

# 西安交通大学

## 毕业设计（论文）

题 目 \_\_\_\_\_

黑体，黑色，居中

\_\_\_\_\_ 学院 \_\_\_\_\_ 专业 \_\_\_\_\_ 班

学生姓名 \_\_\_\_\_ 张三 \_\_\_\_\_

学 号 \_\_\_\_\_ 2170120000 \_\_\_\_\_

指导教师 \_\_\_\_\_

设计所在单位 \_\_\_\_\_

2021 年 6 月



西安交通大学

系（专业） \_\_\_\_\_  
系（专业）主任 \_\_\_\_\_  
批准日期 \_\_\_\_\_

## 毕业设计（论文）任务书

\_\_\_\_\_学院\_\_\_\_\_专业\_\_\_\_\_班 学生\_\_\_\_\_

毕业设计(论文)课题 \_\_\_\_\_

毕业设计(论文)工作自\_\_\_\_\_年\_\_\_\_\_月\_\_\_\_\_日起至\_\_\_\_\_年\_\_\_\_\_月\_\_\_\_\_日止

毕业设计(论文)进行地点： \_\_\_\_\_

课题的背景、意义及培养目标

---

---

---

---

---

---

---

---

设计(论文)的原始数据与资料

---

---

---

---

---

---

---

---

课题的主要任务

---

---

---

---

---

---

---

---

---

---

---

---

课题的基本要求(工程设计类题应有技术经济分析要求)

---

---

---

---

---

---

---

---

完成任务后提交的书面材料要求(图纸规格、数量，论文字数，外文翻译字数等)

---

---

---

---

---

---

---

---

主要参考文献

指导教师\_\_\_\_\_

接受设计(论文)任务日期\_\_\_\_\_

(注：由指导教师填写)

学生签名：\_\_\_\_\_

西 安 交 通 大 学

毕 业 设 计 ( 论 文 ) 考 核 评 议 书

\_\_\_\_\_学院\_\_\_\_\_专业\_\_\_\_\_班

指导教师对学生\_\_\_\_\_所完成的课题为\_\_\_\_\_

的毕业设计(论文)进行的情况, 完成的质量及评分的意见: \_\_\_\_\_

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

指导教师建议成绩: \_\_\_\_\_

指导教师\_\_\_\_\_

年 月 日

毕 业 设 计 ( 论 文 ) 评 审 意 见 书

评审意见: \_\_\_\_\_

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

评阅人建议成绩: \_\_\_\_\_

评阅人\_\_\_\_\_职称\_\_\_\_\_

年 月 日

## 毕业设计(论文)答辩结果

\_\_\_\_\_学院

\_\_\_\_\_专业

毕业设计(论文)答辩组对学生\_\_\_\_\_所完成的课题为

\_\_\_\_\_

的毕业设计(论文)经过答辩，其意见为\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

并确定成绩为\_\_\_\_\_

毕业设计(论文)答辩组负责人\_\_\_\_\_

答辩组成员\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

年 月 日

## 摘 要

本文基于 Overleaf 上开源的西安交通大学博士学位毕业论文 L<sup>A</sup>T<sub>E</sub>X 模板<sup>①</sup>，修改得到的西安交通大学学士学位毕业设计（论文）模板，本着造福校友的原则，将本项目开源。本文的编译环境为 XeLaTeX，TeX Live Version 为 2019。

摘要是论文的高度概括，是全文的缩影，是长篇小说不可缺少的组成部分。要求用中、英文分别书写，一篇摘要不少于 400 字。英文摘要与中文摘要的内容和格式必须一致。

**关键词：**西安交通大学学士学位论文；毕业设计；L<sup>A</sup>T<sub>E</sub>X

---

<sup>①</sup> 链接为：<https://www.overleaf.com/latex/templates/latex-template-for-doctoral-thesis-of-xjtu/bmrqcdhbdrcw>，作者为张明

## ABSTRACT

You will never want to use Word again after you know how to write in  $\LaTeX$ .

