much better than the other two opinion-unaware methods, NIQE and QAC. While this is to be expected given that IL-NIQE may be viewed as a feature-enriched version of NIQE, the increment in performance is quite significant. Third, on TID2013, LIVE and MD1, IL-NIQE achieves performance comparable with the opinion-aware models CORNIA, BRISQUE and BLIINDS2, and performs better than the other two opinion-aware methods over all partition ratios. Finally, IL-NIQE performs significantly better than all of the competing models on CSIQ and MD2, even though it does not need distorted images or human subjective scores on them to train the model.

范例 2：

Experimental results are summarized in Table 6. In Table 6, we list the recognition rate achieved by each method. In addition, we also list the time cost consumed by one identification operation for each method. Given a test sample, the time cost for one identification operation includes the time consumed by the feature extraction and the time consumed by matching the test feature with the gallery feature set. In Fig. 10, the CMC (cumulative match characteristic) curves for all methods evaluated on our dataset are shown.

Based on the evaluation results listed in Table 6 and Fig. 10, we could have the following findings. With respect to the classification accuracy, the proposed method CR_CompCode could achieve comparable results with the state-of-the-art methods. Its rank1 recognition rate is 98.78% on our benchmark dataset. Meanwhile, CR_CompCode runs greatly faster than all the other competing methods. For example, it is about 1979 times faster than SIFT+AlignedCompCode [45]. Under our experimental settings, it costs CR_CompCode only 12.48 ms to complete one identification operation against a gallery set comprising 6000 samples from 600 classes. The low speeds of the other methods evaluated can mainly be attributed to the brute-force matching strategy they adopt for identification. Take SIFT+AlignedCompCode [45] as an example. When it is used for identification, it needs to match the test sample to all the gallery samples and then figures out the most matched one. Meanwhile, when SIFT+AlignedCompCode matches a pair of palmprint samples, the following operations are conducted: SIFT matching, image alignment, CompCode extraction from the test sample, and CompCode matching. That explains the low speed of SIFT+AlignedCompCode for identification.

Based on above discussions, we recommend using the proposed CR_CompCode method for contactless palmprint identification since such an approach can achieve a distinguished high recognition accuracy while having an extremely low computational complexity. It is quite suitable for large-scale identification applications.

范例 3：

From the results presented in Tables X and XI, we can draw a number of conclusions.

1) The classical methods [25], [31], [34], [41] cannot achieve satisfied results. The root cause is that they are based on primitive visual features (edges, lines, or corners) and are not robust to complicated scenes and various imaging conditions.

2) PSD_L exhibits clear performance advantages over its counterparts depending on low-level visual features. The better performance should be attributed to the machine learning (ACF + Boosting) based marking-point detection scheme used in PSD_L.

3) In all cases, the performance of DeepPS surpass that of PSD_L by a large margin. Compared with PSD_L, the superior performance of DeepPS is mainly for two reasons. First, for detecting marking-points, DeepPS adopts a state-of-the-art DCNN-based object detection framework YoloV2, rather than ACF + Boosting as used in PSD_L. The results shown in Fig. 10 and Table VI clearly demonstrate that the YoloV2-based framework works much better than the ACF+Boostingbased

scheme in terms of the detection accuracy and the localization accuracy. Second, to determine the validity and the type of an entrance-line, PSD_L uses a complicated rule-based scheme while DeepPS formulates it as a local image pattern classification problem which can be more concisely and robustly solved by a standard DCNN model.

