

国家自然科学基金资助项目  
信息技术重点图书

# 基于学习的 图像增强技术

*Learning based Image  
Enhancement*

主 编 吴 炜  
副主编 陶青川



西安电子科技大学出版社  
<http://www.xduph.com>

国家自然科学基金资助项目

信息技术重点图书

# 基于学习的图像增强技术

Learning based Image Enhancement

主 编 吴 炜

副主编 陶青川

编 委 庞 宇 严斌宇 孟庆党

西安电子科技大学出版社

## 内 容 简 介

本书是一部关于基于学习的图像增强原理及应用的学术专著,反映了近年来基于学习的图像增强技术的最新研究进展。全书共分为三个部分十四章,第一部分为基础知识(第一章至第三章,介绍图像的基本概念、图像增强的一些基本方法和图像插值技术;第二部分为基于学习的图像分辨率增强(也称为基于学习的图像超分辨率)技术(第四章至第十三章),对基于学习的图像超分辨率技术进行了详细的介绍;第三部分(第十四章)介绍了一种新的图像增强技术——基于视觉美学学习的图像质量评估和增强技术。

本书适合于通信与信息系统、信号处理、计算机应用、模式识别等相关专业的研究人员、工程技术人员、高校教师以及硕/博士研究生学习参考。

## 图书在版编目(CIP)数据

基于学习的图像增强技术/吴炜主编. —西安:西安电子科技大学出版社,2013.2

国家自然科学基金资助项目

ISBN 978 - 7 - 5606 - 2981 - 0

I. ① 基… II. ① 吴 III. ① 机器学习—研究 IV. TP181

中国版本图书馆 CIP 数据核字(2013)第 023789 号

策划编辑 李惠萍

责任编辑 买永莲 李惠萍

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)88242885 88201467 邮 编 710071

网 址 www.xduph.com 电子邮箱 xdupfxb001@163.com

经 销 新华书店

印刷单位 陕西华沐印刷科技有限责任公司

版 次 2013年2月第1版 2013年2月第1次印刷

开 本 787毫米×1092毫米 1/16 印张 14

字 数 318千字

印 数 1~2000册

定 价 25.00元

ISBN 978 - 7 - 5606 - 2981 - 0/TP

**XDUP 3273001 - 1**

\* \* \* 如有印装问题可调换 \* \* \*

本书的研究工作受到国家自然科学基金(联合基于学习的超分辨率技术和多传感器超分辨率技术在红外图像复原中的研究,资助号为 61271330;多视频时空超分辨率重建技术研究,资助号为 61071161)、教育部重点项目(视频超分辨率重建关键技术研究,资助号为 107094)的资助,谨在此特别致谢。

# 前言

自计算机问世以来,人们就希望它能够进行自我学习。然而到目前为止,计算机还不具有与人类(或者其它灵长类生物)同样强大的学习能力,但就特定的任务而言,机器学习已经取得了不错的成绩。特别是近些年来,随着计算机技术的飞速发展,机器学习技术也获得了快速发展,成为人工智能最前沿的研究领域之一。

来自维基百科的机器学习定义包括下面几种:(1)机器学习是一门人工智能的科学,该领域的主要研究对象是人工智能,特别是如何在经验学习中改善具体算法的性能;(2)机器学习是对能通过经验自动改进的计算机算法的研究;(3)机器学习是用数据或以往的经验,以此优化计算机程序的性能标准。一种经常引用的英文定义是:A computer program is said to learn from experience  $E$  with respect to some class of tasks  $T$  and performance measure  $P$ , if its performance at tasks in  $T$ , as measured by  $P$ , improves with experience  $E$ .

机器学习在各个领域获得了广泛的应用,一些典型的应用包括语音识别、数据挖掘、自然语言处理、医学诊断、检测信用卡欺诈等。机器学习在图像处理领域同样获得了广泛的应用,例如目前数码相机中使用的人脸检测技术、笑脸检测技术等,指纹考勤系统中的指纹识别技术,智能交通系统中的车牌识别技术等。

近年来,随着多媒体技术的快速发展以及数码设备的普及,视频图像的应用领域越来越广,但是由于图像采集设备的限制,拍摄的图像质量常常较低,这严重影响了对图像的后续处理。因此有必要对这些低质量的图像进行增强处理,以提高图像质量。探索合适的图像增强算法,获取视觉效果良好的图像,具有十分重要的意义。

目前图像增强方法可分为两大类,传统的图像增强方法和基于学习的图像增强方法。传统的图像增强方法主要是依据信号处理的相关知识对图像进行增强,这种方法的不足是增强的效果有限、增强方式不灵活等。基于学习的图像增强技术是模拟人的学习的过程,通过机器学习技术获取图像增强的某些特征,然后将这些特征应用于图像的增强。该方法在增强过程中可充分考虑人的视觉特性,具有增强方式更灵活、增强效果更好等优点。