**KEY WORDS:** XJTU B.S. thesis;  $\LaTeX$



## 目 录

主要符号表.....	IX
1 图表公式.....	1
1.1 图.....	1
1.2 表.....	1
1.3 公式.....	2
1.3.1 单个公式.....	2
1.3.2 多个公式.....	2
2 论文结构.....	3
2.1 摘要.....	3
2.2 主要符号表.....	3
2.3 正文.....	3
2.3.1 绪论.....	3
2.3.2 课题的研究方法与手段.....	3
2.3.3 结论与展望.....	4
2.3.4 致谢.....	4
2.4 参考文献与注释.....	4
2.5 装订要求.....	5
3 格式要求.....	6
3.1 封面.....	6
3.2 任务书、考核评议书.....	6
3.3 中、英文摘要页.....	6
3.4 中、英文目录页.....	6
3.5 表格.....	6
3.6 图片.....	7
3.7 其他.....	8
3.7.1 页码.....	8
致 谢.....	9
参考文献.....	10
附录 A 公式定理证明.....	11

附录 B 算法与代码.....	13
B.1 算法 .....	13
B.2 代码 .....	13
附录 C 外文文献原文.....	15
附录 D 外文文献译文 .....	17

## 主要符号表

$x$	请输入你对符号的解释与说明文字
$y$	请输入你对符号的解释与说明文字



## 1 图表公式

论文写作过程中最重要的是**图、表、公式**等内容的编排<sup>①</sup>。

### 1.1 图

图 1-1 是用 Tikz 画的，Visio 画不出这么好看的图。

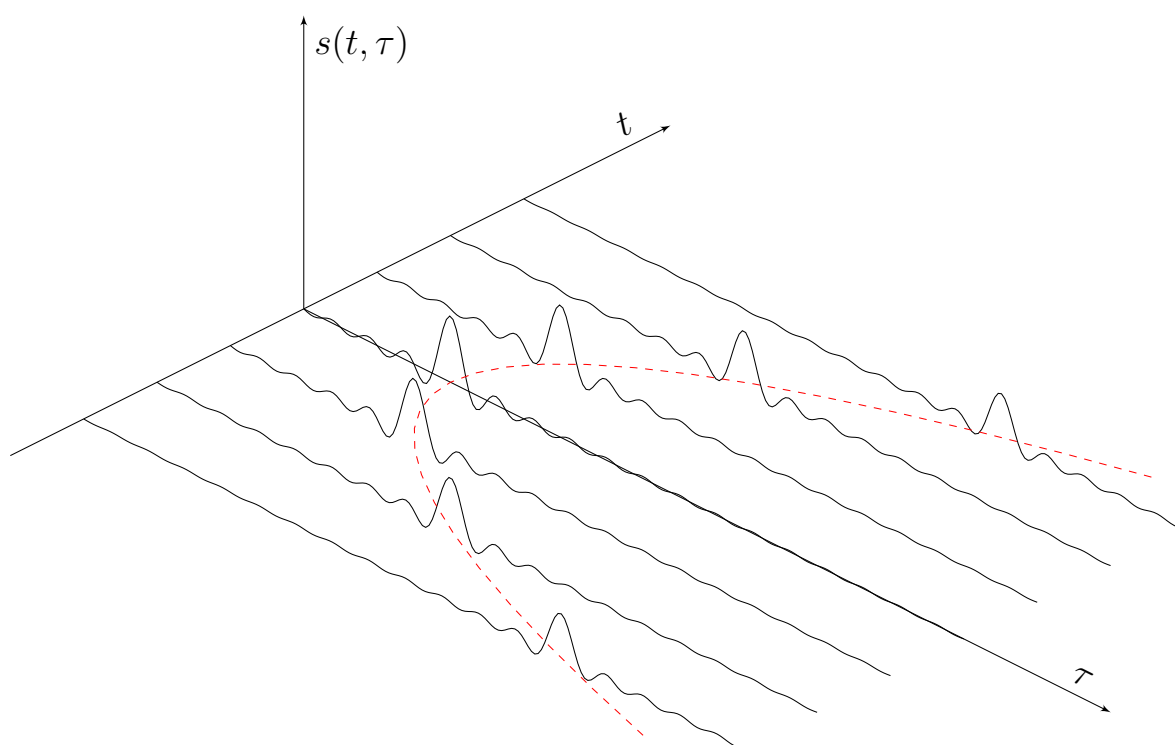


图 1-1 图注是五号字。

### 1.2 表

表格要求采用三线表，如表 1-1 所示。

表 1-1 表题也是五号字

Interference	DOA (deg)	Bandwidth (MHz)	INR (dB)
1	-30	20	60
2	20	10	50
3	40	5	40

<sup>①</sup> 这是一个注脚，可以输入你对内容的补充说明

## 1.3 公式

### 1.3.1 单个公式

单个公式的编号如式 (1-1) 所示，该式是标准正态分布的概率密度函数<sup>[1]</sup>，从公式可知 L<sup>A</sup>T<sub>E</sub>X 排版的公式比 MathType 美观，而且公式编写效率更高。

$$f_Z(z) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|z - \mu_Z|^2}{\sigma^2}\right) \quad (1-1)$$