按照近年来信息领域（尤其是与机器学习有关的领域）论文的写作习惯，在介绍完对照实验结果以后，一般还要进行消融分析实验（Ablation Study），尤其是在算法类的论文里面。消融分析实验的目的主要是要证明你的系统或算法各个模块都是有用的，都可以对最终整体性能的提升做出贡献。比如说，你为了提升 baseline 的性能，给它加了两个模块 A、B，加完之后效果果然提高了很多。于是你急急忙忙开始写论文，写到你的贡献，你给了两条：模块 A，模块 B。但是这样写有个问题：尽管 AB 同时加上去对模型有提升效果，但是你并没有证明 A、B 两个模块分别都是有意义的。所以，为了验证 A、B 两个模块是不是真的都有用，你需要做 ablation study。方法也很简单：在 baseline 的基础上加上模块 A，看效果。在 baseline 的基础上加上模块 B，看效果。在 baseline 的基础上同时加上模块 A 和 B，看效果。然后结果可能是，实验 1 和实验 2 的结果都不如实验 3，那么说明 A 和 B 都是有用的。然而，也有可能你会发现实验 1 的结果和实验 3 一样，甚至更好；这就说明你的想法是有问题的，模块 B 其实并没有起到作用，提升只来自于模块 A。综上所述，ablation study 就是你在同时提出多个思路提升某个模型性能的时候，为了验证这几个思路分别都是有效的，而进行的控制变量实验。如果篇幅允许（一般是期刊长文）消融实验可以做的很全面，但如果是篇幅受限的会议论文，可以精简一些，只做和关键模块（参数）有关的消融实验。但无论如何，你的论文最好是要有消融实验内容，以显示你实验的充分性、完整性、可靠性。下面，我举 2 个简洁的消融实验写法的例子，供大家参考。

范例 1：

Ablation Study. We performed detailed ablation analyses to validate the contribution of each error term during optimization of VIS_{SLAM} in three respects, the revisiting error, the average distance between adjacent parking-slots and the time cost, and the results are presented in Table 5. It can be seen from the table that both the revisiting error and the time cost of a visual-inertial error term based SLAM system can reach satisfied performance, which are 0.199m and 0.045s/frame, respectively. But it is not suitable for autonomous indoor parking, since the VI-SLAM system provides no semantic information of parking-slots during navigation. If we simply incorporate parking-slots in the surround-view image for tracking without optimization, they will compromise the SLAM system and lead to a huge revisiting error. But if the parking-slots are incorporated in optimization, both the revisiting errors and the adjacency gaps can be significantly diminished, confirming the effectiveness of VIS_{SLAM}.

Table 5: Optimization results using various error terms.

Mode	Revisiting error (m)	Adjacency gap (m)	T (s/frame)
V-I	0.199	-	0.045
S	0.317	0.243	0.040
VIS	0.028	0.106	0.067

范例 2：

Ablation Study. We perform an ablation study on MEF dataset to quantitatively evaluate the merit brought by each term in the loss function of RRDNet by combining different loss terms. The results are summarized in Table 2. It can be seen that adding texture enhancement loss term and noise estimation loss term can obviously improve both contrast and naturalness of the restored outputs. The progressive improvement of the performance demonstrates the effectiveness of each loss term.

Table 2. Ablation study.

Settings	L_c	$L_c + L_n$	$L_c + L_t$	$L_c + L_t + L_n$
CPCQI	0.8357	0.9299	1.0178	**1.0981**
NIQE	3.7438	4.1115	3.3467	**3.1803**

实验最后一个部分可以对你方案（方法）失败的情况进行讨论分析。有些同学可能会想：在论文里面写出自己方法失败的情况，不是在暴露自己的缺点从而被审稿人降分吗？实际上，情况并不是这样的。对于信息领域的应用基础研究或应用研究来说，在绝大多数情况下，你的方案或者算法几乎不可能是完美的，当放到实际环境中去应用测试的时候，必然会有一些失败的情况。能够给出典型的失败例子并讨论分析出导致失败的潜在原因，恰恰是论文作者严谨治学态度的最好例证。因此，**如果论文中包含了对失败情况的讨论分析，审稿人会觉得你比较严谨、对问题本质有着深刻的认识和理解，会对你的论文产生一个良好的印象**。下面通过 2 个范例来让我们看一下如何对失败案例进行讨论分析。

范例 1：

Failure Cases of DeepPS. DeepPS currently is not perfect. When the imaging conditions are poor (e.g., there are strong shadows caused by nearby trees), sometimes it will miss a true candidate (false negative). There are two main reasons for missing true candidates. First, the confidence score of a true marking-point is lower than the predefined threshold. Second, the local image pattern classification model may misclassify a valid entrance-line candidate as “invalid.”