本书主要讨论和介绍基于学习的图像增强技术(也称为基于机器学习的图像增强技术),图0-1为基于学习的图像增强技术的一个框架示意图。在该框架中,须事先建立一个已知的训练库,该训练库中包含多个增强前后的图像对(增强的具体方法是未知的)。在需要增强图像时,通过机器学习获得图像增强函数  $F(\cdot)$  并在训练库中图像的辅助下,对输入的图像进行增强。

基于学习的图像增强技术近年来已广泛应用于图像的分辨率增强(也称为超分辨率技术)、图像复原、个性化图像增强以及基于视觉美学的图像增强等。随着机器学习技术和图像增强技术的发展,基于学习的图像增强技术将获得越来越广泛的应用。

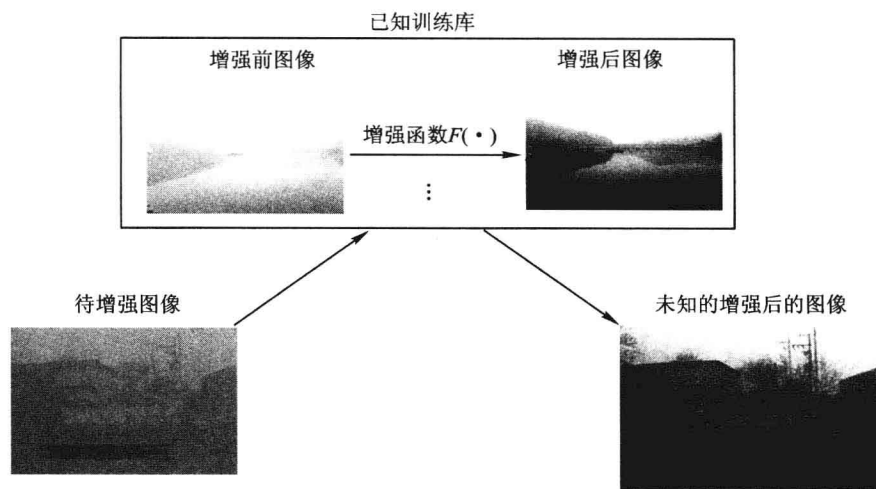


图 0-1 基于学习的图像增强技术框架示意图

本书是一部关于基于学习的图像增强方面的学术专著，全面地讨论了基于学习的图像超分辨率技术(即图像的分辨率增强)和图像复原技术。

本书由吴炜、陶青川、庞宇、严斌宇、孟庆党共同完成，其中吴炜、陶青川分别担任主编和副主编。

我们对参加本书相关工作的雷翔硕士、黄晓强硕士、郑丽贤硕士，以及参与了部分章节写作的窦翔硕士(参与了第一章的写作)、王瑶硕士(参与了第三章的写作)、杨皖钦硕士(参与了第十四章的写作)在此表示感谢。

本书的内容涉及笔者近 10 年来在四川大学电子信息学院工作的一些研究成果。四川大学电子信息学院的何小海教授对本书的工作给予大力支持，提出了不少宝贵的建议和意见，可以说没有他的支持，很难有本书的出版，在此特别表示感谢。另外，在此也感谢陶德元教授、罗代升教授、滕奇志教授对本书工作的支持。

本书在编写和出版的过程中，得到了西安电子科技大学出版社及李惠萍老师的帮助，在此表示感谢！

本书的研究工作受到国家自然科学基金(联合基于学习的超分辨率技术和多传感器超分辨率技术在红外图像复原中的研究，资助号为 61271330；多视频时空超分辨率重建技术研究，资助号为 61071161)、教育部重点项目(视频超分辨率重建关键技术研究，资助号为 107094)的资助，在此表示感谢。

由于作者水平有限，书中难免有不足和不当之处，恳请广大读者批评指正。

吴炜、陶青川于四川大学望江校区

2012 年 7 月

## 作者简介



**吴 炜** 男, 博士, 副教授。1994 年 9 月至 1998 年 7 月就读于天津大学, 获学士学位; 2000 年 9 月至 2003 年 7 月就读于四川大学电子信息学院, 获硕士学位; 2003 年 7 月至今在四川大学电子信息学院任教; 2008 年 7 月在四川大学获通信与信息系统博士学位; 2009 年 10 月至 2010 年 10 月在加拿大国家研究院从事为期一年的博士后研究。主要从事图像处理、模式识别、机器学习等理论和技术研究, 并担任 IEEE Transactions on Instrumentation and Measurement、Machine Vision and Applications、Journal of Electronic Imaging 等国际学术期刊的审稿人和美国 Journal of Pattern Recognition Research 期刊的 Prospective Editor, 主持或参与并完成了包括联合基于学习的超分辨率技术和多传感器超分辨率技术在红外图像复原中的研究(国家自然科学基金: 61271330)、计算光学切片显微三维成像技术(国家自然科学基金: 60372079)、视频超分辨率重建关键技术研究(教育部重点项目资助: 107094)、智能交通系统、景区门票“人票合一”验证管理系统等多项纵向、横向科研项目, 在国内外重要刊物和会议上发表论文 50 余篇(其中被 SCI、EI 收录 20 余篇)。



