

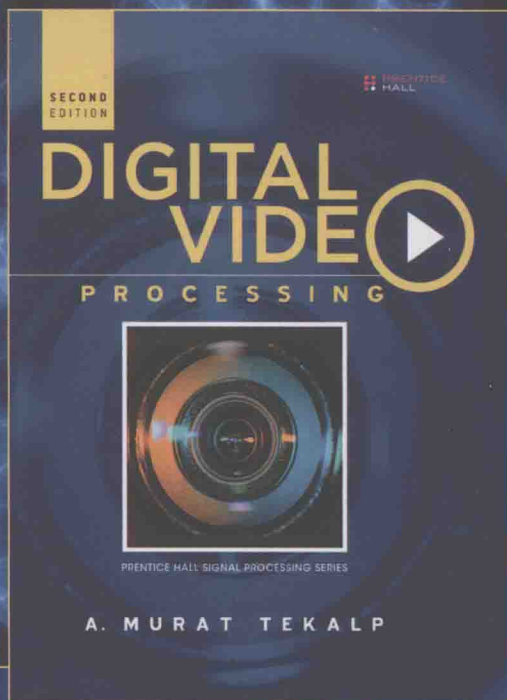
数字视频处理

(原书第2版)

[土耳其] A. 缪拉·泰卡尔普 (A. Murat Tekalp) 著

曹铁勇 张雄伟 杨吉斌 黄辉 李莉 译

*Digital Video
Processing
Second Edition*

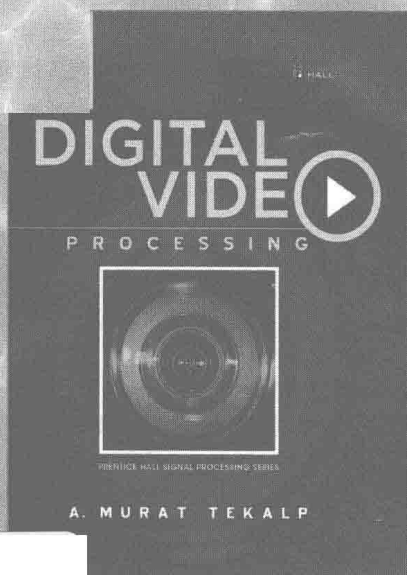


数字视频处理

(原书第2版)

[土耳其] A. 缪拉·泰卡尔普 (A. Murat Tekalp) 著
曹铁勇 张雄伟 杨吉斌 黄辉 李莉 译

*Digital Video
Processing
Second Edition*



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

数字视频处理 (原书第 2 版)/(土) A. 缪拉·泰卡尔普 (A. Murat Tekalp) 著; 曹铁勇等译. —北京: 机械工业出版社, 2017. 9

(国外电子与电气工程技术丛书)

书名原文: Digital Video Processing, Second Edition

ISBN 978-7-111-58042-3

I. 数… II. ① A… ② 曹… III. 数字视频系统—数字信号处理 IV. TN941.3

中国版本图书馆 CIP 数据核字 (2017) 第 231157 号

本书版权登记号: 图字 01-2015-5217

Authorized translation from the English language edition, entitled Digital Video Processing, Second Edition, 9780133991000 by A. Murat Tekalp, published by Pearson Education, Inc., Copyright © 2015.

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from Pearson Education, Inc.

CHINESE SIMPLIFIED language edition published by CHINA MACHINE PRESS / HUAZHANG CO., Copyright © 2018.

本书中文简体字版由 Pearson Education (培生教育出版集团) 授权机械工业出版社在中华人民共和国境内 (不包括台湾地区和香港、澳门特别行政区) 独家出版发行。未经出版者书面许可, 不得以任何方式抄袭、复制或节录本书中的任何部分。

本书封底贴有 Pearson Education (培生教育出版集团) 激光防伪标签, 无标签者不得销售。

本书内容丰富, 组织严谨, 全面覆盖了图像滤波、运动估计、视频分割与跟踪、视频滤波、图像/视频压缩等方面的基本原理和最新成就。然而, 一本书无法覆盖数字视频处理和计算机视觉领域所有的最新成就, 因此本书只对最基本、最常用的技术和算法加以详解, 而对更多的先进算法和最新研究成果只进行简介, 并提供用于自学的参考文献。大部分章节最后都设有习题集 (部分习题集中还含有 MATLAB 练习), 以便读者对所学到的方法进行练习。

出版发行: 机械工业出版社 (北京市西城区百万庄大街 22 号 邮政编码: 100037)

责任编辑: 张梦玲

责任校对: 殷虹

印刷: 北京瑞德印刷有限公司

版次: 2018 年 1 月第 1 版第 1 次印刷

开本: 185mm×260mm 1/16

印张: 19.75

书号: ISBN 978-7-111-58042-3

定价: 89.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

客服热线: (010) 88378991 88361066

投稿热线: (010) 88379604

购书热线: (010) 68326294 88379649 68995259

读者信箱: hzjsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问: 北京大成律师事务所 韩光/邹晓东

出版者的话

文艺复兴以来,源远流长的科学精神和逐步形成的学术规范,使西方国家在自然科学的各个领域取得了垄断性的优势;也正是这样的传统,使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中,美国的产业界与教育界越来越紧密地结合,信息学科中的许多泰山北斗同时身处科研和教学的最前线,由此而产生的经典科学著作,不仅肇划了研究的范畴,还揭示了学术的源变,既遵循学术规范,又自有学者个性,其价值并不会因年月的流逝而减退。

近年,在全球信息化大潮的推动下,我国的信息产业发展迅猛,对专业人才的需求日益迫切。这对我国教育界和出版界都既是机遇,也是挑战;而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下,美国等发达国家在其信息科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此,引进一批国外优秀教材将对我国教育事业的发展起到积极的推动作用,也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自1998年开始,我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力,我们与Pearson、McGraw-Hill、Elsevier、John Wiley & Sons、CRC、Springer等世界著名出版公司建立了良好的合作关系,从他们现有的数百种教材中甄选出Alan V. Oppenheim、Thomas L. Floyd、Charles K. Alexander、Behzad Razavi、John G. Proakis、Stephen Brown、Allan R. Hambley、Albert Malvino、Peter Wilson、H. Vincent Poor、Hassan K. Khalil、Gene F. Franklin、Rex Miller等大师名家的经典教材,以“国外电子与电气技术丛书”和“国外工业控制与智能制造丛书”为系列出版,供读者学习、研究及珍藏。这些书籍在读者中树立了良好的口碑,并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也越来越多被实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑,这些因素使我们的图书有了质量的保证。随着电气与电子信息学科、自动化、人工智能等建设的不断完善和教材改革的逐渐深化,教育界对国外电气与电子信息类、控制类、智能制造类等相关教材的需求和应用都将步入一个新的阶段,我们的目标是尽善尽美,而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正,我们的联系方式如下:

华章网站: www.hzbook.com

电子邮件: hzsj@hzbook.com

联系电话: (010)88379604

联系地址: 北京市西城区百万庄南街1号

邮政编码: 100037



华章科技图书出版中心

译者序

数字图像/视频处理是数字信号处理领域的一个重要方向。自 20 世纪以来,随着计算机技术、电子技术和数学的发展,人们在数字图像/视频处理的增强、复原、检测、分割、识别、跟踪、压缩等各方面都取得了长足的进展,其成果已广泛应用于人们的日常生活、工作学习、军事建设等各个方面。由于其技术应用已深入到工业、农业、矿产、海洋、军事、娱乐等各领域,所以世界上大多数理工科院校都开设了数字图像处理或相关课程,相关专著和教材也已数以千计。但在其中,大多数相关教材都将重点放在数字图像处理、图像分析、图像理解上,以数字视频处理为重点内容的相关教材和专著相对较少。

数字视频处理是与数字图像处理紧密相关的。但与数字图像处理相比,由于输入信号中包含了时间维度的信息,因此数字视频处理具有许多独特的理论与技术,如运动估计、视频分割、3D 场景分析等。我们可以将视频看成静态图像序列,并逐帧独立处理,但若采用基于帧间相关性的多帧联合处理技术,却能够开发出更有效的视频处理算法,例如运动补偿滤波和预测等。此外,运动估计或动态场景分析等任务显然是无法基于单个图像来实现的。

土耳其的 A. Murat Tekalp 教授是数字图像与视频处理领域的资深专家,长期从事数字图像与视频处理方面的本科和研究生教学工作,具有丰富的教学和科研经验。本书是其近 20 多年来在数字图像与视频处理领域的教学结晶,全面覆盖了图像滤波、运动估计、视频分割与跟踪、视频滤波、图像/视频压缩等方面的基本原理和最新成就。全书的主要特点是从问题描述入手,按照模型建立、条件约束、求解实现的顺序,建立了较为完整的问题求解框架,从而有效拓展读者对相关问题的认识深度。为此,作者还在开始部分加入了多维信号与系统的介绍,在附录部分提供了问题求解所涉及的相关矩阵、优化等知识的介绍。对于具体的算法实现过程步骤,本书着墨不多。因此,本书更适合相关专业的研究生作为教材使用。

本书的翻译由曹铁勇、张雄伟、杨吉斌、黄辉、李莉共同完成,其中,前言、1.1 节、1.2.1 节、第 3 章、第 5 章、第 6 章由曹铁勇翻译,第 1 章其余部分、第 2 章由张雄伟翻译,第 8 章和附录部分由杨吉斌翻译,第 4 章由黄辉翻译,第 7 章由李莉翻译。全书由曹铁勇统稿。在翻译过程中,还得到了郑云飞、王文、孙蒙、赵斐、吴其前等的帮助。

在本书翻译过程中,我们力求忠于原著,但由于水平有限,难免会出现错误或遗漏之处,恳请广大读者批评指正。

译者

2017 年 6 月

于南京

前 言

本书于 1995 年出版了第 1 版,是第一本全面介绍数字视频处理的教材,当时根据视频处理领域的重要论题分成了 25 章,在一个学期的课程中,每章可以用一到两次课进行讲授。那个时期的数字视频技术和视频处理算法还不太成熟,数码摄像机和 DVD 刚刚商业化,数字电视标准正在制定,而数字电影还没有纳入议程。因此,与当今的技术水平相比,第 1 版中的一些方法/算法和技术已经过时,比如像素级回归的运动估计、矢量量化、不规则形状压缩、基于模型的编码等已不再先进,还有一些诸如模拟视频/电视和 128K 可视电话等技术则已经淘汰了。同时近 20 年来的重大进展显然也无法体现出来。

第 1 版出版至今已有 20 多年了,在现今这个数字化时代,数字视频已广泛应用于我们的日常生活。信号处理与计算机视觉领域的重大发展促进了视频处理算法的不断成熟,能够应用于不同用途的最常用又有效的算法与技术也更加清晰。因此,现在是本书推出新版的最好时机。本书围绕图像与视频处理的最新发展进行了精心编排,力图成为一本内容全面、结构严谨的教材。

第 2 版大幅度改进了内容与表述风格的组织方式,包含当今最先进的技术、最有效的算法和最新的知识。全书共分 8 章,每章对应一个主题,分别是多维信号处理、图像/视频基础、图像滤波、运动估计、视频分割、视频滤波、图像压缩、视频压缩。每个主题侧重介绍最有效的技术。与第 1 版相比,这一版不是简单的内容增补,而是一次全新的改写。

本书可作为高年级本科生或研究生的数字图像与视频处理课程的教材,要求读者预先掌握微积分、线性代数、概率论和一些基本的数字信号处理概念。具有计算机科学背景但不熟悉信号处理基本概念的读者,则可以跳过第 1 章,从第 2 章开始学习。尽管本书表述严谨,但仍然像一般教材一样从原理开始讲起,因此也可以用作产业界或学术界的工程师和研究人员自学的参考书。本书有助于读者理解图像和视频处理方法的理论基础,学习用最通用、最有效的算法解决常见的图像与视频处理问题。大部分章后提供习题集(部分习题集中还含有 MATLAB 练习),通过这些可加深读者对知识的理解和对方法的掌握。

数字视频处理就是对数字视频比特流的各种操作。所有数字视频应用都离不开压缩。此外,为了获得高质量图像或提取特定信息,数字视频应用也离不开广泛应用于格式转换、增强、复原、高分辨率重建等场合的滤波处理;有些应用还需要用到其他的处理,以实现运动估计、视频分割和 3D 场景分析。视频的帧与帧之间存在着大量的时间相关性(冗余),这使得视频处理不同于静态图像处理。可以将视频看成静态图像序列,并逐帧独立处理。但若采用基于帧间相关性的多帧联合处理技术,我们能够开发出更有效的视频处理算法,例如运动补偿滤波和预测。此外,某些任务(比如运动估计或动态场景分析)显然是无法基于单个图像来进行的。

本书的目的是为读者提供图像(单帧)和视频(多帧)处理方法的数学基础。特别是,本书还回答了以下基本问题:

- 如何从噪声中分离出图像(信号)?
- 内插、复原和超分辨率重建之间是否有内在的联系?
- 对于不同的应用,该如何估计 2D 和 3D 运动?
- 如何将图像和视频分割成感兴趣区域?
- 如何跟踪视频中的对象?
- 与图像滤波相比,视频滤波问题是否更趋向于适应?