Compared with false negatives, in a practical self-parking system, false positives actually are more annoying. Occasionally, DeepPS may return a false positive. The root reason is that the DCNN model may (although quite rare) misclassify with a high confidence score the local image pattern defined by two marking-points. Four examples of false positives detected by DeepPS are shown in Fig. 12. In Fig. 12(a), the “right-angled clockwise” entrance-line is misclassified as “right-angled anticlockwise.” In Fig. 12(b), the “slanted”entrance-line is misclassified as “right-angled.” In Fig. 12(c), the “right-angled anticlockwise” local image pattern defined by $\mathbf{p}_1$ and $\mathbf{p}_2$ is misclassified as “right-angled clockwise.”In Fig. 12(d), the “invalid” local image pattern is

misclassified as "right-angled clockwise." Actually, misclassifying images with high confidences is a known shortcoming of DCNN [69], and in the future we may adopt strategies (e.g., DeepFool [70]) for creating adversarial training samples to solve this issue.

范例 2：

Although D2S1R has achieved remarkable results as a whole, there are still some local failure samples. Several typical failure cases are shown in Fig. 4. By analyzing the samples failed to be located, several typical situations are summarized here:

[1] Narrow long corridor (Fig. 4(a) and Fig. 4(b)). In such a scene, all locations along the corridor have high matching degrees, which is a typical repetitive pattern problem in localization. Due to the limitation of the measurement range, this problem cannot be completely solved by a single short-range LiDAR, and it is necessary to incorporate other sensors or resort to long-range LiDARs.

[2] Candidate location deviation (Fig. 4(c)). One of the possible reasons for this situation is that the partition resolution of the distance histogram is too large, resulting in low degree of discrimination of the location against adjacent locations.

[3] Office area repetition pattern (Fig. 4(d)). It is another typical repetition pattern scenario, i.e. a similar office area appears, resulting in the best candidate location not being within the retrieval range. This can be solved to some extent by increasing the number of candidate locations.

[4] Priori map error (Fig. 4(e)). The localization result of this situation appears as the local deviation of matching degree, which is caused by the lack of local scanning data. To avoid such errors, when constructing a priori map, someone should expand the whole space as much as possible and add enough loops to improve the accuracy of map construction.

F. 结论

结论部分比较容易组织，基本上就是阐明本篇论文主要解决了什么问题、你的贡献在哪里、最终取得的结果和成效是什么，这部分的表述以一般过去时为主。最后还可以提一下将来的工作想法和计划（future work）。

III. 表述的逻辑性

对于初写论文的学生来说，写出来的文字缺乏逻辑性是一个最常见的也是最严重的问题。实际上，书面文字的逻辑性是独立于语言的，它是一个人逻辑思维能力与语言呈现能力的体现。即使用中文来表达，我们很多研究生写出来的材料也是极度缺乏逻辑性，语病一大堆。要想提升文字表达的逻辑性，必须在平时就要加以有意识的训练。

好的论文读上去应该是层次分明、逻辑紧密的。如果你的论文读上去总感觉不太通顺、不太连贯、似是而非的，那么一定是在逻辑上出了问题。**解决逻辑问题的关键在于理顺思维，要把握住整篇论文章节之间、段落之间、意群之间、句子之间、句子内部成分之间的逻辑关系。这个过程往往是一个反复迭代、不断推敲的过程，很难一蹴而就。**“两句三年得，一吟双泪流”，在逻辑推敲这个问题上，我们似乎也值得如此。