**陶青川** 男, 博士, 副教授, 模式识别专业硕士导师。1993 年 9 月至 1997 年 7 月就读于四川大学电子信息学院, 获无线电电子学学士学位; 1997 年 9 月至 2000 年 7 月就读于四川大学电子信息学院, 获模式识别硕士学位; 2000 年 7 月至今在四川大学电子信息学院任教; 2005 年获光学专业博士学位。主要从事图像处理、模式识别、机器学习、信息光学、生物医学等理论和技术研究, 同时具有丰富的信息化项目工程经验。目前主讲研究生的模式识别学位课程, 并主持或参加了激光共焦三维生物医学图像处理与识别(国家自然科学基金: 30070228)、计算光学切片显微三维成像技术(国家自然科学基金: 60372079)、多视频时空超分辨率重建技术研究(国家自然科学基金: 61071161)、人脸识别、步态识别、车牌识别、岩心扫描成像及金相图像分析系统等多项纵向、横向科研项目; 在国内外重要刊物和会议上发表论文近 40 篇(其中被 SCI、EI 收录 10 篇)。

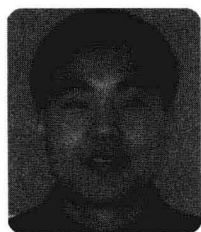


**庞 宇** 男, 博士, 副教授。1996 年 9 月至 2000 年 7 月就读于四川大学, 获学士学位; 2000 年 9 月至 2003 年 7 月就读于电子科技大学, 获硕士学位; 2004 年 1 月~2010 年 6 月分别就读于加拿大 Concordia 大学和 McGill 大学, 获博士学位; 2010 年 9 月至今, 在重庆邮电大学任教。主要从事医学图像处理、远程医疗系统设计、超短距离无线通信、集成电路设计与综合等方面的研究。主持了包括可逆逻辑综合理论

与实现方法研究(国家自然科学基金: 61102075; 重庆市自然科学基金: CSTC2011BB2142)、重庆儿童医院体温组网实时监护系统、NFC 儿童无线体温系统、NFC 基础体温系统等多个纵向和横向项目, 参与了工信部物联网专项与重庆市重点科技攻关项目; 指导研究生 40 余名, 所发表的文章被 SCI 和 EI 收录的有 30 余篇。



**严斌宇** 男, 博士, 讲师。1993 年 9 月至 1997 年 6 月就读于四川大学物理系, 获学士学位; 1999 年 9 月至 2002 年 6 月就读于四川大学计算机学院, 获工学硕士学位; 2002 年 7 月至今, 在四川大学电子信息学院任教, 期间于 2011 年 12 月在四川大学获工学博士学位; 2012 年 7 月至 8 月在美国西密西根大学进行为期一月的访问交流。主要从事计算机网络、无线传感网络等理论和技术研究。参与并完成了国家自然科学基金项目“容延迟移动传感器网络中通信协议的研究(60773168)”, 主持了四川大学青年基金项目“无线传感器网络信任模型研究(2010SCU11006)”, “中石油 MES 系统二期工程分包项目”、“智能型仪器运行状态采集仪研制”等多项纵向、横向科研项目; 在国内外重要刊物和会议上发表论文 10 余篇, 其中被 EI 收录的有 4 篇。



**孟庆党** 男, 通信与信息系统专业数字图像处理方向硕士。1992 年 9 月至 1996 年 7 月就读于河北大学无线电电子学系, 获学士学位; 1996 年 9 月至 1999 年 7 月就读于四川大学电子信息学院, 获硕士学位; 1999 年 7 月至今, 在四川大学电子信息学院任教。主要从事图像处理、计算机网络、信息安全等理论和技术研究, 同时具有丰富的信息系统集成项目工程经验。目前主讲现代通信网、交换技术、移动通信、计算机网络、软交换与 NGN 等课程, 并主持或参加了多项军工科研项目、多项省级网络项目的规划与实施。