- 超分辨率重建为何能够实现?
- 能否从视频片段中得到高质量的静态图像?
- 图像和视频压缩为什么能够实现?
- 如何压缩图像和视频?
- 图像/视频压缩的最新国际标准是什么?
- 3D 视频表示和压缩的最新标准是什么?

图像和视频处理的大都是病态问题(欠定的和/或对噪声敏感的),并且它们的解都依赖于某些图像和视频模型。附录 B 讨论了用于病态问题解的图像建模方法。实际上,图像模型可以分成基于局部平滑、基于变换域稀疏和基于非局部自相似的种类。

图像处理算法大都使用了以上模型中的一种或多种。此外,视频模型还包括基于全局平移或块平移、参数化运动、运动(空间上)的平滑性、时域运动单调性(时域连续或平滑)、3D 空-时频域的平面支撑等类。

各章概述如下。

第 1 章回顾了多维信号、变换和系统的基础知识,它们是一些图像和视频处理方法的理论基础。该章还介绍了格上空-时采样的体制(如逐行和隔行采样),以及采样格式转换理论。读者如果具有计算机科学背景但不熟悉信号处理概念,则可以跳过本章,直接从第 2 章开始学习。

第 2 章给出了数字图像与视频的基础知识。主要内容包括人类视觉、空间频率、彩色模型、模拟和数字视频表示、数字视频标准、3D 立体/多视角视频表示、数字视频质量评估等基本概念,以及一些常见的数字视频应用,如数字电视、数字电影和网络视频流等。

第 3 章介绍了图像(静止帧)滤波类问题,比如图像重采样(抽取与内插)、梯度与边缘检测、增强、去噪、复原等。还介绍了线性移不变滤波器、自适应滤波器和非线性滤波器。附录 B 给出了求解病态逆问题的一般性框架。

第 4 章介绍了 2D 和 3D 运动估计方法。运动估计是数字视频处理的核心,因为运动是视频的显著特征,并且运动补偿滤波是利用时间冗余的最有效方法。再者,许多计算机视觉工作的第一步都是 2D 或 3D 运动估计与跟踪。2D 运动估计一般分为稠密光流或稀疏对应估计两类,可以基于参数法和非参数法来实现。非参数法包括基于图像梯度的光流估计法、块匹配法、像素递归法、贝叶斯法和相位相关法。基于仿射模型或单应性的参数法可以用于图像配准或局部变形估计。3D 运动/结构估计法一般都基于双帧极线约束法(主要针对立体对)或多帧因子分解法。欧几里得 3D 结构重建需要对所有相机进行标定,而投影重建法则无需标定。

第 5 章介绍了图像分割和变化检测,以及基于参数聚类法和贝叶斯法的主要运动或复杂运动分割。我们还讨论了运动估计与分割的同时实现问题。因为双视角运动估计技术对于图像梯度或对应点的估计精度很敏感,因此我们也对长的单目序列或立体像对中分割对象的运动跟踪进行了讨论,其结果更鲁棒。

第 6 章介绍了视频滤波,包括标准转换、去噪和超分辨率重建等内容。首先介绍了运动补偿滤波的基本原理,随后介绍了标准转换问题,包括帧速转换和去隔行等。视频帧的画面中经常存在颗粒,尤其在静止帧模式下观看时更加严重。为此,讨论了用于噪声抑制的运动自适应和运动补偿滤波。最后介绍了一种统一各种视频滤波问题的综合模型,可用于低分辨率视频获取和超分辨率重建。

第 7 章介绍了包括二值图像(传真)和灰度图像在内的静态图像压缩方法与标准,如 JPEG 和 JPEG 2000 等。还特别讨论了无损的图像压缩、有损的离散余弦变换(DCT)编码和小波变换编码等方法。

第 8 章讨论了视频压缩方法和标准,它们是实现数字电视、数字电影等数字视频应用的基础。在简要介绍了视频压缩的不同方法后,详细描述了 MPEG-2、AVC/H.264 和

HEVC(高效视频编码)等标准, 以及这些标准在可伸缩视频编码和立体和多视角视频编码方面的扩展。

本书是近 20 多年来我在数字图像与视频处理领域的教学结晶。本书内容丰富, 结构合理, 全面覆盖了图像滤波、运动估计、视频分割与跟踪、视频滤波、图像/视频压缩等方面的基本原理和最新成就。然而, 一本书无法覆盖数字视频处理和计算机视觉领域所有的最新成就, 因此本书只对最基本、最常用的技术和算法加以详解, 而对更多的先进算法和最新研究成果只进行简介, 并提供用于自学的参考文献。大部分章节最后都包含习题集(部分习题集还包含 MATLAB 练习), 以便读者对所学到的方法进行练习。

教师可以通过申请获得教学资料。表 P.1 提供了一个教学建议, 按照一个学期 14 周、每周两次课、每次课 75 分钟的方式进行组织, 可在一个学期的数字图像与视频处理课程中讲完本书的全部内容。另一种方式是将本书内容分到两个学期中, 这样有更多时间对每个主题的细节进行探讨: 第一学期可以开设数字图像处理课程, 讲解第 1、2、3、7 章中的内容; 第二学期开设数字视频处理课程, 讲解第 4、5、6、8 章的内容。

