

分类号_____ 密级_____
UDC _____

西華大學
碩士學位論文
(學術型)

基于动态阈值神经 P 系统的图像融合研究

作者姓名： 李波
学科、专业： 计算机科学与技术
学号： 212018081200004
指导教师： 彭宏
完成日期： 2021年04月

Classified Index: _____
UDC: _____

Xihua University
Master Degree Thesis

**Image fusion based on dynamic threshold
neural P systems**

Candidate : Li Bo
Major : Computer Science and Technology
Student ID: 212018081200004
Supervisor: Prof. Peng Hong

April, 2021

西华大学学位论文独创性声明

作者郑重声明：所呈交的学位论文，是本人在导师的指导下进行研究工作所取得的成果。尽我所知，除文中已经注明引用内容和致谢的地方外，本论文不包含其他个人或集体已经发表的研究成果，也不包含其他已申请学位或其他用途使用过的成果。与我一同工作的同志对本研究所做的贡献均已在论文中做了明确的说明并表示了谢意。

若有不实之处，本人愿意承担相关法律责任。

学位论文作者签名: 李波 指导教师签名: 彭宏
日期: 2021.6.8 日期 2021.6.8

西华大学学位论文版权使用授权书

本学位论文作者完全了解学校有关保留、使用学位论文的规定，在校攻读学位期间论文工作的知识产权属于西华大学，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅，西华大学可以将本论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复印手段保存和汇编本学位论文。（保密的论文在解密后遵守此规定）

学位论文作者签名: 李波 指导教师签名: 彭宏
日期: 2021.6.8 日期 2021.6.8

摘要

动态阈值神经 P (DTNP) 系统是一种新型分布式和并行计算模型，它结合了神经元的脉冲机制和动态阈值机制。DTNP 系统在理论上已经被证明是图灵通用计算设备。本文主要研究 DTNP 系统与多尺度变换结合在图像融合中的应用，包括多焦点图像融合、红外与可见光图像融合和多模态医学图像融合。本文主要的工作和贡献如下：

- (1) 提出了一种具有局部邻域结构的二维 DTNP 系统。
- (2) 提出了一个基于 DTNP 系统的多焦点图像融合框架，其中 4 个 DTNP 系统用于控制低频和高频 ST 系数的融合。依据该框架，开发出一种基于 DTNP 系统和 ST 的多焦点图像融合算法。在公开的多焦点图像上对比实验结果验证了所提出的融合算法在多焦点图像融合上的可用性和融合性能等方面的优势。
- (3) 提出了一种基于 DTNP 系统和非下采样剪切波变换 (NSST) 的红外和可见光图像融合框架，其中低频系数的融合规则由两个 DTNP 系统控制。依据此框架，开发出一种基于 DTNP 系统和 NSST 变换的红外和可见图像融合算法。使用可见光和红外图像 TNO 公共图像集进行对比实验，验证了所提出的融合算法在融合红外和可见光图像的可用性和融合性能方面的优越性。
- (4) 结合 NSCT 系数的 WLE-INSML 特征，提出了一种基于 DTNP 系统的多模态医学图像融合框架，并开发出相应的多模态医学图像融合算法。在公开的多模态医学图像组成的开放数据集上的对比实验结果验证了所提出的融合算法在多模态医学图像融合上的可用性和融合性能等方面的优势。

关键词： 动态阈值神经 P 系统；图像融合；多尺度变换；局部邻域结构

Abstract

Dynamic threshold neural P (DTNP) systems are a novel kind of distributed and parallel computing models that combine the spiking mechanism and the dynamic threshold mechanism of neurons. DTNP systems have been theoretically proven to be Turing universal computing devices. This paper focuses on the application of combining DTNP systems and multiscale transform in image fusion, including multi-focus images fusion, infrared and visible image fusion, and multi-modal medical image fusion. The main work and contributions of this paper are as follows:

(1) A two-dimensional DTNP system with local neighborhood structure is proposed.

(2) A multi-focus image fusion framework based on DTNP systems is proposed, where four DTNP systems are used to control the fusion of low- and high-frequency ST coefficients. Based on the framework, a multi-focus images fusion algorithm based on DTNP systems and ST is developed. The comparison experiments on public multi-focus images demonstrate the advantages of the proposed fusion algorithm in terms of usability and fusion performance for the fusion of multi-focus images.

(3) A fusion framework for infrared and visible images based on DTNP system and NSST is proposed, where the fusion rules of low-frequency coefficients are controlled by two DTNP systems. Based on the framework, a fusion algorithm based on the DTNP system and the NSST transform for infrared and visible images is developed. Comparison experiments on a publicly available TNO image set of visible and infrared images demonstrate the advantages of the proposed fusion algorithm in terms of usability and fusion performance for the fusion of infrared and visible images.

(4) Combining the WLE-INSML features of NSCT coefficients, a fusion framework based on DTNP system for multi-modal medical images is proposed, and the corresponding fusion algorithm for multi-modal medical images is developed. Comparison experiments on an open dataset composed of public multi-modal medical images demonstrate the advantages of the proposed fusion algorithm in terms of usability and fusion performance for the fusion of multi-modal medical images.

Key Words: Dynamic threshold neural P systems; Image fusion; Multi-scale transform; Local neighborhood topology

目 录

摘要	I
Abstract	II
1 绪 论	1
1.1 研究背景和意义	1
1.2 图像融合研究现状	2
1.3 主要结构	3
2 准备知识	5
2.1 动态阈值神经 P 系统	5
2.1.1 动态阈值 P 系统的定义	5
2.2 多尺度变换	6
2.2.1 Surfacelet 变换	6
2.2.2 非下采样剪切波变换	7
2.2.3 非下采样轮廓波变换	7
2.3 图像融合中常见的源图像	8
2.4 图像融合评价标准	10
3 基于动态阈值 P 系统和 ST 的多焦点图像融合	14
3.1 引言	14
3.2 局部邻域结构的动态阈值 P 系统	14
3.2.1 局部邻域结构的二维 DTNP 系统	14
3.2.2 SML 和 SF 特征与 DTNP 组合	15
3.3 多焦点图像融合框架	16
3.3.1 低频 ST 系数融合规则	17
3.3.2 高频 ST 系数融合规则	18
3.3.3 一致性检验	19
3.4 实验结果与分析	19
3.4.1 评价指标和参数设定	20
3.4.2 与传统七种融合方法比较	21
3.4.3 与基于深度学习的融合方法比较	26
3.4.4 时间效率对比	31
3.5 本章小结	32
4 基于动态阈值神经 P 系统和 NSST 的红外与可见光图像融合	33

4.1	引言	33
4.2	NSST 低频系数与 DTNP 系统的组合	33
4.3	红外可见光图像融合框架	35
4.3.1	低频 NSST 系数融合规则	36
4.3.2	高频 NSST 系数融合规则	36
4.4	实验与分析	37
4.4.1	实验数据集	37
4.4.2	评价指标	38
4.4.3	参数设定	38
4.4.4	定性和定量实验结果比较	38
4.4.5	与三种基于深度学习的融合方法的比较	42
4.5	本章小结	46
5	基于动态阈值神经 P 系统和 NSCT 的多模态医学图像融合	47
5.1	引言	47
5.2	INSMIL 特征与 DTNP 的组合	47
5.3	多模态医学图像的融合框架	48
5.3.1	低频 NSCT 系数融合规则	49
5.3.2	高频 NSCT 系数融合规则	51
5.3.3	融合算法	53
5.4	实验结果	54
5.4.1	参数设定和客观评价指标	55
5.4.2	实验结果分析	55
5.5	本章小结	64
结	论	66
参	考 文 献	67
攻	读 硕士学位期间发表论文及科研成果	74
致	谢	75

1 绪 论

1.1 研究背景和意义

图像处理技术是一门涉及到计算机科学，数学，物理学等多学科交叉领域，诞生于 20 世纪 60 年代。图像处理是通过计算机对图像进行分析和处理，以达到特定目的的方法，包括图像配准、图像降噪、图像融合、图像重建、图像分割等一系列方法。图像处理技术在当前诸如人工智能、远程遥感、医学、军事等各个领域都有广泛的应用。同时也因图像处理的广阔应用场景以及计算机技术的飞速发展，该领域受到了广泛地关注，并在短短的几十年取得了长足的进步^[1-2]。