论文中的逻辑问题无外乎就是两大类，要么就是缺乏逻辑，要么就是逻辑错误。下面，我将分别针对这两方面的问题，来说说我们应该如何应对。

A. 隐秘的钩子：保证逻辑紧密的利器

优秀的成品论文在逻辑上来说应该是浑然一体的。也就是说，论文中的任何一句话都是和论文整体有关系的，不能出现某一句话、某些段落是孤立的、游离在整体之外的。**要做到这一点，论文在各个层级上都要有“粘合剂”，更形象的说，要有“钩子”：每一句话都要有一个隐形“挂钩”挂到其他部分上，从而使得论文前后内容之间都是有关联关系的。**

- Section（sub-section）之间的联系

Section 之间的逻辑关系一般是比较粗线条的，比较松散的。作为总览，在引言这一节的最后往往要介绍一下剩下几个 Section 的主要内容。这样，引言节就和论文其他节“钩”在一起了。章节内容安排的写法可以参考如下范例。

范例：

The remainder of this paper is organized as follows. Section II introduces the related work and our contributions. Section III presents our DCNN-based parking-slot detection approach, DeepPS. Experimental results are presented in Section IV. Finally, Section V concludes the paper.

除了引言之外，其他节之间大部分情况下是并列关系，不需要格外注意他们之间的逻辑关联性。但也有时候（尤其是在介绍某种方案流程的时候），中间的几节之间（或者一个 Section 内部的几个 sub-section 之间）是递进关系，这时候需要注意保持它们之间的关联性。看一下这个范例。我们要描述泊车位检测系统，Section III-A、III-B、III-C 分别按照先后顺序，介绍了它的三个模块。Section III-D 需要总体总结检测系统的流程，可如下表述：

In Sect. III-A ~ III-C, we have presented details about our DCNN-based parking-slot detection approach DeepPS. In order to enable the reader to have a clear and overall understanding of our work, the pipeline of DeepPS is summarized in Table II.

- 段落之间的联系

相对于 Section 之间的关系来说，一个 Section 内部的段落之间的关联关系是更加密切的。这个时候务必要注意段与段之间的起承转合连接。也就是说，一般情况下，下面一段必须要和上面一段有连接关系。这种连接关系要正确地反映出段落之间的并列、转折、递进、因果等逻辑关系。

- 句子之间的联系

意群内部句子之间的逻辑关系是最密切的，务必要通过“钩子”把意群内部的句子相互勾连起来。如果在这个层面上出现逻辑的割裂，是最容易让读者感受到的，所以绝对不能在意群内部出现孤立在外、没有钩子能钩到它的语句。

句子之间的“钩子”可以通过连词、代词、表示逻辑关系的状语来体现。常见的可以作为“钩子”的词语和短语

包括：代词（it, that, this, those, these, ones, them, their, its），such a (an), the former ones, the latter ones, in addition to, besides, another, moreover, meanwhile, furthermore, however, by contrast, on the contrary, but, nonetheless, consequently, to this end, accordingly, thus, hence, as a result, therefore, due to, owing to, thanks to, on one hand…on the other hand, for example, as we all know, unfortunately, one…the other to sum up, as follows 等等。

通过下面这两个范例，来仔细体会一下句子之间的隐形“钩子”。

范例 1：

The working principle of a vision-based approach is fundamentally different from that of a method based on free-space. The goal of a **vision-based approach** is to identify and locate the parking-slots defined by parking line segments painted on the ground. Apparently, the performance of **such approaches** does not depend on the existence or poses of adjacent vehicles. **Moreover**, in most cases, parking line segments can provide more accurate parking information than “free-space.” **Meanwhile**, most car manufacturers have started to produce vehicles equipped with wide FOV (field of view) imaging sensors, usually used in an AVM (around view monitoring) system. For **these reasons**, the vision-based approach has begun to draw a lot of attention recently in the field of parking spot detection, which is also our focus in this paper.