表 P.1 一学期课程的教学建议

每次课的编号	专 题	章/节
1	2D 信号与 2D 变换	1.1 节、1.2 节
2	2D 系统、2D FIR 滤波器、频率响应	1.3 节
3	格上空-时采样	1.4 节、1.5 节
4	数字图像/视频、人类视觉、视频质量	第 2 章
5	矢量-矩阵表示、图像模型、图像/视频处理中病态问题的表达	附录 A、附录 B
6	抽取、内插、多分辨率金字塔	3.2 节
7	梯度估计、边缘/角检测	3.3 节
8	图像增强、点操作、钝化掩模、双边滤波	3.1 节、3.4 节
9	噪声滤波: LSI 滤波器(自适应、非线性、非局部滤波器)	3.5 节
10	图像复原: 迭代法、POCS	3.6 节
11	运动建模、光流法、对应法	4.1 节、4.2 节、4.3 节
12	不同的方法: Lukas-Kanade、参数模型	4.4 节
13	块匹配、参数模型估计中的特征匹配、相位相关法	4.5 节、4.7 节
14	3D 运动估计、对极几何	4.8 节
15	变化检测、视频分割	5.2 节、5.3 节
16	运动跟踪	5.4 节、5.5 节
17	运动补偿滤波、多帧联合去隔行、去噪	6.1 节、6.2 节、6.3 节
18	超分辨率重建	6.5 节
19	数据/图像压缩介绍、信息论概念、熵编码、算术编码	7.1 节
20	无损位平面编码、G3/G4、JBIG 标准	7.2 节
21	预测数据编码、JPEG 标准	7.2 节
22	DCT 与 JPEG 图像压缩	7.3 节
23	小波变换、JPEG 2000 图像压缩	7.4 节
24	MC-DCT、MPEG-1、MPEG-2	8.1 节、8.2 节
25	MPEG-4 AVC/H.264 标准	8.3 节
26	HEVC	8.4 节
27	可分级视频编码、DASH 自适应流、抗误码	8.5 节
28	3D/立体和多视角视频压缩	8.6 节

显然，本书是信号处理和计算机科学相关组织研究成果的荟萃。每章都有很多引用并列出了相关参考文献，但肯定无法涵盖图像和视频领域科研与工业部门杰出研究者的所有成就。此外，对于 ISO 和 ITU 组织中各位科学家经多年工作取得的图像与视频编码的显著成果，这里也难以一一致意。

最后，衷心感谢 Xin Li (美国西弗吉尼亚大学, WVU)、Eli Saber、Moncef Gabbouj、Janusz Konrad 和 H. Joel Trussell 在本书成稿过程中的贡献。同时感谢 Prentice Hall 出版社的 Bernard Goodwin、Kim Boedigheimer 和 Julie Nahil 的帮助与支持。

A. Murat Tekalp

于土耳其伊斯坦布尔 Koc 大学

出版者的话
译者序
前言

第 1 章 多维信号与系统	1
1.1 多维信号	1
1.1.1 有限域信号和周期信号	1
1.1.2 对称信号	3
1.1.3 特殊的多维信号	3
1.2 多维变换	5
1.2.1 连续信号的傅里叶变换	5
1.2.2 离散信号的傅里叶变换	7
1.2.3 离散傅里叶变换	8
1.2.4 离散余弦变换	10
1.3 多维系统	11
1.3.1 冲激响应和 2D 卷积	12
1.3.2 频率响应	13
1.3.3 FIR 滤波器和对称性	14
1.3.4 IIR 滤波器和偏微分方程	15
1.4 多维采样理论	16
1.4.1 格上采样	17
1.4.2 格上采样信号的频谱	19
1.4.3 格上采样的奈奎斯特准则	19
1.4.4 基于格上采样点的重构	22
1.5 采样结构转换	23
参考文献	26
习题集	26
第 2 章 数字图像与数字视频	28
2.1 人类视觉系统和色彩	28
2.1.1 色觉与模型	28
2.1.2 对比灵敏度	30
2.1.3 空-时频率响应	31
2.1.4 立体/深度感知	33
2.2 模拟视频	34
2.2.1 逐行扫描和隔行扫描	34
2.2.2 模拟视频信号格式	35
2.2.3 模拟/数字转换	35
2.3 数字视频	36
2.3.1 空间分辨率和帧率	36
2.3.2 色彩、动态范围和位深	37

2.3.3 彩色图像处理	38
2.3.4 数字视频标准	40
2.4 3D 视频	42
2.4.1 3D 显示技术	43
2.4.2 立体视频	44
2.4.3 多视角视频	45
2.5 数字视频应用	46
2.5.1 数字电视	46
2.5.2 数字电影	48
2.5.3 互联网视频流	49
2.5.4 计算机视觉和场景/活动理解	51
2.6 图像与视频质量	51
2.6.1 可视失真	52
2.6.2 主观质量评价	52
2.6.3 客观质量评价	53
参考文献	54
第 3 章 图像滤波	57
3.1 图像平滑	57
3.1.1 线性移不变低通滤波	57
3.1.2 双边滤波	59
3.2 图像重采样和多分辨率表示	60
3.2.1 图像抽取	60
3.2.2 内插	61
3.2.3 多分辨率金字塔表示	65
3.2.4 小波表示	65
3.3 图像梯度估计、边缘和特征检测	68
3.3.1 图像梯度的估计	69
3.3.2 拉普拉斯估计	71
3.3.3 Canny 边缘检测	72
3.3.4 Harris 角检测	73
3.4 图像增强	74
3.4.1 基于像素的对比度增强	74
3.4.2 用于色调映射和图像锐化的空间滤波	76
3.5 图像去噪	79
3.5.1 图像与噪声模型	80
3.5.2 DFT 域的线性空不变滤波器	81