图像融合是图像处理领域中的一个重要的研究分支，随着传感器技术的发展，已经从单一传感器模式逐渐发展为多种传感器模式，由于各种传感器具有不同的成像原理、不同的工作波长范围等因素^[3]，使得图像融合成为了目前非常热门的研究方向。图像融合广泛的应用于医学^[4]、图像增强^[5]、军事运用，人工智能^[6]等各个领域。图像融合是将各类相同场景下具有不同特点的图像信息融合成一张综合的清晰的图像。随着传感器技术的发展，单一的可见光模式逐渐发展为多种传感器模式。各种传感器具有不同的成像机理、不同的工作波长范围、不同的工作环境与要求，完成不同的功能。由于传感器自身物理特性、成像机理和观察视角等各个方面的种种限制，单一的图像传感器往往不能够从场景中提取足够的信息，以至于很难甚至无法独立获得对一幅场景的全面描述。这就需要研究多源图像融合。利用多个传感器提供的互补信息，融合后的图像包含了更为全面、丰富的信息，其更符合人或机器的视觉特性、更有利于对图像的进一步分析处理以及自动目标识别，在医疗领域主要体现在 CT 和 MRI 等相关医学图像融合用于疾病的诊断和治疗，在军事领域上通过红外与可见光图像融合用于战场观测，武器制导。

虽然图像融合技术在最近几十年取得了巨大的进步，但是由于传感器自身物理特征、成像机理和观察视角等各个方面的因素，目前图像融合的方法依旧无法全面满足当前实际应用的需要，具体表现在融合图像质量、实时性等方面的原因，所以图像融合依旧是目前研究的一个热点问题。

动态阈值 P (Dynamic threshold neural P, DTNP) 系统是一种新颖的分布式、并行计算模型，由 Peng 等人^[7]于 2019 年正式提出。DTNP 系统在理论上已经被证明具有图灵机的计算能力，由于它具有特殊的动态阈值机制，所以 DTNP 系统为处理现实问题提供了新颖的思路。

本文是将 DTNP 系统与图像融合领域结合，提出若干新颖的图像融合框架，并为图像融合领域开辟了一种新的算法，所以，本文的研究具有一定的理论价值和实际意义。

1.2 图像融合研究现状

20 世纪 70 年代末期，Daily^[8]等人首次将雷达图像和 Landsat-MSS 影像融合用以质地分析。20 世纪 80 年代中期，多尺度变换理论被应用于图像融合。20 世纪 90 年代，小波理论被引入到图像融合领域^[9]，图像融合理论得到了进一步丰富和发展了。当前图像融合算法主要有四种类型^[10]：空间域方法（Spatial Method）、多尺度变换（Multi-Scale Transform, MST）、稀疏表示（Sparse Representation, SR）、基于深度学习的融合方法（Deep Learning based fusion method, DL）。

空间域方法是直接对图像的像素值进行操作，常常使用最简单的加权平均或者绝对值最大值的融合规则。虽然这两种类型的融合规则计算简单、运行速度快，但是融合的图像会丢失大量源图像的细节信息，比如轮廓和边缘信息。人们还将源图像划分为多个区域，然后使用各种锐度测量方法空间频率或梯度，然后选择区域进行分块融合^[11-13]。主成分分析（Principle Component Analysis, PCA）^[14]是一种典型的空间域方法，通过降维来寻找图像的主成分，根据主成分来确定各融合图像的权重并完成融合。Wan 等人^[15]提出了基于鲁棒性的主成分分析的多聚焦图像融合方法。Li 等人^[16]首次将图像引导滤波应用到图像融合并取得较好的效果。为了保留纹理信息，提出了一种新的基于形态成分分析（Morphological Component Analysis, MCA）^[17]的融合算法。除此之外，脉冲耦合神经网络（Pulse-coupled neural networks, PCNN）^[18-20]也被运用于图像融合，空间域融合方法通过简单的叠加运算的融合规则不仅会大幅度降低融合图像的对比度，还会丢失大量源图像的细节信息。

近年来，变换域方法受到了广泛的关注。这些方法包含三个步骤：首先将源图像分解为高、低频系数；然后利用不同的融合策略生成融合策略，最后将融合后的系数通过逆变换实现图像的融合^[21]。基于 MST 的方法主要包括拉普拉斯金字塔（laplacian pyramid, LP）^[22-24]、梯度金字塔（gradient pyramid, GP）^[25-26]、小波变换（wavelet transform, WT）^[27]、离散小波变换（discrete wavelet transform, DWT）^[28-30]和双树复小波变换（dual-tree complex wavelet transform, DTCWT）^[31]。然而，这些方法存在着非移位不变性、空间性差和非时间不变性等缺点。为了克服这些缺点，人们引入了一些基于 MST 的融合方法，包括曲线小波变换（curvelet transform, CT）^[31-32]、曲面小波变换（surfacelet transform, ST）^[33]、非下采样剪切变换（nonsubsampled shearlet transform, NSST）^[34]和非下采样轮廓变换（nonsubsampled contourlet transform, NSCT）^[35-36]。整体来看，MST 方法的特征

提取精准度较高，融合效果较好，优于空间域方法。但是，MST 方法也存在一定的局限性，比如存在费时，实时性较差等缺点。

基于 SR 的融合方法在给定的字典中识别出多模态图像的稀疏表示，然后构建融合图像。近年来，人们研究了几种基于 SR 的融合方法^[37-40]。然而，基于 SR 的方法中只采用了一种字典。为了解决这一局限性，在基于 SR 的方法中引入了字典学习^[41-43]。Zhu 等人^[44]提出了一种基于字典学习的图像融合方法，Li 等人^[45]提出了一种基于判别性低秩稀疏字典学习的医学图像融合方法，Liu 等人^[46]介绍了一种用于像素级医学图像融合的稀疏表示模型，称为基于卷积稀疏的形态成分分析。

在最近几年里，基于 DL 的融合方法变得越来越流行，因为它们可以实现有竞争力的融合性能。有几种基于 DL 的融合方法被开发出来，用于多模态图像的融合。Tang 等人^[47]研究了一种用于多焦点图像融合的像素卷积神经网络(CNN)，其中训练了一个模型，根据聚焦、失焦和未知像素的邻域像素信息来学习它们的概率。Gao 等人^[48]研究了一种用于多焦点图像融合的 CNN，其中使用了一个较深的网络来构建初始决策图。Amin-Naji 等人^[49]提出了一种基于 CNN 与集成学习相结合的多焦点图像融合方法。Li 等人^[50]提出了一种 DL 框架，该网络包括编码器和解码器，并将这种融合方法称为 DenseFuse。Zhang 等人^[51]提出了一个使用 CNN 的通用多模态图像融合框架，称为 IFCNN。这些基于 DL 的融合方法已经证明了具有竞争力的融合性能。Ma 等人^[52]提出了一种双判别器条件生成对抗网络 DDcGAN 用于不同分辨率的图像融合。Li 等人^[53]提出了一种基于低秩表示的多级图像分解方法 MDLatLRR。但是基于 DL 的融合方法需要一个耗时的训练过程。

动态阈值神经 P 系统 (DTNP 系统)^[7]是一种分布式并行计算模型和图灵通用计算模型，它结合了神经元的脉冲机制和动态阈值机制。

1.3 主要结构

本文的主要内容是探讨如何运用一个二维 DTNP 系统去处理图像融合问题，其中包括多焦点图像融合、红外与可见光图像融合和医学图像融合。

本文结构如下：

第一章为绪论，详细阐述了本文的研究背景和现状，并简要介绍了本文的主要工作。

第二章为准备基础，先介绍了动态阈值 P 系统的定义，同时对本文使用的 3 种多尺度变换方法进行了简要的介绍，然后介绍了图像融合中常见的源图像，最后介绍了目前图像融合评价标准。

第三章介绍了基于动态阈值 P 系统和 ST 变换的多焦点图像融合。首先介绍 ST 变换与 DTNP 系统结合，然后提出了一种新颖的多焦点图像融合框架，并详细的介绍了低