### 1.3.2 多个公式

多个公式作为一个整体可以进行二级编号，如 (1-2) 所示，该式是连续时间 Fourier 变换的正反变换公式<sup>[2]</sup>。

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (1-2a)$$

$$x(t) = \int_{-\infty}^{\infty} X(f)e^{j2\pi ft} df \quad (1-2b)$$

## 2 论文结构

### 2.1 摘要

摘要是论文的高度概括，是全文的缩影，是长篇小说不可缺少的组成部分。要求用中、英文分别书写，一篇摘要不少于 400 字。英文摘要与中文摘要的内容和格式必须一致。

### 2.2 主要符号表

如果论文中使用了大量的物理量符号、标志、缩略词、专门计量单位、自定义名词和术语等，应将全文中常用的这些符号及意义列出<sup>①</sup>。**如果上述符号和缩略词使用次数不多，可以不设专门的主要符号表，但在论文中出现时须加以说明。**论文中主要符号全部采用法定单位，严格执行国家标准（GB3100~3102—93）有关“量和单位”的规定。单位名称采用国际通用符号或中文名称，但全文应统一，不得两种混用。缩略词应列出中英文全称。

### 2.3 正文

#### 2.3.1 绪论

绪论相当于论文的开头，它是三段式论文的第一段（后二段是本论和结论）。绪论与摘要写法不完全相同，摘要要写得高度概括、简略，绪论可以稍加具体一些，文字以 1000 字左右为宜。绪论一般应包括以下几个内容：

- (1) 为什么要写这篇论文，要解决什么问题，主要观点是什么。
- (2) 对本论文研究主题范围内已有文献的评述（包括与课题相关的历史的回顾，资料来源、性质及运用情况等）。
- (3) 说明本论文所要解决的问题，所采用的研究手段、方式、方法。明确研究工作的界限和规模。概括论文的主要工作内容。

#### 2.3.2 课题的研究方法与手段

课题研究的方法与手段，分别以下面几种方法说明：用实验方法研究课题，应具体说明实验用的装置、仪器、原材料的性能等是否标准，并应对所有装置、仪器、原材料做出检验和标定。对实验的过程和操作方法，力求叙述得简明扼要，对实验结果的记录、分析，对人所共知的或细节性的内容不必过分详述。用理论推导的手段和方法达到研究目的的，这方面内容要精心组织，做到概念准确，判断推理符合客观事物的发展规

<sup>①</sup> [https://es.overleaf.com/learn/latex/List\\_of\\_Greek\\_letters\\_and\\_math\\_symbols](https://es.overleaf.com/learn/latex/List_of_Greek_letters_and_math_symbols)

律，要做到言之有序，言之有理，以论点为中心，组织成完整而严谨的内容整体。用调查研究的方法达到研究目的的，调查目标、对象、范围、时间、地点、调查的过程和方法等，这些内容与研究的最终结果有关系，但不是结果本身，所以一定要简述。但对调查所提的样本、数据、新的发现等则应详细说明，这是结论产生的依据。

### 2.3.3 结论与展望

在写作时，应对研究成果精心筛选，把那些必要而充分的数据、现象、样品、认识等选出来，写进去，作为分析的依据，应尽量避免事无巨细，把所得结果和盘托出。在对结果做定性和定量分析时，应说明数据的处理方法以及误差分析，说明现象出现的条件及其可证性，交代理论推导中认识的由来和发展，以便别人以此为根据进行核实验证。对结果进行分析后所得到的结论和推论，也应说明其适用的条件和范围。恰当运用表和图作结果与分析，是科技论文通用的一种表达方式。

结论与展望：结论包括对整个研究工作进行归纳和综合而得出的总结；所得结果与已有结果的比较；联系实际结果，指出它的学术意义或应用价值和在实际中推广应用的可能性；在本课题研究中尚存在的问题，对进一步开展研究的见解与建议。结论集中反映作者的研究成果，表达作者对所研究课题的见解和主张，是全文的思想精髓，是全文的思想体现，一般应写得概括、篇幅较短。

### 2.3.4 致谢

对于毕业设计（论文）的指导教师，对毕业设计（论文）提过有益的建议或给予过帮助的同学、同事与集体，都应在论文的结尾部分书面致谢，言辞应恳切、实事求是。

## 2.4 参考文献与注释

参考文献是为撰写论文而引用的有关文献的信息资源。参考文献列示的内容务必实事求是。论文中引用过的文献必须著录，未引用的文献不得虚列。遵循学术道德规范，杜绝抄袭、剽窃等学术不端行为。参考文献须是作者亲自考察过的对学位论文有参考价值的文献。参考文献应有权威性，应注意所引文献的时效性。参考文献的数量：一般不少于 10 篇，其中，期刊文献不少于 8 篇，国外文献不少于 2 篇，均以近 5 年的文献为主。注释是正文需要的解释性、说明性、补充性的材料、意见和观点等。