3.5.3 局部自适应滤波	83	4.6.2 贝叶斯运动估计	131
3.5.4 非线性滤波: 顺序统计、小波 收缩和双边滤波	85	4.7 变换域法	132
3.5.5 非局部滤波: NL 均值和 BM3D	87	4.7.1 相位相关法	132
3.6 图像复原	88	4.7.2 空域频谱法	132
3.6.1 模糊模型	89	4.8 3D 运动与结构估计	133
3.6.2 线性空不变模糊退化图像的 复原	91	4.8.1 相机标定	133
3.6.3 盲复原——模糊识别	94	4.8.2 仿射重建	134
3.6.4 空变模糊退化图像的复原	95	4.8.3 投影重建	134
3.6.5 图像修补	97	4.8.4 欧几里得重建	137
参考文献	97	4.8.5 平面视差和相对仿射结构 重建	138
习题集	101	4.8.6 基于立体视觉的稠密结构	139
MATLAB 资源	104	参考文献	139
第 4 章 运动估计	105	习题集	143
4.1 成像	105	MATLAB 资源	144
4.1.1 相机模型	105	第 5 章 视频分割与跟踪	145
4.1.2 3D 运动的光度效应	108	5.1 图像分割	146
4.2 运动模型	108	5.1.1 阈值法	146
4.2.1 投影运动和表观运动	109	5.1.2 聚类	147
4.2.2 3D 刚体运动投影模型	111	5.1.3 贝叶斯法	148
4.2.3 2D 表观运动模型	112	5.1.4 图方法	151
4.3 2D 表观运动估计	114	5.1.5 主动轮廓模型	152
4.3.1 稀疏对应估计、光流估计与 图像配准问题	114	5.2 变化检测	153
4.3.2 光流方程和法向流	115	5.2.1 镜头边界检测	153
4.3.3 位移帧差	116	5.2.2 背景相减	154
4.3.4 运动估计是病态问题: 遮挡和 孔径问题	117	5.3 运动分割	157
4.3.5 分层运动估计	118	5.3.1 主要运动分割	158
4.3.6 运动估计的性能评价	119	5.3.2 多运动分割	159
4.4 差分法	119	5.3.3 基于区域的运动分割: 色彩 与运动融合	164
4.4.1 Lucas-Kanade 法	119	5.3.4 同步进行运动估计与分割	165
4.4.2 Horn-Schunk 运动估计	121	5.4 运动跟踪	167
4.5 匹配算法	123	5.4.1 基于图的空-时分割与跟踪	168
4.5.1 基本的块匹配	124	5.4.2 KLT 跟踪	168
4.5.2 可变大小块匹配	126	5.4.3 MS 跟踪	169
4.5.3 分层块匹配	127	5.4.4 粒子滤波跟踪	170
4.5.4 广义块匹配——局部变形 运动	127	5.4.5 主动轮廓跟踪	171
4.5.5 来自特征对应的单应性 估计	128	5.4.6 2D-mesh 跟踪	173
4.6 非线性优化法	130	5.5 图像和视频抠图	173
4.6.1 像素递归运动估计	130	5.6 性能评估	174
		参考文献	175
		习题集	181
		网络资源	181
		第 6 章 视频滤波	182
		6.1 空-时滤波原理	182

6.1.1 视频的频谱	182	第 8 章 视频压缩	249
6.1.2 运动自适应滤波	184	8.1 视频压缩方法	249
6.1.3 运动补偿滤波	184	8.1.1 帧内压缩、运动 JPEG 2000 和 数字电影	249
6.2 视频格式转换	186	8.1.2 3D 变换编码	250
6.2.1 下变频	187	8.1.3 运动补偿变换编码	251
6.2.2 去隔行	189	8.2 早期视频压缩标准	252
6.2.3 帧率转换	193	8.2.1 ISO 和 ITU 标准	252
6.3 多帧的噪声滤除	196	8.2.2 MPEG-1 标准	252
6.3.1 运动自适应噪声滤除	196	8.2.3 MPEG-2 标准	257
6.3.2 运动补偿噪声滤除	197	8.3 MPEG-4 AVC/ITU-T H.264 标准	261
6.4 多帧复原	200	8.3.1 输入视频格式和数据结构	261
6.4.1 多帧建模	200	8.3.2 帧内预测	262
6.4.2 多帧维纳复原	200	8.3.3 运动补偿	263
6.5 多帧超分辨率	202	8.3.4 变换	264
6.5.1 超分辨率	202	8.3.5 其他工具和改进	265
6.5.2 低分辨率采样建模	204	8.4 高效视频编码标准	266
6.5.3 频域超分辨率	206	8.4.1 视频输入格式和数据结构	266
6.5.4 多帧空域法	208	8.4.2 编码树单元	266
参考文献	211	8.4.3 并行编/解码工具	267
习题集	214	8.4.4 其他工具和改进	268
第 7 章 图像压缩	216	8.5 可分级视频压缩	270
7.1 图像压缩基础	216	8.5.1 时域可分级	270
7.1.1 信息论概念	216	8.5.2 空域可分级	271
7.1.2 图像压缩系统的组成	217	8.5.3 质量可分级	271
7.1.3 量化	218	8.5.4 混合可分级	272
7.1.4 符号编码	220	8.6 立体和多视角视频压缩	272
7.1.5 赫夫曼编码	220	8.6.1 帧兼容立体视频压缩	273
7.1.6 算术编码	222	8.6.2 H.264/AVC 标准的立体和 多视角视频编码扩展	274
7.2 无损图像压缩	224	8.6.3 带深度信息的多视角视频	275
7.2.1 位平面编码	224	参考文献	278
7.2.2 RLC 和 ITU G3/G4 标准	225	习题集	280
7.2.3 自适应算术编码和 JBIG	228	网络资源	280
7.2.4 早期无损预测编码	228	附录 A 图像和视频处理中的矢量和 矩阵运算	281
7.2.5 JPEG-LS 标准	229	附录 B 图像和视频处理中的 病态问题	286
7.2.6 Lempel-Ziv 编码	231	附录 C 马尔可夫随机场和吉布斯随 机场	289
7.3 离散余弦变换编码和 JPEG	232	附录 D 优化方法	293
7.3.1 离散余弦变换	233	附录 E 模型拟合	298
7.3.2 ISO JPEG 标准	234		
7.3.3 编码器控制和压缩失真	238		
7.4 小波变换编码和 JPEG 2000	239		
7.4.1 小波变换和滤波器的选择	239		
7.4.2 ISO JPEG 2000 标准	242		
参考文献	245		
习题集	247		
网络资源	248		