频 ST 系数融合规则、高频 ST 系数融合规则和后续的图像一致性检验。接着，给出了详细的实验结果和分析，对比了 10 种先进的多焦点图像融合方法，其中包括 7 种传统的融合方法和 3 种基于深度学习的方法，此外，还分析对比了时间效率。最后，对本章进行了小结。

第四章介绍了一种基于 DTNP 系统的红外可见光图像融合框架。然后对实验结果进行了详细分析，包括实验数据集、评价指标、参数设定、定性和定量实验结果分析等。对比了 10 种先进的红外可见光图像融合方法。最后，对本章进行了小结。

第五章介绍了一种基于 DTNP 的多源医学图像融合框架，首先介绍 DTNP 与 NSCT 结合。然后对比了 13 种先进的医学图像融合方法，其中包括 9 种传统的融合方法和 4 种基于深度学习的方法，从定性和定量的角度对实验进行了详细的分析。最后，对本章进行了小结。

最后，总结本文内容，提出今后进一步研究的方向。

2 准备知识

2.1 动态阈值神经 P 系统

动态阈值神经 P 系统 (DTNP 系统)^[7]是脉冲神经 P 系统 (SNP 系统)^[54-59]的一种变体。DTNP 系统是由 Peng 等人^[7]提出的一种新型分布式和并行计算模型，它结合了神经元的脉冲机制和动态阈值机制。在之前的工作中，Peng 等人^[7]已经证明了 DTNP 系统是图灵通用的数字生成/接受设备和函数计算设备。下面简要地介绍一维 DTNP 的定义。

2.1.1 动态阈值 P 系统的定义

一个度 m 的 DTNP 系统可表示为：

$$\Pi = (O, \sigma_1, \sigma_2, \dots, \sigma_m, syn, out) \quad (2.1)$$

其中

- (1) $O = \{a\}$ 是单字符集合，其中 a 指一个脉冲；
- (2) $\sigma_i = (u_i, \tau_i, R_i)$ 代表动态阈值神经元，其中 $u_i \geq 0$ 是神经元 σ_i 的脉冲数量， $\tau_i \geq 0$ 是神经元 σ_i 的动态阈值， R_i 表示神经元 σ_i 中所有规则的有限集合，它的表达形式为 $E_i/(a^u, a^\tau) \rightarrow a^p$ ，其中 $E_i = \{u_i \geq \tau_i \wedge u_i \geq u \wedge \tau_i \geq \tau\}$ 是点火条件， $u \geq 1, \tau \geq 0, p \geq 0$ 并且 $p \leq u$ ；如果 $p \geq 1$ 我们称之为点火规则， $p = 0$ 称之为遗忘规则；
- (3) $syn \subseteq \{1, 2, \dots, m\} \times \{1, 2, \dots, m\}$ 表示神经元之间连接(突触)，并且对任意 $1 \leq i \leq m, (i, i) \notin syn$ ；
- (4) $in, out \in \{1, 2, \dots, m\}$ 分别表示输入和输出神经元。

DTNP 系统被视为一个没有环的有向图，图中的节点由神经元代表，而弧线则表示这些神经元之间的突触连接。在这个图中可以考虑一个代表环境的虚拟节点，也可以指定两个不同的神经元为输入神经元和输出神经元，允许系统与环境进行交流。

除了有向图结构外，DTNP 系统有两个要素：数据和规则。数据用来描述每个神经元的状态，规则刻画系统的动态行为。与 SNP 系统不同，DTNP 系统中的每个神经元有一个数据单元 u_i 和一个动态阈值单元 τ_i 。用 $u_i(t)$ 和 $\tau_i(t)$ 分别表示在 t 时刻数据单元和动态阈值单元的脉冲数量。可以采用一个向量来描述系统在 t 时刻的状态： $C_t = (u_1(t), \tau_1(t), \dots, u_m(t), \tau_m(t))$ ，称为一个格局或配置。于是，初始配置记为一个向量 $C_0 = (u_1, \tau_1, \dots, u_m, \tau_m)$ ，即对于所有的 $1 \leq i \leq m$ ，有 $u_i(0) = u_i$ 和 $\tau_i(0) = \tau_i$ 。

DTNP 系统有两种类型的规则：点火规则和遗忘规则。有一个统一的形式 $E_i / (a^u, a^\tau) \rightarrow a^p$ ，其中 $p = 0$ ，称为遗忘规则，否则称为点火规则。特别说明，在本文处理图像时不再严格区分这两种形式，都简称为点火规则。规则的点火条件为 $E_i = \{u_i \geq \tau_i \wedge u_i \geq u \wedge \tau_i \geq \tau\}$ ，将作为缺省条件。DTNP 系统中规则语义被说明如下。假设系统工作在 t 时刻，如果满足 $u_i(t) \geq \tau_i(t)$ 、 $u_i(t) \geq u$ 和 $\tau_i(t) \geq \tau$ ，那么神经元 σ_i 的规则 $E_i / (a^u, a^\tau) \rightarrow a^p$ 将被运用，即神经元 σ_i 点火。当神经元 σ_i 点火时，它将消耗值为 u 的脉冲和值为 τ 的动态阈值。然后，生成值为 p 的脉冲。如果 $p > 0$ ，那么产生的脉冲将会传递到突触后的每个神经元 σ_j ，其中 $(i, j) \notin syn$ 。因此，根据这个点火规则的语义，神经元 σ_i 的状态更新公式如下：

$$\begin{cases} u_i(t+1) = u_i - u + n \\ \tau_i(t+1) = \tau_i - \tau + p \end{cases} \quad (2.2)$$

其中 n 表示从其他神经元收到的脉冲值， p 表示神经元自身产生的脉冲值。

2.2 多尺度变换

2.2.1 Surfacelet 变换

Yu 等人^[60]扩展了方向滤波组 (DFB)^[61]，并提出了一种针对 N 维信号 ($N \geq 2$) 的新型滤波组 NDFB。基于 NDFB 具有多尺度金字塔，Yu 等人^[60]提出了 Surfacelet Transform (ST)，ST 可以用来有效地捕捉和表达多维数据中不同大小的表面类奇点。

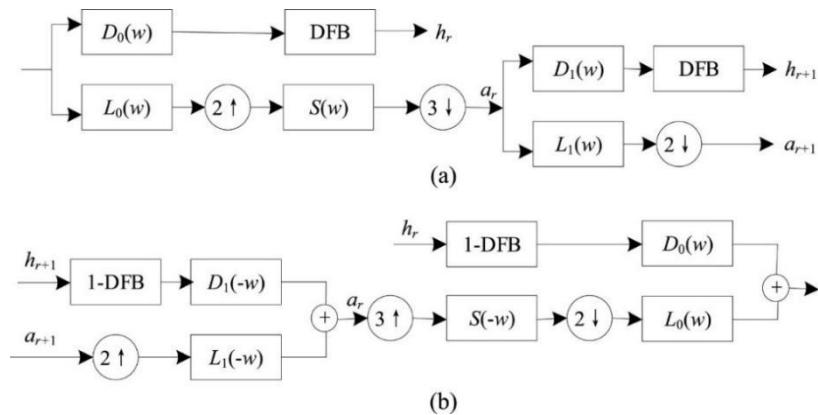


图 2.1 surfacelet transform 的结构^[61]: (a) 分解 (b) 合成

Fig. 2.1 The structure of surfacelet transform^[61]: (a) decomposition and (b) synthesis

其思想与等值线构造类似，但是一个重要的区别是，ST 采用了一种新的多尺度金字塔结构，而不是像等值线那样的 Laplacian 金字塔。虽然这种方法比 Laplacian 金字塔

包含了很多冗余信息，但它可以避免 NDFB 在频域产生的混杂现象。这种特殊的采样方法使得曲面小数的分解系数在空间和频域上都有很好的表达，因此，设计这种滤波器的主要优点是可以避免在频域上产生混杂现象。因此，在空间域的这种滤波器的主要优点是可以严格控制图像的频率响应。图 2.1 为 Surfacelet Transform 结构。