# 目 录

|                    |    |
|--------------------|----|
| 第一章 图像的基础知识        | 1  |
| 1.1 图像信号的基本概念      | 1  |
| 1.1.1 图像的表示        | 1  |
| 1.1.2 图像的数字化过程     | 2  |
| 1.1.3 数字图像的基本类型    | 7  |
| 1.1.4 颜色模式         | 8  |
| 1.1.5 图像分辨率        | 15 |
| 1.2 人眼的视觉原理        | 16 |
| 1.2.1 人眼结构         | 16 |
| 1.2.2 相对视敏度        | 18 |
| 1.2.3 明暗视觉         | 19 |
| 1.2.4 对比灵敏度        | 19 |
| 1.2.5 可见度阈值和马赫带效应  | 20 |
| 1.3 图像质量的评估标准与方法   | 20 |
| 参考文献               | 23 |
| 第二章 常用图像增强技术介绍     | 25 |
| 2.1 图像增强概述         | 25 |
| 2.2 空域图像增强         | 25 |
| 2.2.1 灰度变换         | 25 |
| 2.2.2 直方图均衡        | 28 |
| 2.2.3 空域滤波         | 29 |
| 2.3 频域图像增强         | 30 |
| 2.3.1 频域低通滤波器      | 30 |
| 2.3.2 频域高通滤波器      | 31 |
| 2.4 图像客观评价算法       | 31 |
| 2.4.1 人眼视觉系统       | 31 |
| 2.4.2 归一化灰度差       | 32 |
| 2.4.3 归一化对比度       | 32 |
| 2.4.4 归一化信息熵       | 32 |
| 2.4.5 视频图像质量客观评价函数 | 33 |
| 2.5 本章小结           | 33 |

|                                        |    |
|----------------------------------------|----|
| 参考文献 .....                             | 33 |
| <b>第三章 图像插值技术</b> .....                | 34 |
| 3.1 图像插值放大原理 .....                     | 34 |
| 3.2 传统图像插值算法及原理 .....                  | 35 |
| 3.3 基于边缘的图像插值算法 .....                  | 38 |
| 3.4 实验结果与分析 .....                      | 47 |
| 3.5 本章小结 .....                         | 51 |
| 参考文献 .....                             | 52 |
| <b>第四章 超分辨率技术综述</b> .....              | 53 |
| 4.1 超分辨率的含义及应用 .....                   | 54 |
| 4.2 超分辨率技术的分类 .....                    | 55 |
| 4.3 成像模型 .....                         | 56 |
| 4.4 基于重建的超分辨率 .....                    | 57 |
| 4.4.1 频域算法 .....                       | 58 |
| 4.4.2 空域算法 .....                       | 58 |
| 4.5 基于学习的超分辨率 .....                    | 60 |
| 4.5.1 最大后验概率(MAP)框架下的基于学习的超分辨率理论 ..... | 61 |
| 4.5.2 基于学习的超分辨率算法的类别 .....             | 63 |
| 4.6 本章小结 .....                         | 65 |
| 参考文献 .....                             | 66 |
| <b>第五章 基于多分辨率塔式结构的人脸图像超分辨率技术</b> ..... | 68 |
| 5.1 基于学习的人脸超分辨率系统 .....                | 68 |
| 5.2 幻觉脸技术的复原框架 .....                   | 69 |
| 5.3 图像金字塔模型 .....                      | 70 |
| 5.4 多分辨率塔式结构算法 .....                   | 73 |
| 5.4.1 人脸高斯金字塔 .....                    | 74 |
| 5.4.2 人脸拉普拉斯金字塔 .....                  | 74 |
| 5.4.3 人脸特征金字塔 .....                    | 74 |
| 5.4.4 多分辨率塔式结构算法总结 .....               | 76 |
| 5.5 匹配复原过程 .....                       | 76 |
| 5.5.1 塔状父结构 .....                      | 77 |
| 5.5.2 搜索匹配过程 .....                     | 78 |
| 5.6 算法描述 .....                         | 79 |
| 5.7 基于学习的超分辨率图像的集成优化 .....             | 80 |
| 5.7.1 超分辨率复原的贝叶斯框架 .....               | 80 |
| 5.7.2 单目标优化算法 .....                    | 81 |
| 5.8 实验结果与分析 .....                      | 84 |
| 5.8.1 多分辨率塔式结构算法实验结果与分析 .....          | 84 |
| 5.8.2 集成优化实验结果与分析 .....                | 86 |