一维信号处理与多维信号处理之间存在着一些根本性的不同,因为多维 z 变换的零点和极点不是孤立点而是函数,且多变量多项式通常无法进行因式分解。

对于数字图像和视频处理来说,多维(Multi-Dimensional, MD)信号与系统理论是基础。数字图像是关于两个离散空间变量的二维(2D)序列,也可以表示成数组(矢量或矩阵)形式;数字视频是关于两个空间变量和一个时间变量的三维(3D)函数;3D视频则是关于三个空间变量和一个时间变量的四维(4D)函数。用于处理这些信号的数字滤波器是MD系统。MD变换有助于人们理解空间域频率和时间域频率的概念,也有助于理解物理意义上的频率和归一化频率。本章介绍数字图像与视频处理的基础,包括MD信号、变换和系统的基本概念,以及MD模拟信号的格上采样和采样格式转换等。

1D信号处理和MD信号处理之间存在着一些根本性的不同,这是因为:(1)MD z 变换的零点和极点不是孤立点而是函数;(2)多变量多项式通常无法进行因式分解。因此,部分1D信号的处理方法是无法简单地推广到MD信号处理的;但同时,MD信号处理中的新算法也可能无法对应地用到1D信号处理中。本章将以2D信号处理为例,给出MD信号、系统和变换的主要结论,这些结论能够简单地推广到更高维,只不过表示更加烦琐和抽象,也无法形象化地呈现出来。

1.1 多维信号

MD信号是指具有 $M \geq 2$ 个独立的连续、离散或混合变量的多变量函数或序列。

定义:模拟MD信号 $s(\mathbf{x}) = s(x_1, x_2, \dots, x_M)$ 是具有 M 个连续变量的函数,其中 $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_M]^T$ 。函数 $s(\mathbf{x})$ 可以是标量形式(如灰度图像),也可以是矢量形式(如彩色图像)。

定义:离散MD信号 $s(\mathbf{n}) = s(n_1, n_2, \dots, n_M)$ 是定义于格上的MD序列,是关于 M 元整型矢量 $\mathbf{n} = [n_1 \ n_2 \ \dots \ n_M]^T$ 的偏序集。同样, $s(\mathbf{n})$ 可以是标量形式或矢量形式。

定义:混合MD信号是具有 M 个变量的函数,其中既有连续变量,也有离散变量。模拟视频信号就是一种混合信号。

1.1.1 有限域信号和周期信号

本节讨论MD有限域信号和MD周期信号,两者之间是同构的。

有限域信号

照相机的传感器尺寸是有限的,视频记录时间是有限的,我们感兴趣的许多MD信号都是有限的。换句话说,它们定义于有限的空间/时间域内,我们称这个域为“支撑”。信号的支撑可能是MD矩形,也可能是任意形状。

如果一个2D信号定义于从原点出发的两条直线之间,那么就称这个信号具有楔形支撑,如图1.1d所示。1/4平面支撑(见图1.1a)、半平面支撑(见图1.1b)和非对称半面支撑(Non-Symmetric Half-Plane, NSHP)(见图1.1c)都是楔形支撑的特例。

例 1.1 数字图像具有有限的1/4平面支撑,如图1.1a所示,即定义于一个矩形区域 $0 \leq n_1 \leq N_1 - 1, 0 \leq n_2 \leq N_2 - 1$ 内,其中 N_1 和 N_2 分别代表图像在水平方向和垂直方向的

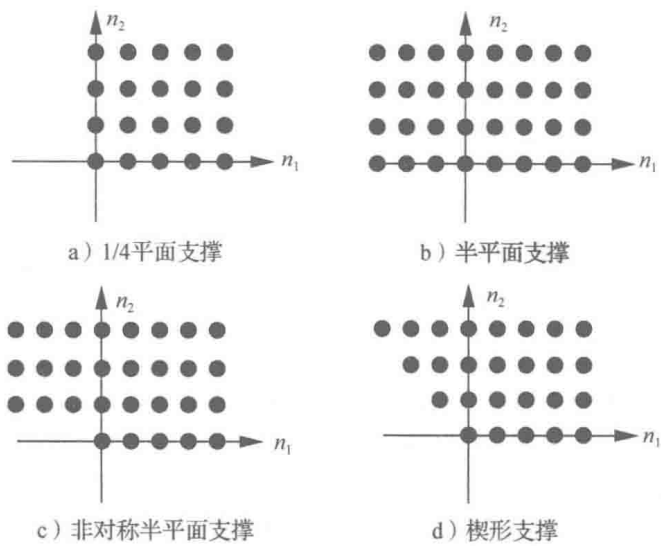


图 1.1 2D 信号的支撑

像素数。

周期信号

对于 MD 序列 $\hat{s}(n)$ ，其中 $n=[n_1 \ n_2 \ \cdots \ n_M]^T$ ，若其对于所有的 M 元整型矢量 r ，均满足下式，则称其为是周期的，对应周期矩阵为 N 。

$$\hat{s}(n) = \hat{s}(n + Nr) \tag{1.1a}$$

其中， N 为 $M \times M$ 大小的周期矩阵，且 $\det(N) \neq 0$ 。

对于 2D 信号 $r=[r_1 \ r_2]^T$ ，且 $N=[n_1 \ | \ n_2]=\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}$ ，矢量方程式(1.1a)可以写成如下标量形式：

$$s(n_1, n_2) = \begin{cases} s(n_1 + r_1 N_{11}, n_2 + r_1 N_{22}), r_2 = 0 \\ s(n_1 + r_2 N_{12}, n_2 + r_2 N_{22}), r_1 = 0 \end{cases} \tag{1.1b}$$

整型矢量 $n_1^\ominus=[N_{11} \ N_{21}]^T$ 和 $n_2=[N_{12} \ N_{22}]^T$ 表示的是某个周期中样点与其他周期中对应样点间的替代关系。换句话说，2D 信号在位移矢量 n_1 和 n_2 的整数倍位置上不断复制自身，如图 1.2 所示。周期定义了 (n_1, n_2) 面内的一个单位域(unit cell)，这个单位域在整个平面上不断重复。需要注意的是， N 不是唯一的：任意两个线性独立矢量 n_1 和 n_2 只要在不同周期内能够指向相同的样点，都能表示相同的周期模式。