2.2.2 非下采样剪切波变换

非下采样剪切波变换（NSST）是一种先进的多尺度几何分析工具^[62]，是对剪切波（Shearlet）变换的改进。NSST 具有平移不变性、多方向性、计算复杂性低等特点，被用于图像去噪、目标识别、缺陷提取、图像增强、图像融合等领域^[63-65]。NSST 变换主要分为以下两步：多尺度分解和多方向分解；NSST 通过非下采样金字塔滤波器（Nonsubsampled Pyramid, NSP）进行 n 次可以得到 1 个低频系数和 n 个高频系数，NSST 通过剪切滤波器（Shearlet Filter, SF）实现多方向分解^[66]。图 2.2 为 NSST 的分解过程。

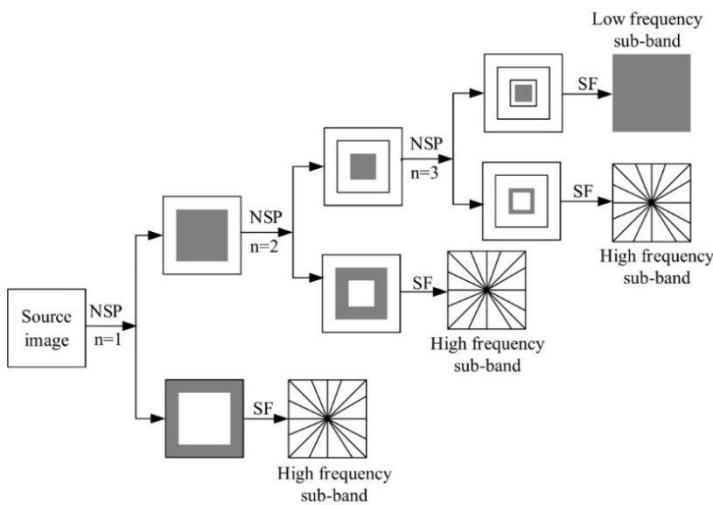


图 2.2 NSST 分解过程^[62]

Fig. 2.2 The decomposition process of NSST^[62]

2.2.3 非下采样轮廓波变换

2005 年 Cunha 和 Zhou 等人^[67]受到非下采样小波变换的启发，首次提出非下采样轮廓波（Nonsubsampled Contourlet Transform, NSCT）。NSCT 是基于非下采样金字塔（Nonsubsampled Pyramid, NSP）和方向分解的非下采样方向滤波组（Nonsubsampled Directional Filter Bank, NSDFB）的一种变换^[68]。由于 NSCT 取消了下采样的操作，这

样就能够有效的消除图像在奇异处引起的伪吉布斯现象，NSCT 还具有平移不变，多尺度，多方向性等一系列优良的特性^[69]。

NSP 将图像分解为高频和低频两个部分，实现图像的多尺度分解；然后，由 NSDFB 将高频频子带分解为多个方向子带，实现图像的多方向分解；最后，对每一层的低频部分继续进行如上分解，这样便实现了 NSCT 的多尺度分解^[70]。若对 NSDFB 进行 l 级方向分解，包含 1 个低频子带图像和 l 个高频频子带图像。NSCT 的分解过程如图 2.3 所示。

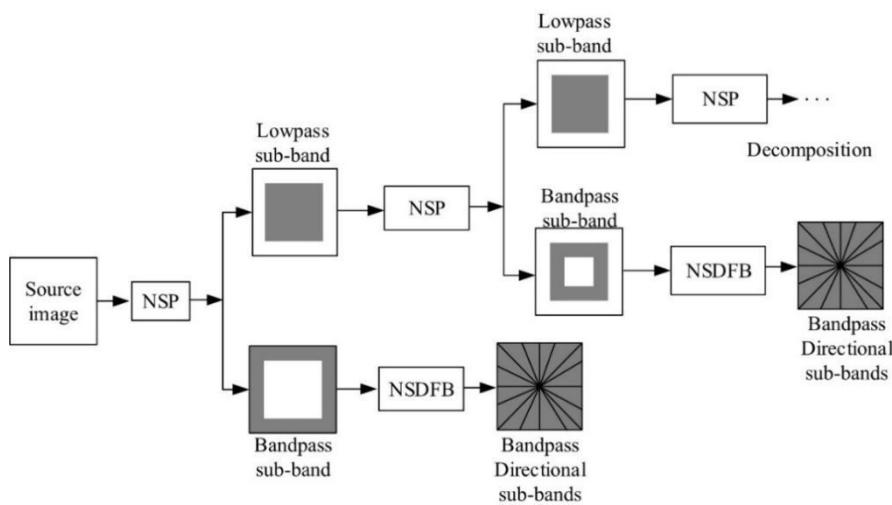


图 2.3 NSCT 分解过程^[68]

Fig. 2.3 The decomposition process of NSCT^[68]

2.3 图像融合中常见的源图像

由于不同类型的传感器所拍摄的图像具有很大差异，因此在设计图像融合算法尤其是多模态图像融合算法的过程中需要考虑到不同模态的图像的特性和共性才能设计出更好的算法。因此，本节将对常见的源图像类型进行简要介绍。

当前常见的源图像主要有多焦点图像、红外图像、可见光图像、计算机断层扫描图像（Computed Tomography, CT）、核磁共振图像（Magnetic Resonance Imaging, MRI）^[71]、正电子发射型计算机断层显像（Positron Emission Computed Tomography, PET）^[72]、单光子发射计算机断层成像（Single-Photon Emission Computed Tomography, SPECT）^[73]。

由于每个带有光学相机的成像设备具有有限的景深，因此相机捕获的图像不能全部聚焦。因此，具有特定景深的对象是清晰的，而其他对象则是模糊的。如图 2.4 所示，图（a）的左侧是清晰聚焦，右侧是未聚焦模糊的。图（b）则刚好相反。

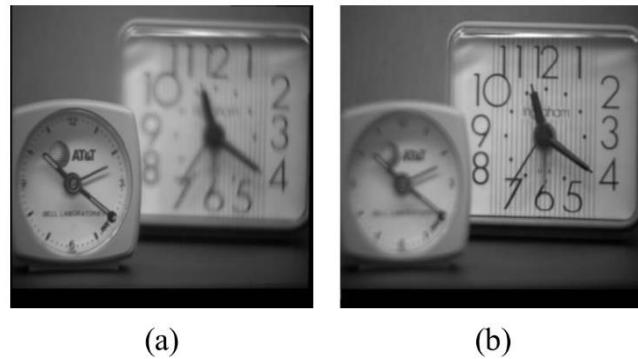


图 2.4 多焦点图像

Fig. 2.4 Multi-focus images

红外图像是指波长为 0.76-1000 微米之间的电磁波图像，物体温度越高则图像对应区域的光强就越高。由于这类电磁波是人眼无法感知到的，并且具有许多可见光不具备的特性，因此这类图像包含了大量可见光图像不具备的信息。比如在夜间条件下，由于可见光比较少因此拍摄的图像十分的模糊，但是红外图像却几乎可以不受影响。因此在实际的融合中通常将这类图像和可见光图像进行融合，得到质量更高的图像，红外图像如图 2.5 (a) 所示。

由于人眼能够感知到的光波十分有限，波长为 400-760 纳米，而这一类光就被称为可见光。因而可见光图像就是人类最敏感的一类图像，因此我们更易察觉可见光图像中诸如细节，纹理和颜色等各个方面信息。因此，大多数图像融合算法所生成的融合图像都尽可能倾向于可见光图像已达到更好的视觉效果，可见光图像示例如图 2.5 (b) 所示。