参考文献格式应符合国家标准 GB/T-7714-2005《文后参考文献著录规则》。中国国家标准化管理委员会于 2015 年 5 月 15 日发布了新的标准 GB/T 7714-2015《信息与文献参考文献著录规则》。因为二者的差别非常小，所以采用了新的标准。标准的 BiBTeX 格式网上资源非常多，本文使用了李泽平开发的版本<sup>[3]</sup>。本模板中使用 bib<sub>La</sub>TeX 提供参考文献支持。请使用 \upcite 命令进行引用。

如，本文参考了如下文献<sup>[4] [5]</sup>。



## 2.5 装订要求

论文的查重部分和不查重部分的内容如表2-1所示。

表 2-1 论文查重部分与不查重部分内容对照表

查重部分	不查重部分
封面	致谢
中文摘要	外文原文（附录）
英文摘要	外文译文（附录）
目录	有关算法（附录）
正文	有关图纸（附录）
参考文献	计算机源程序（附录）

### 毕业设计(论文)装订次序要求

- 第一 封面
- 第二 任务书（双面打印）
- 第三 考核评议书（背面是答辩结果）
- 第四 中文摘要
- 第五 英文摘要
- 第六 目录
- 第七 正文（含绪论和结论）
- 第八 致谢
- 第九 参考文献
- 第十 附录（含外文原文及其译文、有关图纸、计算机源程序等）

## 3 格式要求

### 3.1 封面

采用西安交通大学毕业设计（论文）统一封面（模板见毕业设计管理系统——资料下载），用 A3 纸（封面封底连在一起，从左侧包住论文）。封面上所填内容居中排列三字号加粗，论文题目不能超过 35 个汉字。封面可去教材科购买。

### 3.2 任务书、考核评议书

双面打印。《考核评议书》、《评审意见书》和《答辩结果》必须分别由指导教师、评阅人和答辩组据实填写。

### 3.3 中、英文摘要页

摘要页由摘要正文、关键词等组成。

摘要 / ABSTRACT 按一级标题编排中文“摘要”二字，二字间距为两个字符。英文为“ABSTRACT”。摘要正文，中文每段开头左起空两字符起排，段与段之间不空行；英文每段开头左对齐顶格编排，段与段之间空一行。小四号字。

关键词/KEY WORDS”正文内容下，空一行，左对齐顶格编排“关键词/KEY WORDS”（小四号，加粗），后接冒号，其后为具体关键词（小四号），每一关键词之间用分号分开，最后一个关键词后无标点符号。英文每组 KEY WORDS 的第一个字母为大写，其余为小写。

资助申明：资助申明编排在摘要第一页的页脚处。

### 3.4 中、英文目录页

中文目录页应放在奇数页上起排。“目录”二字按一级标题编排，两字间距两个字符。目录正文，包括编号、标题及其开始页码。一般只列到三级标题。目录中标题的编号应与正文中标题的编号一致；第一级标题左对齐顶格编排；与上一级标题相比，下一级标题左缩进一个字符起排；标题与页码之间用“……”连接。页码右对齐顶格编排；建议采用文本编辑软件的“目录自动生成功能”生成目录。如果有英文目录，英文目录的内容、格式均须与中文目录一致。

### 3.5 表格

图、表、公式等的序号用阿拉伯数字分章连续编号，如图 1-1、表 1-1、表 2-1 等，但不出现“公式”两字，将编号置入小括号中，如（1-2）等。图、表和公式等与正文

之间间隔 0.5 行。图应有图题，表应有表题，并分别置于图号和表号之后。图号和图题置于图下方的居中位置，表号和表题应置于表上方的居中位置。引用图或表应在图题或表题右上角标出文献来源。若图或表中有附注，采用英文小写字母顺序编号，附注写在图或表的下方。物理量及量纲均按国际标准（SI）及国家规定的法定符号和法定计量单位标注，禁止使用已废弃的符号和计量单位。物理量的符号由斜体字母标注，单位的符号使用正体字母标注，量与单位间用斜线隔开。例如： $I/A$ ， $\rho/kg \cdot m^{-3}$ ， $F/N$ ， $v/m \cdot s^{-1}$  等等。