周期矩阵为对角矩阵 $N=\begin{bmatrix} N_1 & 0 \\ 0 & N_2 \end{bmatrix}$ 的情况是一个特例，如图 1.2b 所示。此时位移矢量 n_1 和 n_2 分别是沿着水平轴和垂直轴的。这种情况下，认为 $s(n_1, n_2)$ 具有矩形周期，并满足：

$$s(n_1, n_2) = s(n_1 + N_1, n_2) = s(n_1, n_2 + N_2), \text{对于所有 } (n_1, n_2)$$

例 1.2 对于 2D 离散复指数信号

$$s(n_1, n_2) = e^{j(\omega_1 n_1 + \omega_2 n_2)}$$

若下式成立，则其在 (n_1, n_2) 内具有矩形周期，周期为 (N_1, N_2) ：

$$\frac{\omega_1}{2\pi} = \frac{k_1}{N_1} \text{ 和 } \frac{\omega_2}{2\pi} = \frac{k_2}{N_2}$$

其中： N_1 、 N_2 、 k_1 和 k_2 为无量纲的整型数； ω_1 和 ω_2 的单位是 rad(弧度)。

⊖ 原书有误， n_1 、 n_2 应为列矢量。——译者注

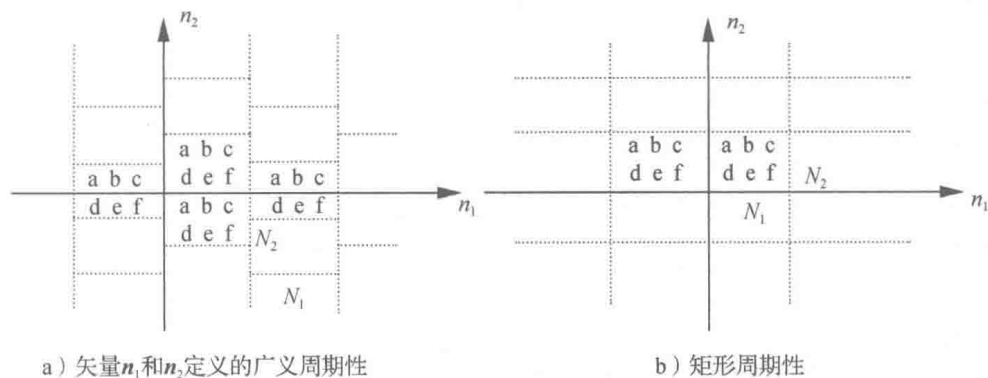


图 1.2 2D 信号的周期性

周期信号和有限域信号之间是同构的, 因为已知一个矩形周期信号, 该周期信号的主周期可以定义为一个有限域信号; 而已知一个有限域信号, 通过下式的周期扩展, 也可以定义周期信号:

$$\tilde{s}(n_1, n_2) = \sum_{i_1=-\infty}^{\infty} \sum_{i_2=-\infty}^{\infty} s(n_1 - i_1 N_1, n_2 - i_2 N_2) \quad (1.2)$$

1.1.2 对称信号

2D 空间和 MD 空间中有多种对称形式。最常见的 2D 空间的对称形式包括双重对称、四重对称和圆周对称。对称信号可以是有限域的, 也可以是无限域的。下面, 我们定义关于原点的对称。恰当地移动坐标轴就可以定义关于任意点的对称。

双重对称

双重矩形对称也称为无对称性的半平面对称, 其表示式为:

$$s(n_1, n_2) = s(-n_1, -n_2) \quad (1.3a)$$

图 1.1c 所示支撑中用圆点标出了其典型系数。支撑区域中的其他系数则由对称性来确定。

四重对称

四重矩形对称是一种更严格的对称形式, 其表示式为:

$$s(n_1, n_2) = s(-n_1, n_2) = s(n_1, -n_2) \quad (1.3b)$$

此时, 其典型系数的支撑是 $1/4$ 平面。

圆周对称

若信号 $s(n_1, n_2)$ 仅仅是其到原点的距离 $n_1^2 + n_2^2$ 的函数, 则称其为圆周对称。圆周对称隐含着四重对称。

1.1.3 特殊的多维信号

某些特殊信号对于图像和视频处理滤波器的理解起着重要作用, 包括可分离信号、空间频率模式、MD 脉冲、MD 单位阶跃信号等, 分别介绍如下。

可分离信号

若 MD 信号(函数)满足下式, 则其是可分离的:

$$s(n_1, n_2, \dots, n_M) = s_1(n_1) s_2(n_2) \cdots s_M(n_M) \quad (1.4)$$

以 2D 信号为例: 一个有限支撑 2D 信号 $s(n_1, n_2)$ 可以用矩阵 S 来表示。如果这个信号是可分离的, 那么矩阵 S 就可以写成外积 $S = s_1 s_2^T$ 的形式, 其中, 矢量 s_1 和 s_2 分别表示 1D 信号 $s_1(n_1)$ 和 $s_2(n_2)$ 的采样值。注意, 普通的 $N_1 \times N_2$ 矩阵的自由度为 $N_1 N_2$, 而这里外积的自由度为 $N_1 + N_2$ 。

例 1.3 2D 离散复指数信号是可分离的, 因为

$$s(n_1, n_2) = e^{j(\omega_1 n_1 + \omega_2 n_2)} = e^{j\omega_1 n_1} e^{j\omega_2 n_2} = s_1(n_1) s_2(n_2)$$

空间频率模式

水平空间频率模式可定义为:

$$s(n_1, n_2) = \cos(\omega_1 n_1)$$

其适用于所有的行(图像扫描行)。

垂直空间频率模式可定义为:

$$s(n_1, n_2) = \cos(\omega_2 n_2)$$

其适用于所有的列。

具有 45° 角方向的空间频率模式可定义为:

$$s(n_1, n_2) = \cos(\omega(n_1 - n_2))$$

MD 脉冲

MD 脉冲定义为:

$$\delta(n_1, n_2, \dots, n_M) = \begin{cases} 1, & n_1 = n_2 = \dots = n_M = 0 \\ 0, & \text{其他} \end{cases} \quad (1.5a)$$