|                                                   |            |
|---------------------------------------------------|------------|
| 5.9 本章小结 .....                                    | 87         |
| 参考文献 .....                                        | 88         |
| <b>第六章 基于 Contourlet 变换的人脸图像超分辨率研究 .....</b>      | <b>89</b>  |
| 6.1 Contourlet 变换的基本理论 .....                      | 90         |
| 6.1.1 方向滤波器组 .....                                | 90         |
| 6.1.2 Contourlet 变换的特性分析 .....                    | 93         |
| 6.2 基于 Contourlet 变换的人脸图像超分辨率 .....               | 95         |
| 6.2.1 特征提取 .....                                  | 95         |
| 6.2.2 匹配复原 .....                                  | 97         |
| 6.2.3 算法描述 .....                                  | 98         |
| 6.3 实验结果与分析 .....                                 | 99         |
| 6.4 本章小结 .....                                    | 101        |
| 参考文献 .....                                        | 102        |
| <b>第七章 基于改进的非下采样 Contourlet 变换的人脸图像超分辨率 .....</b> | <b>103</b> |
| 7.1 非下采样 Contourlet 变换 .....                      | 103        |
| 7.1.1 非下采样金字塔 .....                               | 104        |
| 7.1.2 非下采样方向滤波器组 .....                            | 105        |
| 7.2 改进的非下采样 Contourlet 变换 .....                   | 106        |
| 7.3 算法实现 .....                                    | 108        |
| 7.4 实验结果与分析 .....                                 | 109        |
| 7.5 本章小结 .....                                    | 112        |
| 参考文献 .....                                        | 112        |
| <b>第八章 基于马尔可夫随机场的超分辨率技术研究 .....</b>               | <b>114</b> |
| 8.1 马尔可夫随机场模型 .....                               | 114        |
| 8.2 特征表示 .....                                    | 117        |
| 8.3 基于马尔可夫随机场模型的超分辨率学习算法 .....                    | 118        |
| 8.4 实验结果与分析 .....                                 | 120        |
| 8.5 本章小结 .....                                    | 123        |
| 参考文献 .....                                        | 123        |
| <b>第九章 基于重构方法的超分辨率研究 .....</b>                    | <b>124</b> |
| 9.1 基于主成分分析重构的超分辨率算法 .....                        | 124        |
| 9.1.1 基于整幅图像的 PCA 重构算法 .....                      | 126        |
| 9.1.2 基于分块的 PCA 重构算法 .....                        | 126        |
| 9.2 基于流形学习重构的算法 .....                             | 127        |
| 9.2.1 LLE 算法的基本原理 .....                           | 127        |
| 9.2.2 基于流形学习的超分辨率基本原理 .....                       | 129        |
| 9.2.3 特征提取 .....                                  | 130        |
| 9.2.4 算法实现 .....                                  | 131        |
| 9.3 实验结果与分析 .....                                 | 132        |

|                                          |            |
|------------------------------------------|------------|
| 9.4 本章小结 .....                           | 137        |
| 参考文献 .....                               | 137        |
| <b>第十章 基于超完备字典的图像稀疏表示理论的超分辨率复原 .....</b> | <b>138</b> |
| 10.1 概述 .....                            | 138        |
| 10.1.1 信号的稀疏表示及其研究现状 .....               | 138        |
| 10.1.2 信号稀疏性表示 .....                     | 139        |
| 10.1.3 超完备字典的基本概念 .....                  | 139        |
| 10.2 信号稀疏分解算法 .....                      | 140        |
| 10.2.1 引言 .....                          | 140        |
| 10.2.2 框架算法 .....                        | 141        |
| 10.2.3 匹配追踪算法 .....                      | 141        |
| 10.2.4 最佳正交基算法 .....                     | 142        |
| 10.2.5 全局最优算法 .....                      | 142        |
| 10.3 超完备字典学习算法 .....                     | 143        |
| 10.3.1 常用的字典学习算法 .....                   | 143        |
| 10.3.2 超完备字典学习算法的比较 .....                | 145        |
| 10.4 基于图像稀疏表示的单幅图像超分辨率算法 .....           | 147        |
| 10.4.1 自训练字典学习的算法框架 .....                | 148        |
| 10.4.2 由粗到精的图像放大过程 .....                 | 150        |
| 10.4.3 低分辨率和高分辨率超完备字典学习算法 .....          | 151        |
| 10.5 实验结果及分析 .....                       | 152        |
| 10.5.1 文本图像放大实验 .....                    | 153        |
| 10.5.2 与其他基于学习超分辨率算法对比 .....             | 154        |
| 10.5.3 图像特征提取算法对重建效果的影响 .....            | 156        |
| 10.5.4 目标放大倍数对重建效果的影响 .....              | 157        |
| 10.5.5 超完备字典尺寸对本章算法的影响 .....             | 158        |
| 10.6 本章小结 .....                          | 159        |
| 参考文献 .....                               | 160        |
| <b>第十一章 基于回归方法的超分辨率图像复原研究 .....</b>      | <b>162</b> |
| 11.1 支持向量回归 .....                        | 163        |
| 11.2 核偏最小二乘法回归 .....                     | 164        |
| 11.2.1 偏最小二乘法介绍 .....                    | 165        |
| 11.2.2 核偏最小二乘法 .....                     | 166        |
| 11.3 基于回归方法的超分辨率复原的基本原理 .....            | 167        |
| 11.3.1 超分辨率图像复原原理 .....                  | 167        |
| 11.3.2 特征表示 .....                        | 167        |
| 11.3.3 基于回归的图像超分辨率复原算法 .....             | 169        |
| 11.4 基于支持向量回归方法的实验结果与分析 .....            | 170        |
| 11.4.1 算法的性能 .....                       | 170        |