表格应紧跟文述编排。表格中一般是内容和测试项目由左至右横读，数据依序竖读，应有自明性。表的各栏均应标明“量或测试项目、符号、单位”。只有在无必要标注的情况下方可省略。表内同一栏的数字必须上下对齐。表内不宜用“同上”、“同左”和类似词，一律填入具体数字或文字。表内“空白”代表未测或无此项，“…”代表未发现，“0”代表实测结果确为零。如数据已绘成曲线图，可不再列表。

1) 表格转页接排时，在随后的各页上应重复表的编号。编号后跟（续），如表 1（续），续表均应重复表头和关于单位的陈述。

2) 一律使用三线表，与文字齐宽，上下边线，线粗 1.5 磅，表内线，线粗 1 磅。在三线表中可以加辅助线，以适应较复杂表格的需要。

3) 使用他人表格须注明出处。

4) 表中用字一般为五号字。如排列过密，用五号字有困难时，可小于五号字，但不小于七号

### 3.6 图片

1) 一幅图如有若干幅分图，均应编分图号，用 (a),(b),(c),…按顺序编排；

2) 插图须紧跟文述。在正文中，一般应先见图号及图的内容后见图，特殊情况须延后的插图不应跨节；

3) 提供照片应大小适宜，主题明确，层次清楚，利于复制，金相照片一定要有比例尺；

4) 图应具有“自明性”，即只看图、图题和图例，不阅读正文，就可理解图意。

5) 图中的标目是说明坐标轴物理意义的项目，由物理量的符号或名称和相应的单位组成。

6) 图中用字一般为五号字，如排列过密，用五号字有困难时，可小于五号字，但不得小于七号字。

7) 图的大小一般为宽 6.67cm × 高 5.00cm。特殊情况下，也可宽 9.00cm × 高 6.75cm，或宽 13.5cm × 高 9.00cm。同类图片的大小应一致。图片的编排应美观、整齐。

## 3.7 其他

### 3.7.1 页码

论文页码的第一页**从正文开始用阿拉伯数字**标注，直至全文结束。**正文前的内容**（除封面）用**罗马数字**单独标注页码。页码位于页面底端，对齐方式为“外侧”，页码格式为最简单的数字，不带任何其它的符号或信息。页码不能出现缺页和重复页。附录（含外文原文及其译文、有关图纸、计算机源程序等）必须与论文装订在一起，附录的页码必须接着参考文献的页码连续编写。

## 致 谢

感谢那些你想感谢的对论文有帮助的人。

## 参考文献

- [1] Dimitris G M, Vinay K I, Stephen M K. Statistical and adaptive signal processing[M]. Norwood: Artech House, Inc., 2005.
- [2] Vetterli M, Kovacevic J, K Goyal V. Foundations of signal processing[M]. Cambridge: Cambridge University Press, 2014.
- [3] Lee Z. GB/T 7714-2015 参考文献 BiBTeX 样式[EB/OL]. 2016. [github.com/ustctug/ustcthesis](https://github.com/ustctug/ustcthesis).
- [4] Zhou X, Wang Z, Ye H, et al. Ego-planner: An esdf-free gradient-based local planner for quadrotors [J/OL]. IEEE Robotics and Automation Letters, 2021, 6(2):478-485. DOI: 10.1109/LRA.2020.3047728.
- [5] Thrun S. Probabilistic robotics[J/OL]. Communications of the ACM, 2002, 45(3):52-57. DOI: 10.1145/504729.504754.

## 附录 A 公式定理证明

附录编号依次编为附录 A, 附录 B。附录中的图、表、公式另行编排序号, 编号前加“附录 A-”字样。

排版数学定理等环境时最好给环境添加结束符, 以明确定理等内容的起止标志, 方便阅读。例如定义的结束符采用  $\diamond$ , 例子的结束符采用  $\blacklozenge$ , 定理的结束符采用  $\square$ , 证明的结束符采用  $\blacksquare$ 。

**定义 A.1 (向量空间):** 设  $X$  是一个非空集合,  $\mathbb{F}$  是一个数域 (实数域  $\mathbb{R}$  或者复数域  $\mathbb{C}$ )。如果在  $X$  上定义了加法和数乘两种运算, 并且满足以下 8 条性质:

