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 LaTeX 模板[①]，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。

关 键 词：西安交通大学学士学位论文；毕业设计；LaTeX

[①] 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

$\boldsymbol{x}$	请输入你对符号的解释与说明文字
$\boldsymbol{y}$	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[①]。

1.1 图

图 1-1 是用 Tikz 画的，Visio 画不出这么好看的图。

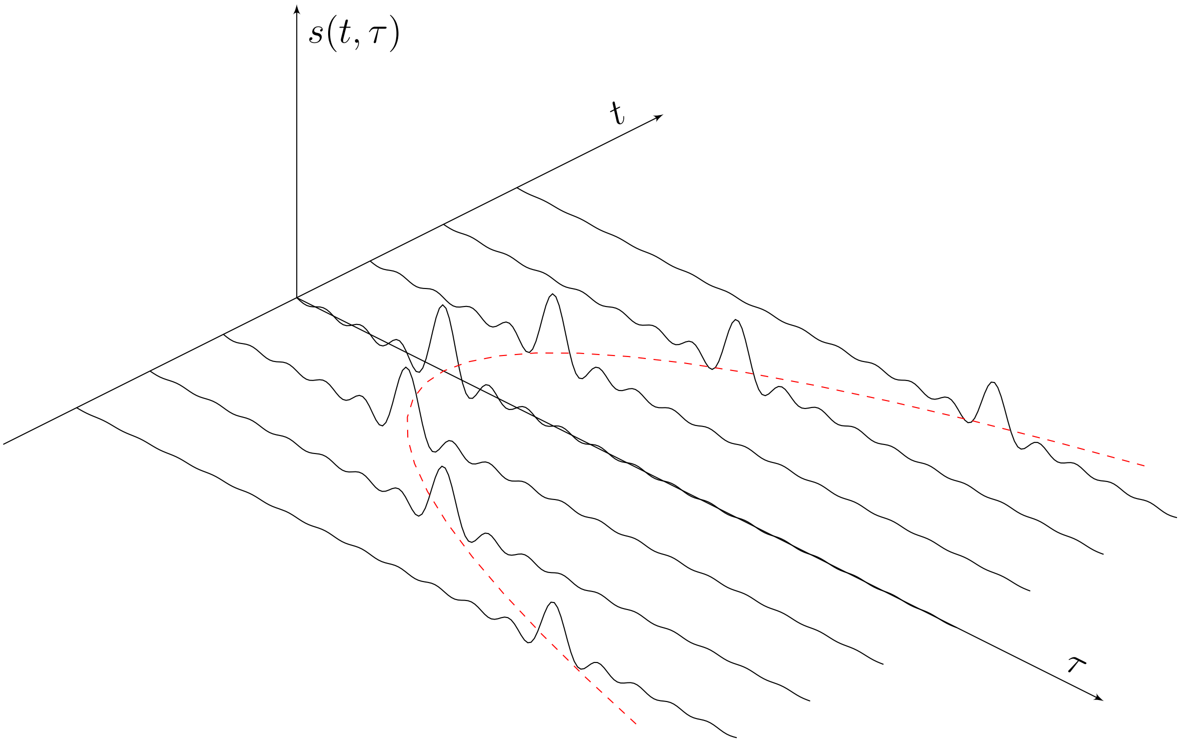

图 1-1 图注是五号字。

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

[①] 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 [1]，从公式可知 LaTeX 排版的公式比 MathType 美观，而且公式编写效率更高。

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

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 [2]。

$$X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t} \mathrm{d}t \tag{1-2a}$$

$$x(t) = \int_{-\infty}^{\infty} X(f) e^{j2\pi f t} \mathrm{d}f \tag{1-2b}$$

2 参考文献

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 [3]。

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.