|                                                |            |
|------------------------------------------------|------------|
| 11.4.2 算法参数分析                                  | 173        |
| 11.5 基于核偏最小二乘法的超分辨率实验结果与分析                     | 174        |
| 11.6 本章小结                                      | 176        |
| 参考文献                                           | 176        |
| <b>第十二章 基于多分辨率金字塔和 LLE 算法的人脸图像超分辨率算法</b>       | <b>178</b> |
| 12.1 先验模型                                      | 178        |
| 12.1.1 高斯金字塔                                   | 178        |
| 12.1.2 拉普拉斯金字塔                                 | 178        |
| 12.1.3 特征金字塔                                   | 179        |
| 12.2 先验模型复原过程                                  | 179        |
| 12.2.1 塔状父结构                                   | 180        |
| 12.2.2 匹配复原                                    | 181        |
| 12.2.3 算法实现                                    | 181        |
| 12.3 先验模型和测量模型                                 | 182        |
| 12.4 实验结果及分析                                   | 183        |
| 12.5 本章小结                                      | 184        |
| 参考文献                                           | 185        |
| <b>第十三章 基于马尔可夫模型与 Contourlet 变换的图像超分辨率复原算法</b> | <b>186</b> |
| 13.1 算法的基本原理                                   | 186        |
| 13.2 Contourlet 系数块结构                          | 187        |
| 13.3 马尔可夫模型                                    | 188        |
| 13.4 基于 MRF 模型和 Cotourlet 变换的超分辨率学习算法          | 190        |
| 13.5 实验结果与分析                                   | 192        |
| 13.6 本章小结                                      | 199        |
| 参考文献                                           | 199        |
| <b>第十四章 基于视觉美学学习的图像质量评估和增强</b>                 | <b>200</b> |
| 14.1 基于学习的美学                                   | 202        |
| 14.1.1 用户调查                                    | 202        |
| 14.1.2 视觉美学特征                                  | 203        |
| 14.2 重建照片,增强照片质量                               | 204        |
| 14.2.1 算法 I 优化对象位置                             | 205        |
| 14.2.2 算法 II 平衡视觉重量                            | 206        |
| 14.3 实验结果与分析                                   | 207        |
| 14.4 本章小结                                      | 209        |
| 参考文献                                           | 209        |

# 第一章 图像的基础知识

图像来自于自然界，是人类获取信息和交换信息必不可少的方式之一。图像通过人的视觉感知获得，它可以记录在纸上，通过相机拍摄在相片上，显示在计算机、电视等显示设备上。

图像依据不同的存储方式，一般可以分为模拟图像和应用十分广泛的数字图像。图像最初的形态是连续的模拟量，也就是模拟图像，它的各点亮度信息由光、电等物理量的强弱来记录。为了便于数字设备(计算机存储)的存储和显示，以及进一步的处理，采用数字来表示图像的亮度信息，这样的图像就是数字图像。

数字图像处理(Digital Image Processing)就是利用计算机及其相关技术，先把模拟图像转化为数字图像，然后对数字图像施加某种或某些运算和处理，从而达到某种预期的处理目的。随着数字计算机及其相关技术的飞速发展，数字图像处理已经成为了一门独立的、有着强大研究潜力的学科。

## 1.1 图像信号的基本概念

文字、语音和图像信息是人们在日常的生活和工作中经常接触的信息形式。其中，语音和简单的图像(图形)是人类早期用于信息交流的主要方式。人类从自然界获得的信息大部分都来自于视觉(即图像)。图像与文字、语音相比，具有直观生动、具体形象等许多显著的优点。

图像是由光能量进入人的视觉系统所重现出来的视觉信息。光能量或者由发光物体直接发出，或者由光源发出后照射在物体上经过反射、衍射、折射等形成。图像来自于自然界或者人为的处理和合成，其原始的形态是连续变换的模拟量。与文字、语音信息相比较，图像信息主要具有信息量大、直观性强以及模糊性、实体性和形象化等特点。

在科技急速发展的现代社会，计算机、微电子、网络、信息处理等技术的发展使人类进入了信息时代，图像处理的各种应用在人们的社会生活、科学工作中发挥着越来越重要的作用。最原始的模拟图像，不便于传输、显示和进一步处理，为提高图像的质量、减小成本和图像的规模等，一般先通过计算机技术将模拟图像数字化。基于学习的图像针对的就是数字化图像。

### 1.1.1 图像的表达

一幅图像实际上记录的是物体辐射能量的空间分布(可以看做空间各点光强度的集合)，这个分布是空间坐标、时间和波长的函数。图像的亮度一般可以用多变量表示为

$$I = f(x, y, z, \lambda, t) \quad (1-1)$$

其中,  $x, y, z$  表示空间某个点的坐标,  $\lambda$  为光的波长,  $t$  为时间轴坐标。当  $z=z_0$  (常数) 时, 表示二维图像; 当  $t=t_0$  (常数) 时, 表示静态图像; 当  $\lambda=\lambda_0$  (常数) 时, 表示单色图像。