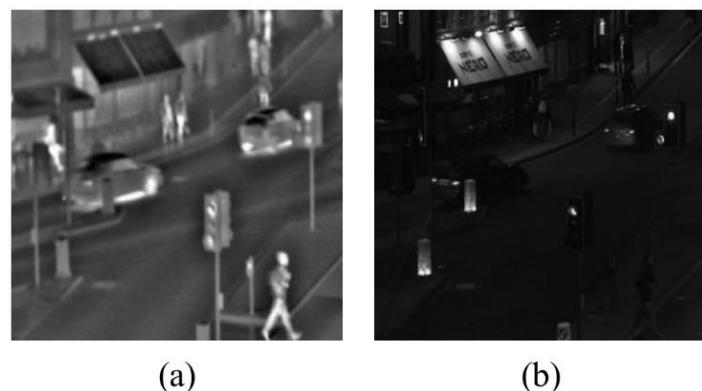


图 2.5 红外与可见光图像

Fig. 2.5 Infrared and visible images

计算机断层扫描（Computed tomography, CT）图像就是通过 X 光线衰减系数的差异性所生成的图像^[74]，如图 2.6（a）所示。核磁共振图像则是利用了人体不同组织，器官的氢原子具有磁性的原子核的核密度的差异性所绘制的图像^[75]，核磁共振图像如图 2.6（b）所示。正电子发射断层成像术和单光子发射计算机断层成像术都是对从病人体内发射的 γ 射线成像^[76-77]。图 2.6 中的（c）（d）分别为 PET 图像和 SPECT 图像。

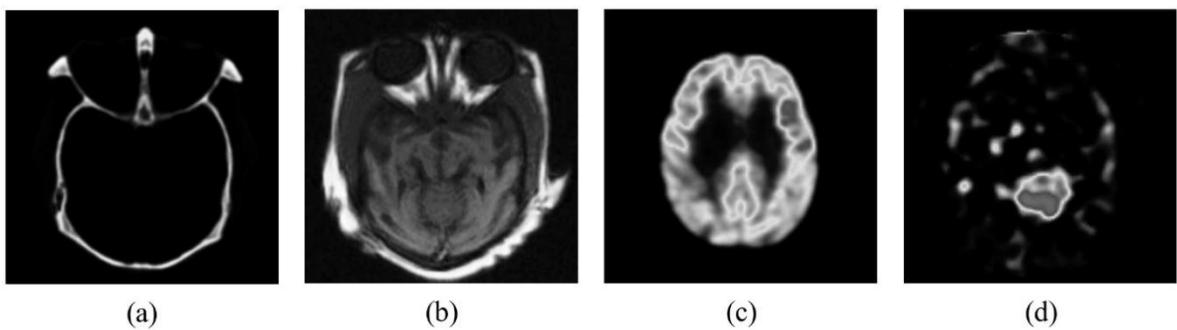


图 2.6 医学图像
Fig. 2.6 Medical images

正是由于不同模态的源图像存在的差异性，图像融合算法可以按照需要融合的源图像的种类分为单模态图像融合和多模态图像融合两大类^[78]。单模态图像融合就是融合由相同类型的传感器所拍摄的图像，多焦点图像融合就是一种常见的单模态图像融合。同理多模态图像融合则是融合两种或者以上类型的传感器所拍摄的图像，红外与可见光图像融合就是常见的多模态图像融合。

2.4 图像融合评价标准

图像融合质量的优劣主要从三方面来评价：（1）是否保留了所有源图像信息；（2）融合图像是否加入了额外的噪音；（3）是否消除了冗余信息和噪音。当前图像融合的评估方式分为两类^[79]：主观评价和客观评价。主观评价一般分为很差、差、一般、好和非常好 5 个等级。但是当融合图像质量十分接近时，这种判断方式的主观性就非常强从而无法准确地判断融合图像的质量。目前大多数图像融合算法用人眼已经不能区分图像融合质量的优劣，所以主观评价方法并没有大规模使用^[80]。客观评价方式是利用某种数学算法对图像质量做出定量评价，客观评价在图像融合中占据更加重要的地位，因为通过评价指标衡量融合图像的质量能够避免主观的问题而且更加方便和快速。在本文中，使用了目前最流行的 10 种客观指标，它们分别是 Q_{abf} （petrovic metric）^[81]、互信

息 (Mutual information, MI)^[82]、特征互信息 (Feature Mutual information, FMI)^[83]、标准差 (Standard deviation, SD)^[84]、信息熵 (entropy, EN)^[85]、交叉熵 (Cross Entropy, CE)^[86]、 Q_w ^[87]、 Q_e ^[87]、差异相关性之和 (sum of the correlations of differences, SCD)^[88]、多层级结构相似性 (Multiscale structural similarity, $MSSIM$)^[89]。

Q_{abf} ^[81]衡量的是融合图像保留源图像梯度信息的丰富程度, 相应的数学表达形式如下, 其中 $Q_g^{XF}(i,j)$ 和 $Q_a^{XF}(i,j)$ 分别表示源图像 X 和融合图像 F 在对应点的梯度和方向信息, A, B 表示源图像, M 和 N 则表示图像的尺寸。

$$Q_{abf} = \frac{\sum_{i=1}^M \sum_{j=1}^N Q^{AF}(i,j)w^A(i,j) + Q^{BF}(i,j)w^B(i,j)}{\sum_{i=1}^M \sum_{j=1}^N w^A(i,j) + w^B(i,j)} \quad (2.3)$$

$$Q^{XF}(i,j) = Q_g^{XF}(i,j) + Q_a^{XF}(i,j) \quad (2.4)$$

互信息 (Mutual information, MI)^[82]度量的是融合图像包含源图像信息量的丰富程度。其中 X 代表源图像, $p_X(i)$ 和 $p_F(i)$ 分别表示源图像和融合图像的直方图, $p_{X,F}(i,j)$ 表示图像 X 和 F 的联合密度。对于 p_X 或者 p_F 而言, 一共有 $\{0,1,\dots,255\}$ 共 256 个像素取值的概率, 对于 $p_{X,F}$ 来说, 一共有 256×256 个像素取值的概率, 分别为 $\{(0,0), (1,1), \dots, (255, 255)\}$ 。 MI 定义如下:

$$MI(A,B,F) = MI_{A,F} + MI_{B,F} \quad (2.5)$$

$$MI_{X,F} = \sum_{i=1}^L \sum_{j=1}^L P_{X,F}(i,j) \log_2 \left(\frac{P_{X,F}(i,j)}{P_X(i)P_F(j)} \right) \quad (2.6)$$

特征互信息 (Feature Mutual information, FMI)^[83]是基于互信息 (Mutual information, MI)^[82]和特征信息的一种评价指标, FMI 值越大, 融合图像包含源图像的结构信息也就越丰富。相应定义公式 (2.7) 所示:

$$FMI = MI_{A^f,F^f} + MI_{B^f,F^f} \quad (2.7)$$

标准差 (Standard deviation, SD)^[84]反映融合图像灰度离散分布, SD 值越大融合图像亮度越高。相应的定义如下:

$$SD = \sqrt{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (F(i,j) - u)^2} \quad (2.8)$$

其中 $F(i,j)$ 表示融合图像在 (i,j) 处的像素值, M 和 N 表示图像的大小, u 表示融合图像像素的均值。 u 的定义如下:

$$u = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x_{i,j} \quad (2.9)$$

基于信息理论的信息熵 (entropy, EN)^[85]是衡量融合图像信息的丰富程度, 相应定义如下:

$$EN = -\sum_{i=0}^{L-1} p_i \log_2 p_i \quad (2.10)$$

对于一张灰度图，像素的取值为 0-255，那么 $L = 256$ 。 p_i 表示灰度值为 i 的概率，可以由灰度值为 i 的像素个数 N_i 与所有像素 N 之比计算，即 $p_i = \frac{N_i}{N}$ 。对于 RGB 图像一般是将其转为灰度图像再计算熵。

