

Intelligence of Fusion

Deep Learning-based Image Fusion

Chinese Monograph

Version: v25.10.22

Preview Version

The full book is available at:

<https://xingchenzhang.github.io/imagefusionbook>

Xingchen Zhang

Department of Computer Science

University of Exeter

Abstract

In recent years, research on image fusion has developed rapidly, with many studies based on deep learning models published in leading journals and conferences. These advances have attracted growing interest from the computer vision community. However, most existing monographs were published before deep learning became prominent in the field and mainly focus on traditional methods.

This book, *Intelligence of Fusion: Deep Learning-Based Image Fusion*, aims to fill this gap by providing a comprehensive overview of deep learning-based image fusion. It is divided into three parts. The first part introduces the background and fundamental concepts of image fusion and deep learning. The second part presents key methods and techniques, including performance evaluation approaches, representative algorithms for visible-infrared, multi-focus, and multi-exposure fusion, as well as general and task-driven fusion frameworks. The third part focuses on practical implementation, applications, and future directions, including benchmark usage, real-world applications, and recent research trends.

The book also contains appendices listing relevant journals and conferences, open-source code links, and advice on writing image fusion papers.

Designed for students, researchers, and practitioners in image fusion, computer vision, and multimodal learning, this book offers both theoretical foundations and practical guidance for advancing work in this field.

目 录

前 言	V
第一部分 图像融合背景与概念	1
第一章 绪论	1
1.1 引言	1
1.2 图像融合的基本概念	3
1.3 图像融合的一个特点和两个目的	4
1.4 图像融合的分类	5
1.5 图像融合中的配准	9
1.6 基于深度学习的图像融合	10
1.7 本书的写作目的	11
1.8 本书主要内容与特色	12
第二章 人工智能基础知识	14
2.1 什么是深度学习?	14
2.1.1 优化问题	16
2.1.2 基于梯度的优化	16
2.2 深度学习三要素	16
2.2.1 数据	16
2.2.2 算力	17
2.2.3 算法	18
2.3 深度学习的分类	18
2.3.1 监督学习	18
2.3.2 无监督学习	19

2.3.3	强化学习	19
2.3.4	分类问题和回归问题	20
2.3.5	判别式模型和生成式模型	21
2.4	深度学习算法的常规设计流程	21
2.5	图灵测试	21
2.6	常用深度学习框架简介	23
2.7	小结	26

第二部分 图像融合方法与技术 27

第三章 基于人工智能的图像融合概述 28

3.1	传统图像融合方法简介	28
3.1.1	图像融合方法的三个步骤	29
3.1.2	传统图像融合方法的类别	29
3.1.3	传统图像融合方法的缺点	30
3.2	基于深度学习的图像融合发展状况概述	30
3.2.1	基于深度学习的目的	30
3.2.2	有监督方法和无监督方法	30
3.3	常用于图像融合的深度学习模型	31
3.3.1	卷积神经网络	31
3.3.2	生成对抗网络	31
3.3.3	Transformer	31
3.3.4	扩散模型	32
3.4	常用于图像融合的重要深度学习技术	32
3.4.1	注意力机制	32
3.4.2	残差连接	33
3.4.3	稠密连接	33
3.4.4	自动网络架构搜索	34
3.4.5	其他重要技术	34
3.5	与多模态机器学习的关系	34
3.6	基于深度学习的图像融合发展趋势	35
3.6.1	多种深度学习模型被用于图像融合	35

3.6.2	从非端到端的方法到端到端的方法	36
3.6.3	从特定图像融合方法到通用图像融合方法	36
3.6.4	从生成融合图像到改进下游应用	36
3.6.5	新型图像融合类型开始出现	37
3.7	小结	37
第四章	图像融合算法性能评价	38
4.1	图像融合算法评价的特殊性	38
4.2	当前的主要图像融合评价方法	39
4.2.1	图像融合定性评价方法	39
4.2.2	图像融合定量评价方法	41
4.2.3	图像融合评价方法现状	43
4.3	其他评价方法	44
4.4	近年来的发展特点	45
4.5	图像融合评价方法的发展趋势	53
4.5.1	设计更好评价基准	53
4.5.2	基于具体应用的性能评价	54
4.6	小结	55
第五章	可见光与红外图像融合	56
5.1	红外图像：从另一个视角感知世界	56
5.2	可见光与红外图像融合概述	58
5.3	传统融合方法概述	60
5.4	使用深度学习做图像融合的动机	60
5.5	基于深度学习的融合方法发展历程概述	61
5.6	基于深度学习的可见光与红外图像融合方法分类	62
5.7	基于深度学习的可见光与红外图像融合方法介绍	62
5.7.1	单分支模型和双分支模型	62
5.7.2	基于卷积神经网络的图像融合方法	63
5.7.3	基于自编码器的图像融合方法	67
5.7.4	基于生成式对抗网络的图像融合方法	68
5.7.5	基于 Transformer 的图像融合方法	72
5.7.6	基于扩散模型的图像融合方法	72