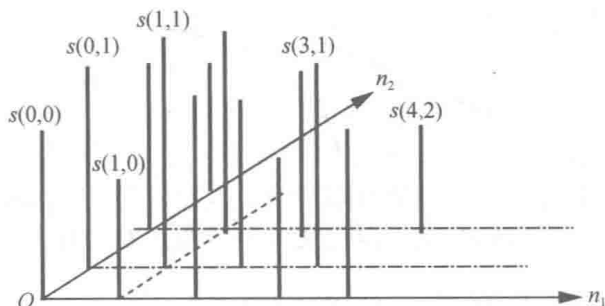


图 1.3 基于 2D 脉冲的 2D 信号表示

2D 脉冲表示平面上具有单位幅度的一个点源。离散图像(2D 信号)可以表示成加权平移脉冲之和, 其表示式为:

$$s(n_1, n_2) = \sum_{k_1=-\infty}^{\infty} \sum_{k_2=-\infty}^{\infty} s[k_1, k_2] \delta(n_1 - k_1, n_2 - k_2) \quad (1.5b)$$

在某些推导中, 这是 2D 信号常用的表示方式, 如图 1.3 所示。

例 1.4 2D 脉冲是可分离的, 因为可以写成:

$$\delta(n_1, n_2) = \delta_1(n_1) \delta_1(n_2)$$

其中: $\delta_1(n_1) = \begin{cases} 1, & n_1 = 0 \\ 0, & \text{其他} \end{cases}$ 是 1D 的克罗内克(Kronecker) δ 函数。

MD 单位阶跃

MD 单位阶跃常用于表示 MD 信号或 MD 系统冲激响应的支撑。它可以定义于任何形式的支撑之上。比如, 对于图 1.1 所示支撑, 我们可定义 2D 的 $1/4$ 平面单位阶跃、半平面单位阶跃或楔形单位阶跃。

例 1.5 具有 $1/4$ 平面支撑的 2D 单位阶跃定义为:

$$u_{QP}(n_1, n_2) = \begin{cases} 1, & n_1 \geq 0, n_2 \geq 0 \\ 0, & \text{其他} \end{cases}$$

具有半平面支撑的 2D 单位阶跃定义为:

$$u_{HP}(n_1, n_2) = \begin{cases} 1, & n_2 \geq 0 \\ 0, & \text{其他} \end{cases}$$

在图 1.1a 和 1.1b 所示支撑中圆点处的样点值为“1”，其他地方的为“0”。

1.2 多维变换

与矢量类似，信号也可以在一些基本函数上展开。MD 连续(离散)傅里叶变换将 MD 连续(离散)信号在 MD 连续(离散)复指数函数上展开，这些复指数函数构成正交基。为方便起见，这里我们只讨论 2D 傅里叶变换，但是 2D 傅里叶变换的所有定义和特性都可以推广到 MD 傅里叶变换。在第 3 章介绍有限冲激响应(FIR)图像滤波和多分辨率/多尺度信号表示之后，我们将讨论 2D 小波变换。

1.2.1 连续信号的傅里叶变换

虽然数字图像和视频是离散变量(序列)的函数，但由于离散变量信号是对连续变量信号进行采样得到的，因此，本节首先介绍连续变量函数的傅里叶变换(FT)，以引入 1.4 节中将用到的符号。

定义：2D 连续信号 $s(x_1, x_2)$ 的傅里叶变换可表示为：

$$S(u_1, u_2) = \iint_{(x_1, x_2)} s(x_1, x_2) e^{-j(u_1 x_1 + u_2 x_2)} dx_1 dx_2 \quad (1.6a)$$

2D 傅里叶反变换可表示为：

$$s(x_1, x_2) = \frac{1}{(2\pi)^2} \iint_{(u_1, u_2)} S(u_1, u_2) e^{j(u_1 x_1 + u_2 x_2)} du_1 du_2 \quad (1.6b)$$

其中： u_1 和 u_2 表示实的空域频率变量，其单位为周期数/距离。还可以再加一个时间变量，例如 $s(x_1, x_2, t)$ 。时间域频率的单位是 Hz。

2D 傅里叶变换的复数形式为：

$$S(u_1, u_2) = |S(u_1, u_2)| e^{j\theta(u_1, u_2)} = S_R(u_1, u_2) + jS_I(u_1, u_2) \quad (1.6c)$$

其中： $S_R(u_1, u_2)$ 和 $S_I(u_1, u_2)$ 分别表示实部和虚部； $|S(u_1, u_2)|$ 和 $\theta(u_1, u_2)$ 分别表示傅里叶幅度和傅里叶相位。

傅里叶变换的收敛性

2D 信号可能是无限的。因此，我们必须考虑二重积分式(1.6a)存在的条件。

- 一致收敛：如果 $\iint_{x_1 x_2} |s(x_1, x_2)| dx_1 dx_2 < \infty$ ，即 $s(x_1, x_2)$ 是绝对可积的，则其傅里叶变换存在，且是 u_1 和 u_2 的连续函数(即二重积分是一致收敛的)。
- 均方收敛：若 $S(u_1, u_2)$ 存在但不连续，此时适用一种弱收敛形式，称为均方收敛。例如， $s(x_1, x_2) = \sin(x_1) \sin(x_2) / (x_1 x_2)^\ominus$ 不是绝对可积的，但是在均方意义上其傅里叶变换收敛。此时可在不连续点周围观察到吉布斯(Gibbs)效应。
- 广义收敛：在这种情况下，一致收敛和均方收敛都不适用，但 $S(u_1, u_2)$ 仍可以用狄拉克(Dirac) δ 函数 $\delta(u_1, u_2)$ 进行定义。比如，对于所有 (x_1, x_2) ， $s(x_1, x_2) = 1$ 不是绝对可积的，但其傅里叶变换从广义上可以定义为 $S(u_1, u_2) = \delta(u_1, u_2)$ 。

连续信号多维傅里叶变换的特性

下面我们将讨论一些 MD 傅里叶变换独有的特性，这些特性在 1D 变换中没有对应。

分解为 1D 变换 MD 傅里叶变换可以分解成一系列针对每个变量的独立 1D 傅里叶变换。这里用 2D 情况进行说明。式(1.6a)可以表示为：

$$\begin{aligned} S(u_1, u_2) &= \iint_{(x_1, x_2)} s(x_1, x_2) e^{-ju_1 x_1} e^{-ju_2 x_2} dx_1 dx_2 \\ &= \int_{x_2} \left\{ \int_{x_1} s(x_1, x_2) e^{-ju_1 x_1} dx_1 \right\} e^{-ju_2 x_2} dx_2 \end{aligned}$$

$\ominus$ 原书为 $s(x_1, x_2) = (\sin x_1)/x_1 (\sin x_2)/x_2$ ，意义不明。——译者注