交叉熵 (Cross Entropy, CE) [86] 表示生成图像与源图像信息的差异。交叉熵大表示生成图像与源图像差异更大，反之表示生成图像与源图像差异更小。假设 A 和 B 是源图像，F 是生成图像，公式如下：

$$CE = \frac{CE_{A,F} + CE_{B,F}}{2} \quad (2.11)$$

其中 $CE_{A,F}$ 和 $CE_{B,F}$ 定义如下：

$$CE_{A,F} = \sum_{i=0}^{L-1} p_{Ai} \log \frac{p_{Ai}}{p_{Fi}} \quad (2.12)$$

$$CE_{B,F} = \sum_{i=0}^{L-1} p_{Bi} \log \frac{p_{Bi}}{p_{Fi}} \quad (2.13)$$

Q_w [87] 基于 UIQI 理论所设计的图像融合质量评价指标。首先分别对融合图像和源图像进行分块，然后计算对应块的指标值，最后将各个块的值相加得到最终的值，其中 ω 表示图像的局部滑窗， W 表示整幅图像的窗口， $Q_0(A, F|\omega)$ 和 $Q_0(B, F|\omega)$ 分别表示融合图像在源图像 A 和 B、在 ω 滑窗的值，而 $\lambda(\omega) (0 \leq \lambda(\omega) \leq 1)$ ， $c(\omega)$ 分别表示在 ω 窗口的权重值和显著性值。通常， Q_w 值越大则融合图像的质量就越好。相应公式如下所示：

$$Q_w(A, B, F) = \sum_{\omega \in W} c(\omega) (\lambda(\omega) Q_0(A, F|\omega) + (1 - \lambda(\omega)) Q_0(B, F|\omega)) \quad (2.14)$$

$$\lambda(\omega) = \frac{s(A|\omega)}{s(A|\omega) + s(B|\omega)} \quad (2.15)$$

Q_e [87] 是根据人类视觉系统的特点对图像评价指标 Q_e 进行改进的一种图像融合评价指标，相应的数学表达形式如下。其中 A' , B' 和 C' 分别图像 A, B, F 对应的边缘权重矩阵， α 表示权重参数。

$$Q_e(A, B, F) = Q_w(A, B, F) \cdot Q_w(A', B', F')^\alpha \quad (2.16)$$

差异相关性之和 (sum of the correlations of differences, SCD) [88] 利用融合图像与源图像相关值的总和作为融合图像的质量度量。它的定义如下：

$$SCD = r(D_1, A) + r(D_2, B) \quad (2.17)$$

其中 A、B 和 F 分别表示源图像 A、源图像 B 和融合图像 F， $D_1 = F - A$, $D_2 = F - B$ ； $r(\cdot)$ 表示计算图像之间的相关性，它的定义如下：

$$r(D_k, S_k) = \frac{\sum_i \sum_j (D_k(i,j) - \bar{D}_k)(S_k(i,j) - \bar{S}_k)}{\sqrt{(\sum_i \sum_j (D_k(i,j) - \bar{D}_k))^2} (\sum_i \sum_j (S_k(i,j) - \bar{S}_k))^2} \quad (2.18)$$

其中, $K = 1, 2$, \bar{D}_k 和 \bar{S}_k 表示 D_k 、 S_k 的平均像素值。

多层级结构相似性 (Multiscale structural similarity, *MSSSIM*)^[89] 用来衡量源图像和融合图像的相似性。其中 $\alpha = \beta = \gamma = 1$, M 表示图像的最大尺度, *MSSSIM* 的定义如下:

$$MSSSIM(X, F) = [l_M(X, F)]^{\alpha M} \prod_{i=1}^M [c_i(X, F)]^{\beta_i} [s_i(X, F)]^{\gamma_i} \quad (2.19)$$

$$MSSSIM(A, B, F) = MSSSIM(A, F) + MSSSIM(B, F) \quad (2.20)$$

其中 A, B 和 F 分别表示源图像和融合图像, $l(\cdot)$ 、 $c(\cdot)$ 、 $s(\cdot)$ 分别表示求亮度、对比度和结构相似性, 它们的定义如下:

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (2.21)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (2.22)$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \quad (2.23)$$

其中 μ 表示对应块的均值, σ 表示标准差, σ_{xy} 表示协方差, C_1 , C_2 , C_3 表示对应项的参数。

3 基于动态阈值 P 系统和 ST 的多焦点图像融合

3.1 引言

多焦点图像融合由于其在图像处理和计算机视觉中的可用性和有效性，已经成为一个新的研究热点^[90]。由于每个光学相机的成像设备的景深都是有限的，相机所拍摄的图像不可能全部对焦。因此，具有特定景深的物体很清晰，但其他物体却很模糊。多焦点图像融合是将两幅或多幅具有不同景深的源图像进行合并，以产生更清晰的图像。由于融合后的图像具有更多的细节信息，所以更适合人类的视觉系统。本章讨论将如何应用动态阈值神经 P 系统处理多焦点图像的融合问题，提出了一种基于 DTNP 系统在表面波变换（surfacelet transform, ST）的图像融合方法。首先，开发了一种基于局部邻域结构的动态阈值神经 P 系统，并以此为基础构建了一种在 ST 的多焦点图像融合框架，其中 DTNP 系统用于控制 ST 低频和高频系数的融合。将在 20 幅公开的多焦点图像上对所提出的融合方法进行了 5 个融合质量指标的评估，并与 10 种最先进的融合方法进行了比较。

3.2 局部邻域结构的动态阈值 P 系统

3.2.1 局部邻域结构的二维 DTNP 系统

在 2.1 小节已经介绍了一维 DTNP 的定义，但是为了将 DTNP 系统运用到图像融合问题，我们构造了具有邻域结构的二维 DTNP 系统。如图 3.1 所示，神经元 σ_{ij} 首先接收来自邻域神经元的脉冲，然后在点火规则执行后将产生的脉冲送回给每个邻域神经元。因此，用双向箭头来描述神经元与其邻近神经元之间的突触连接。神经元 σ_{ij} 由三部分组成：一个数据单元 u_{ij} 、一个动态阈值单元 τ_{ij} ，和一条激发规则 $E/(a^u, a^\tau) \rightarrow a^p$ 。此外，神经元 σ_{ij} 还有一个外部输入（对应于特征矩阵的一个元素），见图 3.1 的左边部分。

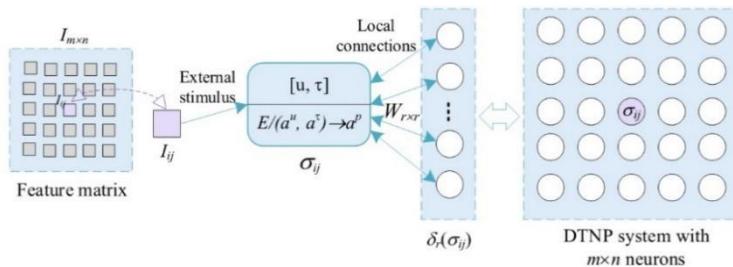


图 3.1 局部邻域结构的二维 DTNP 系统Π

Fig. 3.1 2-dimensional DTNP system Π with local neighborhood structure

3.2.2 SML 和 SF 特征与 DTNP 组合

假设 I 是大小为 $m \times n$ 的多焦点图像，其中 $(SML_{ij})_{m \times n}$ 和 $(SF_{ij})_{m \times n}$ 是图像的低频和高频 ST 系数矩阵。图 3.2 展示了一个 DTNP 系统 Π 与特征矩阵 $(SML_{ij})_{m \times n}$ 或 $(SF_{ij})_{m \times n}$ 的对应关系，其中特征矩阵有 $m \times n$ 行元素， Π 包含 $m \times n$ 个神经元，并且 $1 \leq i \leq m$ 、 $1 \leq j \leq n$ 。

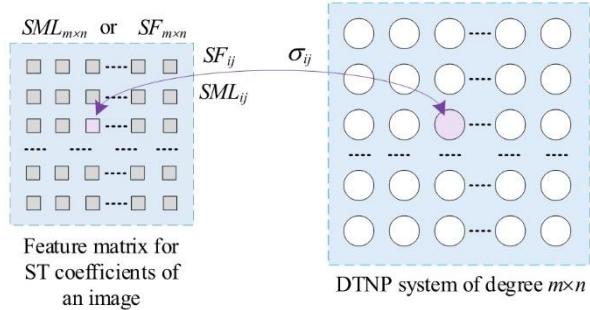


图 3.2 一个 DTNP 系统和相应的 ST 系数的特征矩阵

Fig. 3.2 An DTNP system Π and the corresponding feature matrix of ST coefficients.