范例 2：

VI-SLAM systems can be roughly categorized as loosely-coupled and tightly-coupled ones according to their ways of sensor fusion. Typical studies belonging to the **former ones** include [12] and [23]. **In [12]**, Munguía *et al*. fused measurements from different sensors and the system’s optimization was performed based on Kalman filtering in a loosely-coupled manner. **In [23]**, Weiss and Siegwart estimated the full and metric scaled state of a camera-IMU device in real time by decoupling of the visual pose estimate and the filter state estimation. Since **both the systems** in [12, 23] separately estimate the motion of the IMU and the camera, **they** fail to obtain highly consistent localization results due to the lack of complementary information.

By contrast, the **latter ones** are currently popular schemes adopted in the field of autonomous driving, which combine the state of the IMU and the camera to perform state estimation to ensure a consistent localization result [9–11, 15, 17, 18, 20, 29]. Eminent studies along **this technical line** are briefly reviewed here. MSCKF (multistate constraint Kalman filter) is a real-time visual-inertial navigation system based on EKF (extended Kalman filter) [11]. **It** can provide accurate poses in large-scale environments and the time complexity of the MSCKF algorithm is only related to the number of features. **However**, the back-end of **MSCKF** is based on Kalman filter in which global information cannot be explored for optimization. OKVIS (open keyframe-based visual-inertial SLAM) [9] predicts the current state based on IMU measurement, and performs feature extraction and feature matching based on prediction. **But it** does not support relocation, and there is no loop-closing. In [15], Mur-Artal and Tardos incorporated IMU measurement into the ORB-SLAM system [7, 13, 14, 16], which is widely used in the field of autonomous driving. VINS (a monocular visual-inertial system) [17, 18] is a robust and versatile monocular visual-inertial state estimator. **Its** front-end resorts to KLT (Kanade-Lucas-Tomasi) tracker [1] to track Harris corner points [8], and **its** back-end makes use of sliding windows for optimization. The pre-integration of the IMU and the vision-IMU alignment ensure **VINS’** robustness and stability. S-MSCKF (stereo multi-state constraint Kalman filter) [20] is **a binocular version of MSCKF**, resorting to Fast corner [19] and KLT tracker [1] for tracking. PIRVS (PerceptIn robotics vision system) [29] tightly couples vision sensors and IMUs while loosely coupling with other sensors. It needs to be noted that maps constructed by **these VI-SLAM systems** only provide geometric information, lacking of a semantic understanding of the environment.

下面，我再举几个典型的逻辑缺失的例子，这都来自我们研究生论文的初稿。

范例 1：

S. Hold *et al*. proposed a method of online extrinsic calibration for the binocular camera in [4]. They adopt a conventional detector to detect the lane and sample a series of feature points with a scanning line. By Fast Fourier transform (FFT) they measure the distance between lane points and finally, they can solve the camera’s extrinsic. Ground lanes are considered as parallel lines, which is a strong hypothesis and limits the scope of application.（最后一句没有钩子，游离在外了！）

范例 2：

Knorr *et al*.[8] construct a recursive optimization algorithm. Relative camera poses are corrected by Extended Kalman Filter and the relationship between the multi-camera system and the ground is re-calculated by homography estimation. The solution proposed in [9] relies on two parallel lanes on the flat ground. The relative pose between the camera system coordinate and the world coordinate can be obtained in vanishing point calculation. It’s worth mentioning that the authors consider the camera system as a whole so relative camera poses are not optimized.（全部句子都没有钩子，都是孤立的。）

范例 3：

VINS [20, 21] is a robust and versatile monocular visual inertial state estimator. The front end uses KTL optical flow method [1] to track Harris corner points [8], and the back end uses sliding windows for optimization. The pre-integration of the IMU and the vision-IMU alignment ensure the robustness and stability of the system.（全部句子都没有钩子，都是孤立的。）