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 ◊，例子的结束符采用 ♦，定理的结束符采用 □，证明的结束符采用 ■。

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)。 ◊

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：

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

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

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

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：

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

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。 ♦

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

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

并且满足 $\|f\| = \|x_f\|$。 □

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

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: $\mathbf{x}(k),\quad \mu,\quad \mathbf{w}(0)$

Output: $y(k),\quad \varepsilon(k)$

for $k = 0, 1, \cdots$ **do**

- $y(k) = \mathbf{w}^H(k)\mathbf{x}(k)$ `// output signal`
- $\varepsilon(k) = d(k) - y(k)$ `// error signal`
- $\mathbf{w}(k+1) = \mathbf{w}(k) + \mu\varepsilon^*(k)\mathbf{x}(k)$ `// weight vector update`

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

```verilog
module stap_lms
#(
parameter         M          = 4,    // number of antennas
                  L          = 5,    // length of FIR filter
                  W_IN       = 18,   // wordlength of input data
                  W_OUT      = 18,   // wordlength of output data
                  W_COEF     = 20    // wordlength of weights
)(
output  signed  [W_OUT-1:0]   y_i,   // in-phase component of STAP output
output  signed  [W_OUT-1:0]   y_q,   // quadrature component of STAP output
output                        vout,  // data valid flag of output (high)
input           [M*W_IN-1:0]  u_i,   // in-phase component of M antennas
input           [M*W_IN-1:0]  u_q,   // quadrature component of M antennas
input                         vin,   // data valid flag for input (high)
input                         clk,   // clock signal
input                         rst    // reset signal (high)
);
```

附录 C 外文文献原文

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

Frederik Warburg[†,*], Søren Hauberg[†], Manuel López-Antequera[‡], Pau Gargallo[‡],
Yubin Kuang[‡], and Javier Civera[§]

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

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.

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.[1]

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 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[2]. MSLS covers the following causes of appearance change: different seasons, changing weather conditions, varying illumination at different times of the day, dynamic

*The main part of this work was done while Frederik Warburg was an intern at Mapillary.

[1] www.mapillary.com/datasets/places

[2] See the video accompanying the paper for an overview and sample images.

1

						Type of appearance changes						
Name	Environment	Total length	Geographical coverage	Temporal coverage	Frames	Seasonal	Weather	Viewpoint	Dynamic	Day/night	Intrinsics	Structural
Nordland [36, 37]	Natural + urban	728 km	182 km	1 year	$\sim$115K	✓	✗	✗	✗	✗	✗	✗
SPED [12]	Urban	-	-	1 year	$\sim$**2.5M**	✓	✓	✗	✓	✓	✗	✗
KITTI [20]	Urban + suburban	39.2 km	1.7 km	3 days	$\sim$13K	✗	✗	✓	✓	✗	✗	✗
Eynsham [14]	Urban + suburban	70 km	35 km	1 day	$\sim$10K	✗	✗	✗	✓	✗	✗	✗
St. Lucia [21]	Suburban	47.5 km	9.5 km	1 day	$\sim$33K	✗	✗	✗	✓	✗	✗	✗
NCLT [9]	Campus	148.5 km	5.5 km	15 mon.	$\sim$300K	✓	✗	✓	✓	✗	✗	✗
Oxford RobotCar [32]	Urban + suburban	1.000 km	10 km	1 year	$\sim$27K	✓	✓	✓	✓	✓	✗	✓
VL-CMU [8]	Urban + suburban	128 km	8 km	1 year	$\sim$1.4K	✗	✗	✓	✓	✗	✗	✗
FAS [34]	Urban + suburban	120 km	70 km	3 years	$\sim$43K	✓	✓	✓	✓	✗	✗	✓
Garden Point [41]	Urban + campus	$<$12 km	4 km	1 week	$\sim$600	✗	✗	✓	✗	✓	✗	✗
SYNTHIA [44]	Urban	6 km	1.5 km	-	$\sim$200K	✓	✓	✓	✓	✓	✗	✗
GSV [56]	Urban	-	-	-	$\sim$60K	✗	✗	✗	✗	✗	✗	✗
Pittsburgh 250k [51]	Urban	-	-	-	$\sim$254K	✗	✗	✓	✓	✗	✗	✗
TokyoTM/247 [50]	Urban	-	-	-	$\sim$174K	✓	✗	✓	✓	✓	✗	✓
TB-places [28]	Gardens	$<$100m	$<$100m	1 year	$\sim$60K	✗	✗	✓	✓	✗	✗	✗
Mapillary SLS (Ours)	Urban + suburban	**11,560 km**	**4,228 km**	**7 years**	$\sim$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], Fischer 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