1. 加法交换律,  $\forall x, y \in X, x + y = y + x \in X$ ;
2. 加法结合律,  $\forall x, y, z \in X, (x + y) + z = x + (y + z)$ ;
3. 加法的零元,  $\exists 0 \in X$ , 使得  $\forall x \in X, 0 + x = x$ ;
4. 加法的负元,  $\forall x \in X, \exists -x \in X$ , 使得  $x + (-x) = x - x = 0$ 。
5. 数乘结合律,  $\forall \alpha, \beta \in \mathbb{F}, \forall x \in X, (\alpha\beta)x = \alpha(\beta x) \in X$ ;
6. 数乘分配律,  $\forall \alpha \in \mathbb{F}, \forall x, y \in X, \alpha(x + y) = \alpha x + \alpha y$ ;
7. 数乘分配律,  $\forall \alpha, \beta \in \mathbb{F}, \forall x \in X, (\alpha + \beta)x = \alpha x + \beta x$ ;
8. 数乘的幺元,  $\exists 1 \in \mathbb{F}$ , 使得  $\forall x \in X, 1x = x$ ,

那么称  $X$  是数域  $\mathbb{F}$  上的一个**向量空间** (linear space)。  $\diamond$

**例 A.1 (矩阵空间):** 所有  $m \times n$  的矩阵在普通矩阵加法和矩阵数乘运算下构成一个向量空间  $\mathbb{C}^{m \times n}$ 。如果定义内积如下:

$$\langle A, B \rangle = \text{tr}(B^H Q A) = \sum_{i=1}^n b_i^H Q a_i \quad (\text{附录 A-1})$$

其中  $a_i$  和  $b_i$  分别是  $A$  和  $B$  的第  $i$  列, 而  $Q$  是 HPD 矩阵, 那么  $\mathbb{C}^{m \times n}$  构成一个 Hilbert 空间。当  $Q = I$  时

$$\langle A, B \rangle = \text{tr}(B^H A) \quad (\text{附录 A-2})$$

称为 Frobenius 内积, 对应的范数称为 Frobenius 范数, 即矩阵所有元素模平方之和再开方:

$$\|A\|_F = \sqrt{\text{tr}(A^H A)} = \sqrt{\sum_{j=1}^n \sum_{i=1}^m |a_{ij}|^2} \quad (\text{附录 A-3})$$

如果  $m = n$ , 那么所有  $m \times m$  的 Hermite 矩阵构成  $\mathbb{C}^{m \times m}$  的子空间。但是所有  $m \times m$  的 HPD 矩阵并不构成子空间, 因为 HPD 矩阵对线性运算不封闭。  $\blacklozenge$

**定理 A.1 (Riesz 表示定理):** 设  $H$  是 Hilbert 空间,  $H^*$  是  $H$  的对偶空间, 那么

对  $\forall f \in H^*$ , 存在唯一的  $x_f \in H$ , 使得

$$f(x) = \langle x, x_f \rangle, \quad \forall x \in H \quad (\text{附录 A-4})$$

并且满足  $\|f\| = \|x_f\|$ 。

□

**证明：** 先证存在性，再证唯一性，最后证  $\|f\| = \|x_f\|$ 。

■



## 附录 B 算法与代码

对于数学、计算机和电子信息专业，算法和代码也是经常用到的排版技巧。

### B.1 算法

算法描述使用 `algorithm2e` 宏包，效果如算法 B-1 所示。

---

```

Input:  $\mathbf{x}(k)$ ,  $\mu$ ,  $\mathbf{w}(0)$ 
Output:  $y(k)$ ,  $\varepsilon(k)$ 
1 for  $k = 0, 1, \dots$  do
2    $y(k) = \mathbf{w}^H(k)\mathbf{x}(k)$  // output signal
3    $\varepsilon(k) = d(k) - y(k)$  // error signal
4    $\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\varepsilon^*(k)\mathbf{x}(k)$  // weight vector update
5 end

```

---

算法 B-1 LMS 算法详细描述

### B.2 代码

源代码使用 `listings` 宏包，LMS 算法的 Verilog 模块端口声明如代码 B-1 所示。

代码 B-1 空时 LMS 算法 Verilog 模块端口声明

```

1  module stap_lms
2  #(
3  parameter    M          = 4, // number of antennas
4              L          = 5, // length of FIR filter
5              W_IN       = 18, // wordlength of input data
6              W_OUT      = 18, // wordlength of output data
7              W_COEF     = 20  // wordlength of weights
8  )(
9  output signed [W_OUT-1:0] y_i, // in-phase component of STAP output
10 output signed [W_OUT-1:0] y_q, // quadrature component of STAP output
11 output                                vout, // data valid flag of output (high)
12 input  [M*W_IN-1:0] u_i, // in-phase component of M antennas
13 input  [M*W_IN-1:0] u_q, // quadrature component of M antennas
14 input                                vin, // data valid flag for input (high)
15 input                                clk, // clock signal
16 input                                rst  // reset signal (high)
17 );