当一幅图像为平面单色静止图像时, 空间坐标变量  $z$ 、波长  $\lambda$  和时间变量  $t$  可以从函数中去除, 一幅图像可以用二维函数  $f(x, y)$  来表示:

$$f(x, y) = i(x, y)r(x, y) \quad (1-2)$$

其中,  $i(x, y)$  为照射分量(入射分量), 是入射到景物上的光强度;  $r(x, y)$  为反射分量, 是受到物体反射的光强度, 其大小限制在  $0 \sim 1$  之间。

$i(x, y)$  的性质取决于照射源, 而  $r(x, y)$  取决于成像物体的特性。

由于  $I$  表示的是物体的反射、投射或辐射能量, 因此它是正的、有界的, 即

$$0 \leq I \leq I_{\max} \quad (1-3)$$

其中,  $I_{\max}$  表示  $I$  的最大值,  $I=0$  表示绝对黑色。

式(1-1)是一个多变量函数, 不易于分析, 需要采用一些有效的方法进行降维。由三基色原理知,  $I$  可表示为三个基色分量的和:

$$I = I_R + I_G + I_B \quad (1-4)$$

式中,

$$\begin{cases} I_R = f_R(x, y, z, \lambda_R) \\ I_G = f_G(x, y, z, \lambda_G) \\ I_B = f_B(x, y, z, \lambda_B) \end{cases} \quad (1-5)$$

其中,  $\lambda_R, \lambda_G, \lambda_B$  为三个基色的波长。

由于上式中的每个彩色分量都可以看做一幅黑白(灰度)图像, 所以在以后的讨论中, 所有对于黑白图像的理论和方法都适用于彩色图像的每个分量。

一般地, 一个完整的图像处理系统输入和显示的都是便于人眼观察的连续图像(模拟图像)。为了便于数字存储和计算机处理, 可以通过模/数转换(A/D)将连续图像变为数字图像。反过来, 通过数/模转换(D/A)也可以将数字图像还原为模拟图像。

### 1.1.2 图像的数字化过程

一般的图像(模拟图像)无法直接用计算机进行处理, 为此必须将各类模拟图像转化为数字图像。数字图像的生成过程如图 1-1 所示。数字图像可以理解为对二维函数  $f(x, y)$  进行采样和量化(即离散处理)后得到的图像, 通常用二维矩阵来表示一幅数字图像。对一幅图像进行数字化的过程就是在计算机内生成一个二维矩阵的过程。因此, 采样(选取一种离散的栅格去表示一幅图像)和量化(将颜色和亮度映射成整数)是图像数字化的两个主要过程。

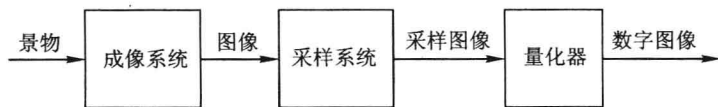


图 1-1 数字图像生成过程

连续图像数字化的结果，将产生能够用于计算机处理的离散数据结构，一般为一个矩阵。假如一幅连续图像  $f(x, y)$  被取样，则产生的数字图像有  $M$  行和  $N$  列。坐标  $(x, y)$  的值变成离散值，通常将这些离散坐标用整数表示(见图 1-2)。

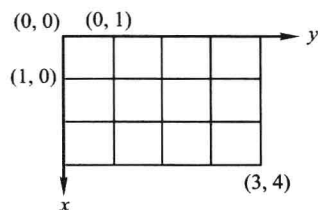


图 1-2 图像的坐标

当然，图像数字化的过程与采样间隔(频率)有关。采样频率太低，则图像会产生混叠现象而失真。另外，灰度或颜色数目的选取也与重建后的图像质量有关。一般地，数量级在 8 位(256 色)的黑白图像是比较好的，而高质量的彩色图像需要使用三基色叠加的办法。

模拟图像数字化后，二维矩阵的每一个位置称为像素(Pixel)，每个像素都包括位置信息和灰度。对于单色即灰度图像而言，每个像素的亮度用一个数值来表示，通常数值范围为 0~255，即可用一个字节来表示，0 表示黑，255 表示白，而其它数值表示灰度级别，如图 1-3 所示。