5.8	可见光与红外图像融合的发展特点	74
5.9	未来发展趋势	80
5.10	小结	83
第六章	多聚焦图像融合	84
6.1	多聚焦图像融合概述	84
6.2	传统多聚焦图像融合方法概述	85
6.3	基于深度学习的多聚焦融合方法	86
6.3.1	基于深度学习的多聚焦图像融合方法的分类	86
6.3.2	基于有监督学习的融合方法	87
6.3.3	基于无监督学习的融合方法	91
6.4	训练数据的获取	92
6.4.1	有监督多聚焦图像融合方法	92
6.4.2	无监督多聚焦图像融合方法	93
6.5	多聚焦图像融合的发展趋势	93
6.6	小结	94
第七章	多曝光图像融合	95
7.1	多曝光图像融合概述	95
7.2	多曝光图像融合的特点	96
7.3	多曝光图像融合方法的分类	96
7.3.1	融合两张源图像的方法和融合多张源图像的方法	96
7.3.2	静态图像融合和动态图像融合	97
7.3.3	传统方法和深度学习方法	98
7.4	基于深度学习的融合方法	98
7.4.1	有监督学习方法	98
7.4.2	无监督学习方法	99
7.5	多曝光图像融合的发展趋势	102
7.6	小结	103
第八章	通用图像融合方法	104
8.1	传统通用图像融合方法	104
8.2	基于深度学习的通用图像融合方法	105
8.2.1	概述与分类	105

8.2.2 实现方式	106
8.3 通用图像融合方法的优缺点	107
8.3.1 优点	107
8.3.2 缺点	107
8.4 小结	108
第九章 应用驱动的图像融合方法	109
9.1 应用驱动的图像融合方法的优势	110
9.2 应用驱动的可见光与红外图像融合方法	111
9.2.1 行人检测驱动的图像融合方法	111
9.2.2 语义分割驱动的图像融合方法	112
9.2.3 通用目标检测驱动的图像融合方法	113
9.2.4 人群计数驱动的图像融合方法	113
9.3 小结	113
第三部分 图像融合实践、应用与展望	116
第十章 图像融合实践	117
10.1 编程语言及深度学习框架选择	117
10.1.1 编程语言选择	117
10.1.2 深度学习框架选择	118
10.2 使用 VIFB 进行可见光红外图像融合	119
第十一章 可见光与红外图像融合的应用	123
11.1 红外图像的常见应用总结	123
11.2 红外图像的缺点	126
11.3 像素级可见光和红外图像融合的应用	127
11.4 其他层级的可见光与红外图像融合的应用	132
11.4.1 基于特征级融合的应用	132
11.4.2 基于决策级融合的应用	132
11.4.3 基于多个层级融合的应用	133
11.5 可见光与红外图像融合的应用小结	133
11.6 展望	134

11.7 小结	136
第十二章 多聚焦图像融合的应用	137
12.1 多聚焦图像融合的应用概述	137
12.2 基于多聚焦图像融合的远距离人脸检测	139
12.3 基于多聚焦图像融合的光学显微图像融合	139
12.4 基于多聚焦图像融合的深度估计	141
12.5 小结	143
第十三章 多曝光图像融合的应用	144
13.1 多曝光图像融合的应用概述	144
13.2 基于多曝光图像融合的语义分割	144
13.3 提升显微图像质量	146
13.4 小结	148
第十四章 图像融合的前沿进展	149
14.1 与其他任务相结合	149
14.2 通用图像融合方法	150
14.3 关于评价基准的研究	150
14.4 基于具体应用的融合方法性能评价	150
14.5 将图像配准和图像融合进行结合	151
14.6 其他类型的图像融合	152
14.6.1 可见光图像与近红外图像融合	152
14.6.2 偏振图像融合	153
14.6.3 RGB 图像融合	153
14.6.4 可见光图像与深度图像融合	154
14.6.5 可见光图像与事件相机数据融合	155
14.6.6 多视角图像融合	156
14.7 小结	156
第十五章 总结与展望	157
15.1 总结	157
15.2 待解决的问题	158
15.3 展望	160

目 录	XVI
附录 A 图像融合相关的学术期刊和学术会议	166
附录 B 图像融合相关开源代码下载链接	168
附录 C 图像融合论文写作经验	169
索 引	173
参考文献	174