```



## 附录 C 外文文献原文

**Mapillary Street-Level Sequences: A Dataset for Lifelong Place Recognition**

Frederik Warburg<sup>†,\*</sup>, Søren Hauberg<sup>†</sup>, Manuel López-Antequera<sup>‡</sup>, Pau Gargallo<sup>‡</sup>,  
Yubin Kuang<sup>‡</sup>, and Javier Civera<sup>§</sup>

<sup>†</sup>Technical University of Denmark, <sup>‡</sup>Mapillary AB, <sup>§</sup>University of Zaragoza  
<sup>†</sup>{frwa,sohau}@dtu.dk, <sup>‡</sup>{manuel, pau, yubin}@mapillary.com, <sup>§</sup>jcivera@unizar.es

**Abstract**

*Lifelong place recognition is an essential and challenging task in computer vision, with vast applications in robust localization and efficient large-scale 3D reconstruction. Progress is currently hindered by a lack of large, diverse, publicly available datasets. We contribute with Mapillary Street-Level Sequences (MSLS), a large dataset for urban and suburban place recognition from image sequences. It contains more than 1.6 million images curated from the Mapillary collaborative mapping platform. The dataset is orders of magnitude larger than current data sources, and is designed to reflect the diversities of true lifelong learning. It features images from 30 major cities across six continents, hundreds of distinct cameras, and substantially different viewpoints and capture times, spanning all seasons over a nine-year period. All images are geo-located with GPS and compass, and feature high-level attributes such as road type.*

*We propose a set of benchmark tasks designed to push state-of-the-art performance and provide baseline studies. We show that current state-of-the-art methods still have a long way to go, and that the lack of diversity in existing datasets has prevented generalization to new environments. The dataset and benchmarks are available for academic research.<sup>1</sup>*

**1. Introduction**

Visual place recognition is essential for the long-term operation of Augmented Reality and robotic systems [31]. However, despite its relevance and vast research efforts, it remains challenging in practical settings due to the wide array of appearance variations in outdoor scenes, as seen in the examples extracted from our dataset in Figure 1.

Recent research on place recognition has shown that features learned by deep neural networks outperform traditional

<sup>\*</sup>The main part of this work was done while Frederik Warburg was an intern at Mapillary.

<sup>1</sup>[www.mapillary.com/datasets/places](http://www.mapillary.com/datasets/places)



Figure 1: Mapillary SLS contains imagery from 30 major cities around the world; **red** stands for training cities and **blue** for test cities. See four samples from San Francisco, Trondheim, Kampala and Tokyo with challenging appearance changes due to viewpoint, structural, seasonal, dynamic, and illumination.

hand-crafted features, particularly for drastic appearance changes [5, 31, 55]. This has motivated the release of several datasets for training, evaluating and comparing deep learning models. However, such datasets are limited, in at least three aspects. First, none of them covers the many appearance variations encountered in real-world applications. Second, many of them have insufficient size for training large networks. Finally, most datasets are collected in small areas, lacking the geographical diversity needed for generalization.

This paper contributes to the progress of lifelong place recognition by creating a dataset addressing all the challenges described above. We present **Mapillary Street-Level Sequences (MSLS)**, the largest dataset for place recognition to date, with the widest variety of perceptual changes and the broadest geographical spread<sup>2</sup>. MSLS covers the following causes of appearance change: different seasons, changing weather conditions, varying illumination at different times of the day, dynamic

<sup>2</sup>See the video accompanying the paper for an overview and sample images.

Name	Environment	Total length	Geographical coverage	Temporal coverage	Frames	Type of appearance changes						
						Seasonal	Weather	Viewpoint	Dynamic	Day/night	Intrinsics	Structural
Nordland [36, 37]	Natural + urban	728 km	182 km	1 year	~115K	✓	✗	✗	✗	✗	✗	✗
SPED [12]	Urban	-	-	1 year	~2.5M	✓	✓	✗	✓	✓	✗	✗
KITTI [20]	Urban + suburban	39.2 km	1.7 km	3 days	~13K	✗	✗	✓	✓	✗	✗	✗
Eynsham [14]	Urban + suburban	70 km	35 km	1 day	~10K	✗	✗	✗	✓	✗	✗	✗
St. Lucia [21]	Suburban	47.5 km	9.5 km	1 day	~33K	✗	✗	✗	✓	✗	✗	✗
NCLT [9]	Campus	148.5 km	5.5 km	15 mon.	~300K	✓	✗	✓	✓	✗	✗	✗
Oxford RobotCar [32]	Urban + suburban	1,000 km	10 km	1 year	~27K	✓	✓	✓	✓	✓	✗	✓
VL-CMU [8]	Urban + suburban	128 km	8 km	1 year	~1.4K	✗	✗	✓	✓	✗	✗	✗
FAS [34]	Urban + suburban	120 km	70 km	3 years	~43K	✓	✓	✓	✓	✗	✗	✓
Garden Point [41]	Urban + campus	<12 km	4 km	1 week	~600	✗	✗	✓	✗	✓	✗	✗
SYNTHIA [44]	Urban	6 km	1.5 km	-	~200K	✓	✓	✓	✓	✓	✗	✗
GSV [56]	Urban	-	-	-	~60K	✗	✗	✗	✗	✗	✗	✗
Pittsburgh 250k [51]	Urban	-	-	-	~254K	✗	✗	✗	✓	✗	✗	✗
TokyoTM/247 [50]	Urban	-	-	-	~174K	✓	✗	✓	✓	✓	✗	✓
TB-places [28]	Gardens	<100m	<100m	1 year	~60K	✗	✗	✓	✓	✗	✗	✗
<b>Mapillary SLS (Ours)</b>	Urban + suburban	<b>11,560 km</b>	<b>4,228 km</b>	<b>7 years</b>	~1.68M	✓	✓	✓	✓	✓	✓	✓