现在我们解释 DTNP 系统 Π 的工作机制。神经元 σ_{ij} 与邻近的神经元有局部的连接，并且神经元 σ_{ij} 有两部分输入：从其相邻神经元收到的脉冲值和外部输入（比如特征值 SML_{ij} 或者 SF_{ij} ）。神经元 σ_{ij} 的点火规则有 $E/(\alpha^u, \alpha^\tau) \rightarrow \alpha^p$ 形式，其中神经元点火条件为 $E \equiv (u_i(t) \geq \tau_i(t)) \wedge (u_i(t) \geq u) \wedge (\tau_i(t) \geq \tau)$ 。如果满足点火条件，那么神经元 σ_{ij} 消耗数据单元中的脉冲值 u 和动态阈值单元中的脉冲值 τ ，然后产生一个值为 p 的脉冲并发送给周围的神经元。根据点火机制，神经元的状态方程可以表达如下：

$$u_{ij}(t+1) = \begin{cases} u_{ij}(t) - u + SML_{ij}(\text{or } SF_{ij}) & \text{if } \sigma_{ij} \text{ fires} \\ & + \sum_{\sigma_{kl} \in \delta_r} w_{kl} P_{kl}(t), \\ u_{ij}(t) + SML_{ij}(\text{or } SF_{ij}) & \text{otherwise} \\ & + \sum_{\sigma_{kl} \in \delta_r} w_{kl} P_{kl}(t), \end{cases} \quad (3.1)$$

$$\tau_{ij}(t+1) = \begin{cases} \tau_{ij}(t) - \tau + p, & \text{if } \sigma_{ij} \text{ fires} \\ \tau_{ij}(t), & \text{otherwise} \end{cases} \quad (3.2)$$

其中， $P_{kl}(t)$ 是神经元 σ_{ij} 从周围神经元 σ_{kl} 接收到的脉冲值， $w_{kl}(t)$ 是局部权重， SML_{ij} (SF_{ij}) 是外部输入，并且 p 是神经元 σ_{ij} 点火产生的脉冲值。

3.3 多焦点图像融合框架

我们提出了一种基于 DTNP 系统的 ST 域多焦点图像融合框架，如图 3.3 所示。该融合框架包括四个部分。（1）ST 变换；（2）融合规则；（3）逆 ST 变换；（4）优化。在图 3.3 中，源图像 A 和 B 是两幅多焦点图像。

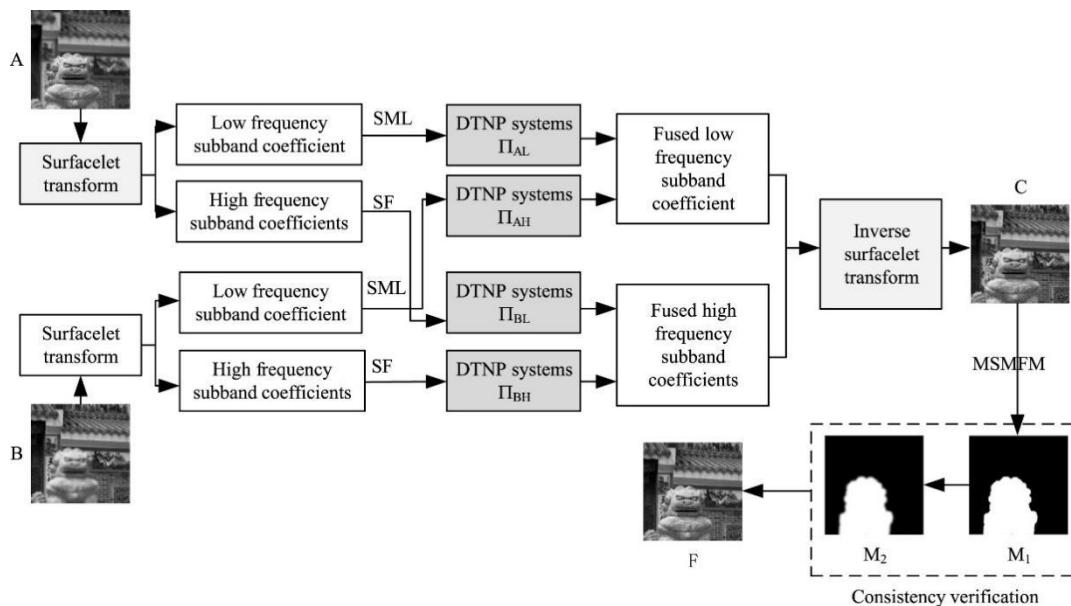


图 3.3 基于 DTNP 系统在 ST 变换域的融合框架，其中 A、B 为源图像，C 为初始融合图像，F 为最终融合图像

Fig. 3.3 The proposed fusion framework based on DTNP systems in ST domain, where A and B are two source images, C is initial fusion image, and F is final fusion image

首先利用 ST 将多焦点图像分解为 ST 系数。然后，将 ST 系数进行融合，生成融合后的 ST 系数。低频系数和高频系数分别由两组基于 DTNP 系统的融合规则进行融合，每组融合规则使用两个 DTNP 系统。DTNP 系统是所提出的融合框架中的核心部分，它们分别用于控制低频系数和高频 ST 系数的融合。两个 DTNP 系统 Π_{AL} 和 Π_{BL} 是合并多焦点图像的低频系数，而另外两个 DTNP 系统 Π_{AH} 和 Π_{BH} 应用于合并多焦点图像的高频系数。合并后的 ST 系数通过逆 ST 变换重建出融合图像。最后，对得到的融合图像进一步优化。根据融合框架，所提出的多焦点图像融合方法由以下步骤实现。

- (1) ST 变换：利用 ST 将多焦点图像 A 和 B，分别分解为 ST 系数，从而得到多焦点图像的低频 ST 系数和高频 ST 系数。

- (2) 低频 ST 系数的融合：基于多焦点图像 A 和 B 的低频 ST 系数，我们计算出两个特征矩阵 SML^{AL} 和 SML^{BL} ，并且将它们作为 DTNP 系统 Π_{AL} 和 Π_{BL} 的外部输入。两个 DTNP 系统从初始状态开始，不断地执行直到达到最大迭代次数 t_{max} 为止而停止，最后将 DTNP 系统的输出作为低频融合规则的控制信号（见公式（3.5））。
- (3) 高频 ST 系数的融合：基于多焦点图像 A 和 B 的高频 ST 系数，我们计算出两个特征矩阵 SF^{AH} 和 SF^{BH} ，并且将它们作为 DTNP 系统 Π_{AH} 和 Π_{BH} 的外部输入。两个 DTNP 系统从初始状态开始，不断地执行直到达到最大迭代次数 t_{max} 为止而停止，最后将 DTNP 系统的输出作为高频融合规则的控制信号（见公式（3.9））。
- (4) 逆 ST 变换：利用逆变换将融合后的 ST 系数转换回来，重建出融合图像。
- (5) 一致性验证：优化初始融合图像，我们将使用公式（3.10）-（3.12）。

在下文中，我们详细介绍了用于低频和高频 ST 系数的两种融合规则以及优化的一致验证过程。

3.3.1 低频 ST 系数融合规则

低频系数表达了图像的近似信息，因为它们包含了大部分的能量。基于 SML (sum-modified Laplacian) 的方法可以用来选择变换域中包含更多细节信息的系数。为了更好地适应人类的视觉系统，SML 被用来表达边缘特征，它被视为 DTNP 系统的外部输入。SML 的定义为：

$$SML_{l_0}(i, j) = \sum_a \sum_b ML_{l_0}(i + a, j + b) \quad (3.3)$$

其中，滑动窗口大小为 3×3 ，ML 由以下公式给出：

$$\begin{aligned} ML_{l_0}(i, j) = & |2C_{l_0}(i, j) - C_{l_0}(i - 1, j) - C_{l_0}(i + 1, j)| \\ & + |2C_{l_0}(i, j) - C_{l_0}(i, j - 1) - C_{l_0}(i, j + 1)| \end{aligned} \quad (3.4)$$