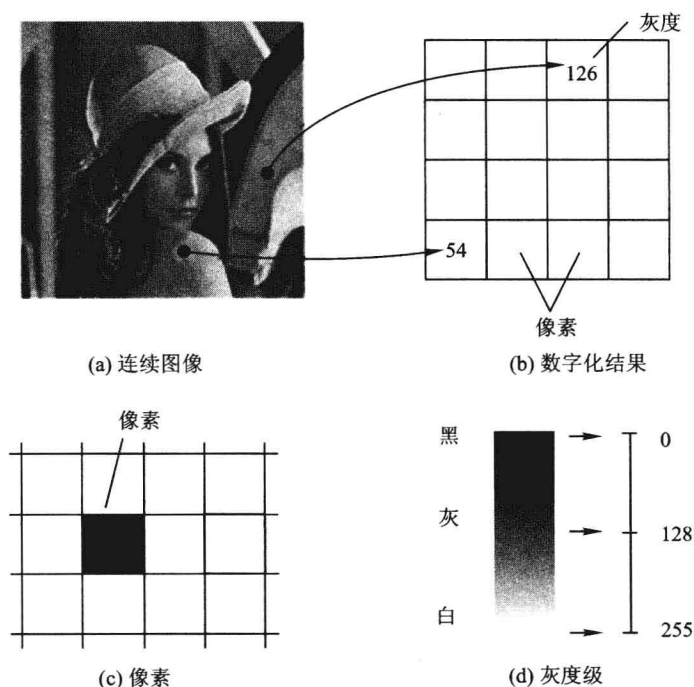


图 1-3 图像的数字化过程

### 1. 采样

采样(Sampling)就是对图像空间坐标的离散化，它决定了图像的空间分辨率(图像中

可分辨的最小细节,主要由采样间隔值决定)。用一个网格把待处理的图像覆盖,然后把每一小格上模拟图像的各个亮度取平均值,作为该小方格的值;或者把方格的交叉点处模拟图像的亮度值作为该方格交叉点上的值,如图 1-4 所示。

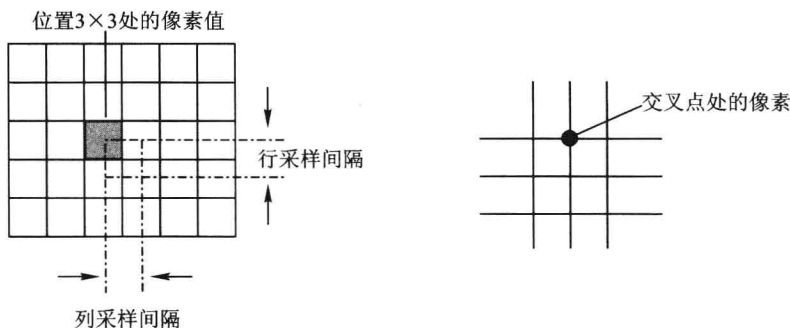


图 1-4 图像采样

图像平面分割成离散点的集合,结果是一个样点值阵列,故又叫点阵取样。每个离散点即像素,可用离散坐标  $(i, j)$  表示。点阵采样的数学描述为

$$S(x, y) = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta(x - i\Delta x, y - j\Delta y) \quad (1-6)$$

$$\begin{aligned} f_p(x, y) &= f(x, y) \cdot S(x, y) = f(x, y) \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \delta(x - i\Delta x, y - j\Delta y) \\ &= \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} f(i\Delta x, j\Delta y) \cdot \delta(x - i\Delta x, y - j\Delta y) \end{aligned} \quad (1-7)$$

上式对二维离散函数  $f(x, y)$  进行均匀采样,  $S(x, y)$  为二维离散采样函数,  $\Delta x$  和  $\Delta y$  为相应方向上的采样间隔,  $f_p(x, y)$  为采样函数。

采样函数的频谱经过傅立叶变换可以表示为

$$F_p\{f_p(x, y)\} = \frac{1}{\Delta x} \frac{1}{\Delta y} \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} F\left(u - i \frac{1}{\Delta x}, v - j \frac{1}{\Delta y}\right) \quad (1-8)$$

即采样函数的频谱是原函数频谱的周期拓展,  $\Delta u = \frac{1}{\Delta x}$  和  $\Delta v = \frac{1}{\Delta y}$  是其拓展的周期。

采样函数的频谱决定了图像的空间分辨率。空间分辨率是图像中可分辨的最小细节,主要由采样间隔值决定。同一个景物,对其进行采样的空间分辨率越高,采样间隔就越小,数字化后的图像就越能反映出景物的细节,图像的质量也就越高。一幅用二维数组  $f(M, N)$  表示的数字图像,像素总数  $M \times N$  称为图像的空间分辨率。同一个物体,采样的空间分辨率越高,获得的图像矩阵  $M \times N$  也就越大,反之则越小。

所以,要使数字图像保有一定的空间分辨率,从而不丢失原连续图像所表达的信息,同时能确保不失真或有限失真地恢复原图像,图像采样间隔就必须依一定的规则选取合适的值。这个规则就是二维采样定理。

由二维采样定理可知,如果二维信号  $f(x, y)$  的二维傅立叶频谱  $F(u, v)$  满足

$$F(u, v) = \begin{cases} F(u, v) & |u| < U_c, |v| \leq V_c \\ 0 & |u| > U_c, |v| > V_c \end{cases} \quad (1-9)$$