Table 1: **Summary of place recognition datasets.** Geographical coverage is the length of unique traversed routes. Total length is the geographical coverage multiplied by the number of times each route was traversed. Temporal coverage is the time span from the first recording of a route to the last recording. “-” stands for “not applicable”.

objects such as moving pedestrians or cars, structural modifications such as roadworks or architectural work, camera intrinsics and viewpoints. Our data spans six continents, including diverse cities like Kampala, Zurich, Amman and Bangkok.

In addition to the dataset, we make several contributions related to its experimental validation. We benchmark particularly challenging scenarios such as day/night, seasonal and temporal changes. We tackle a wider set of problems not limited to image-to-image localization by proposing six variations of MultiViewNet [16] to model sequence-to-sequence place recognition. Moreover, we formulate two new research tasks: sequence-to-image and image-to-sequence recognition, and propose several feature descriptors that extend pretrained image-to-image models to these two new tasks.

## 2. Related Works

**Place Recognition.** Place recognition consists of finding the most similar place of a query image within a database of registered images [31, 55]. Traditional visual place descriptors are based on aggregating local features using bag-of-words [45], Fisher vectors [39] or VLAD [25]. Other hand-crafted approaches exploit geometric and/or temporal consistency [15, 17, 33] in image sequences. Torii et al. [50] synthesizes viewpoint changes from panorama images with associated depth. These synthetic images make the place descriptor, DenseVLAD [4, 26], more robust to viewpoint and day/night changes.

As in other computer vision tasks, deep features have demonstrated better performance than hand-crafted ones [55]. Initially, features from existing pre-trained networks were used for single-view place recognition [7, 11, 46–48]. Later works

demonstrated that the performance improves if the networks are trained for the specific task of place recognition [5, 22, 30]. One of the recent successes is NetVLAD [5, 55], which uses a base network (e.g. VGG16) followed by a generalized VLAD layer (NetVLAD) as an image descriptor. Other works, such as R-MAC [49] and Chen et al. [13], extract regions directly from the CNN response maps to form place descriptors.

Recent deep-learning-based methods exploit the temporal, spatial, and semantic information in images or image sequences. Radenovic et al. [42] proposes a pipeline to obtain large 3D scene reconstructions from unordered images and uses these 3D reconstructions as ground truth for training a Generalized Mean (GeM) layer with hard positive and negative mining. Garg et al. [18], on the other hand, uses single-view depth predictions to recognize places revisited from opposite directions. Also, addressing extreme viewpoint changes, Garg et al. [19] suggests semantically aggregating salient visual information. The 3D geometry of a place is also used by PointNetVLAD [2] that combines PointNet and NetVLAD to form a global place descriptor from LiDAR data. MultiViewNet [16] investigates different pooling strategies, descriptor fusion and LSTMs to model temporal information in image sequences. This research is, however, hindered by the lack of appropriate datasets.

**Place Recognition Datasets.** Table 1 summarizes a set of relevant place recognition datasets. Below we highlight more details and compare our contributions against existing datasets.

**Nordland [36, 37]** contains 4 sequences of a 182km-long train journey, traversed once per season. It captures seasonal changes but contains small variations in viewpoint, camera intrinsics, time of day or structural changes.

## 附录 D 外文文献译文

**题目：**Mapillary 街景级别序列：终身位置识别数据集

**作者：**Frederik Warburg et al.

**摘要：**你的翻译内容

引论

你的翻译