范例 4：

Bosse and Zlot designed three key point extraction algorithms, and proposed five key point description models. The advantages and disadvantages of different models and the design principles of 2D key points were discussed. It has certain

guidance for the extraction of 2D key points and the design of descriptors [20]. （全部句子都没有钩子，都是孤立的；it 指代不清。）

B. 避免似是而非的逻辑关系

逻辑关系不成立是论文逻辑问题的另一个重要方面。产生这个问题的原因也是由于你的思路比较混乱，或者说语言表达功力不够。这个问题的具体表现形式在于：句子之间或者一个句子各个内部成分之间英文字面所体现出来的逻辑关系实际上是不成立的。带有这些毛病的段落，读起来会让人感觉似是而非的。用我的话说就是无逻辑！

所以，论文写完之后，要仔细体会所表达的逻辑关系（递进、对照、转折、因果等等）是否是存在的。对于一个长句来说，需要检查：1）动作是主语发出的吗？；2）动作的宾语对吗？3）现在分词做状语的时候，分词的隐含主语与句子的主语一致吗？

下面，我举几个典型的逻辑错误的范例，请大家体会。

范例 1：

Monocular and binocular visual cameras constitute the basic configuration to build such a semantic SLAM system. However, due to its inability to perceive ground information around the vehicle, a surround-view camera system capturing images from a top-down viewpoint, namely surround-view bird's-eye image, is necessarily called for.（due to 用法逻辑错误。）

范例 2：

Since semantic features in the underground garage typically exist on the ground, semantic SLAM system fails to track consistently during navigation. However, a surround-view system typically consists of four cameras mounted around the vehicle capturing images from a top-down viewpoint.（however 用法错误，这两句没有转折关系。）

IV. 进阶：语言的润色

在表达准确、逻辑紧密的前提下，可以考虑对论文的语言文字进行润色，使之读上去更加的优美，让人赏心悦目。不过这个过程往往比较耗时，也要靠平时有意识的积累。

A. 同意词语的替换

对语言驾驭能力的一个重要体现就是看你会不会用一些更加高级的词语，而不是总停留在中小学词汇量的水平上。对一些常用的语意，要注意平时多积累一些不同形式的表达，以便在同一个论文中不同位置上用不同语言形式来表达同一个语意，增加语言的多样性与丰富性。下面列举一些常用语意的多种表达：

- 使用，利用

use, make use of, make fully use of, explore, exploit, resort to, adopt

- 然而，但是

however, nonetheless, whereas, unfortunately

- 由于，鉴于

because, because of, since, for, as, owing to, due to, thanks to, as a result of

- 除此之外，更进一步

in addition, furthermore, moreover

- 因此

thus, hence, therefore, consequently, accordingly, as a result

- 优点，优势

advantage, merit, superiority

- 表明，证明，暗示着

show, demonstrate, indicate, imply, corroborate, exhibit

- 建立，构建

build, construct, establish

- 方法，方案

method, approach, scheme, solution

- 介绍，描述，讨论

present, introduce, state, discuss

- 解决，处理

solve, resolve, conquer, deal with, cope with

- 对手，同类型的对照方法

competitor, rival, counterpart

- 变为，成为，结果是

become, turn out

- 目前，当前

at present, currently, thus far

- 通过，利用，经由

via, by, by means of

- 至于…，关于…方面

in terms of, with respect to

B. 句式的多样性

为了增加论文的可欣赏性，我们要有意识的增加句式的多样性，不能都是简单的短陈述句。要灵活运用多种句子结构，比如说：状语从句、定语从句、主语从句、同位语从句、现在分词结构做状语、with 和 without 结构等等。

V. 结论

本文根据作者自己的认知和经验，总结了英文科技文写作的要点、注意事项、常见问题，也提出了一些系统性的论文撰写工作方法，希望能给初学者们一些启示。《史记 •吕不韦列传》：“布咸阳市门，悬千金其上，延诸侯游士宾客有能增损一字者予千金。”。类比一下，如果你的科技文写作水平能达到《吕氏春秋》的程度，你就成功了！