其中， $C_{l_0}(i, j)$ 是低频 ST 系数在位置 (i, j) 的值。

假设两个 DTNP 系统 Π_{AL} 和 Π_{BL} 具有局部拓扑结构，并且它们与两个多焦点源图像 A 和 B 的低频系数相关联。根据公式（3.3），计算源图像 A 和 B 低频系数的 SML 值，记为 $SML_{l_0}^{AL}(i, j)$ 和 $SML_{l_0}^{BL}(i, j)$ ，并将它们作为 Π_{AL} 和 Π_{BL} 的外部输入。这两个 DTNP 系统从初始状态开始直到达到最大迭代次数而停止， P_{AL} 、 P_{BL} 是与 Π_{AL} 和 Π_{BL} 关联的脉冲矩阵，并

且 $P_{AL} = \left(p_{l_0}^{AL}(i, j) \right)_{m \times n}$ 、 $P_{BL} = \left(p_{l_0}^{BL}(i, j) \right)_{m \times n}$ 。在此特别的说明一下， $p_{l_0}^{AL}(i, j)$ 和 $p_{l_0}^{BL}(i, j)$ 表示神经元 σ_{ij} 在 Π_{AL} 和 Π_{BL} 产生的脉冲累加值。基于这两个脉冲矩阵，我们定义低频 ST 系数的规则如下：

$$C_{l_0}^C(i, j) = \begin{cases} C_{l_0}^{AL}(i, j), & \text{if } p_{l_0}^{AL}(i, j) \geq p_{l_0}^{BL}(i, j) \\ C_{l_0}^{BL}(i, j), & \text{if } p_{l_0}^{AL}(i, j) < p_{l_0}^{BL}(i, j) \end{cases} \quad (3.5)$$

其中 $C_{l_0}^{AL}(i, j)$ 和 $C_{l_0}^{BL}(i, j)$ 表示多焦点图像在位置 (i, j) 的低频 ST 系数， $C_{l_0}^C(i, j)$ 表示融合图像 C 在位置 (i, j) 的低频 ST 系数，并且 $1 \leq i \leq m$ 、 $1 \leq j \leq n$ 。

3.3.2 高频 ST 系数融合规则

通常绝对值较大的高频系数包含图像的边缘和纹理。空间频率 (Spatial Frequency, SF) 可以用来表达图像的空间分布。因此，可用空间频率来表达图像的边缘和纹理特征，并且作为 DTNP 系统的外部输入。SF 定义如下：

$$SF_{lr}(i, j) = \sqrt{RF_{lr}(i, j)^2 + CF_{lr}(i, j)^2} \quad (3.6)$$

其中 RF 和 CF 由以下公式给出：

$$RF_{lr}(i, j) = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [C_{lr}(i, j) - C_{lr}(i, j-1)]^2} \quad (3.7)$$

$$CF_{lr}(i, j) = \sqrt{\frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [C_{lr}(i, j) - C_{lr}(i-1, j)]^2} \quad (3.8)$$

其中 $C_{lr}(i, j)$ 表示在 l 层 r 方向位置为 (i, j) 的高频 ST 系数，并且 $M = N = 3$ 。

假设两个 DTNP 系统 Π_{AH} 和 Π_{BH} 具有局部拓扑结构，并且它们与两个多焦点源图像 A 和 B 的高频系数相关联。基于公式 (3.6)，计算源图像 A 和 B 高频系数的 SF 值，记为 $SF_{lr}^{AH}(i, j)$ 和 $SF_{lr}^{BH}(i, j)$ ，并将它们作为 Π_{AH} 和 Π_{BH} 的外部输入。这两个 DTNP 系统从初始状态开始直到达到最大迭代次数而停止， P_{AH} 、 P_{BH} 是与 Π_{AH} 和 Π_{BH} 关联的脉冲矩阵，并且 $P_{AH} = \left(p_{lr}^{AH}(i, j) \right)_{m \times n}$ 、 $P_{BH} = \left(p_{lr}^{BH}(i, j) \right)_{m \times n}$ ，其中 $p_{lr}^{AH}(i, j)$ 和 $(p_{lr}^{BH}(i, j))$ 表示神经元 σ_{ij} 在 Π_{AH} 和 Π_{BH} 产生的脉冲累加值。基于这两个脉冲矩阵，我们定义高频 ST 系数的规则如下：

$$C_{lr}^C(i, j) = \begin{cases} C_{lr}^{AH}(i, j), & \text{if } p_{lr}^{AH}(i, j) \geq p_{lr}^{BH}(i, j) \\ C_{lr}^{BH}(i, j), & \text{if } p_{lr}^{AH}(i, j) < p_{lr}^{BH}(i, j) \end{cases} \quad (3.9)$$

其中 $C_{lr}^{AH}(i, j)$ 和 $C_{lr}^{BH}(i, j)$ 表示多焦点图像在 l 层 r 方向位置为 (i, j) 的高频系数， $C_{lr}^C(i, j)$ 表示融合图像 C 层 r 方向位置为 (i, j) 的高频 ST 系数。

3.3.3 一致性检验

变换域方法的一个常见缺点是，它们可能无法从多焦点源图像中提取某些重要细节。为了克服这个缺点，我们采用了一种优化模型。Zhang 等人^[91]提出了一种多尺度形态焦点测量（MSMFM）来评估融合图像。假定 $A(i,j)$ 和 $B(i,j)$ 表示源图像 A 和图像 B 在位置 (i,j) 的像素值， $C(i,j)$ 表示初始融合图像 C 在位置 (i,j) 的像素值。计算 MSMFM 的具体细节可以参见文献^[91]，其中两个差异图像被定义为：

$$\begin{aligned} D_1 &= \{D_1(i,j)\}_{m \times n}, \quad D_1(i,j) = C(i,j) - A(i,j) \\ D_2 &= \{D_2(i,j)\}_{m \times n}, \quad D_2(i,j) = C(i,j) - B(i,j) \end{aligned} \quad (3.10)$$

通过 MSMFM 计算两个差异图像 D_1 和 D_2 ，于是可以得到一个二值矩阵 M_1 ，它的定义如下：

$$M_1(i,j) = \begin{cases} 1, & \text{if } \text{MSMFM}(D_1(i,j)) \leq \text{ifMSMFM}(D_2(i,j)) \\ 0, & \text{otherwise} \end{cases} \quad (3.11)$$

由于聚焦区域和非聚焦区域之间的边界存在一些伪影，我们使用一个滤波器来改进初始决策图。具体来说，我们使用一个有效的边缘引导滤波器^[92]将引导图像中的结构信息转换为输入图像的过滤结果。我们将 M_1 经过引导滤波器得到最终的决策矩阵 M_2 ，引导滤波器有两个参数：局部窗口大小 r' 和正则化参数 ϵ ，我们根据决策矩阵 M_2 ，最终的融合图像 $F = \{F(i,j)\}_{m \times n}$ 被定义如下：

$$F(i,j) = M_2(i,j)A(i,j) + (1 - M_2(i,j))B(i,j) \quad (3.12)$$

其中， $F(i,j)$ 表示融合图像 F 在位置 (i,j) 的像素， $A(i,j)$ 和 $B(i,j)$ 分别表示源图像 A 和 B 在位置 (i,j) 的像素，特别说明，在我们实验中设置 $r' = 3$ 和 $\epsilon = 5$ 。

3.4 实验结果与分析

在我们的实验中，使用公开图像数据集中的 20 对多焦点图像作为测试图像，用于评估所提出的和比较的融合方法。图 3.4 显示了这些多焦点源图像，在每一对图像中，左侧为近处聚焦图像，右侧为远处聚焦图像。源图像均为灰色图像且其大小见表 3.1。

我们对所提出的融合方法在图像数据集上进行了评估，并与 7 种传统的融合方法和 3 种基于深度学习的融合方法进行了比较。这 7 种传统方法包括 curvelet transform (CVT)^[93]，multi-resolution singular value decomposition (MSVD)^[94]，Laplacian pyramid with sparse representation (LP-SR)^[95]，guided filtering based fusion (GFF)^[16]，dense scale invariant feature transform (DSIFT)^[96]，boundary finding (BF)^[91]，spatial frequency-motivated pulse

coupled neural networks in nonsubsampled contourlet transform domain (NSCT-PCNN)^[97]。3 种基于深度学习的方法为 CNN^[98], ECNN^[49]和 IFCNN^[51]。

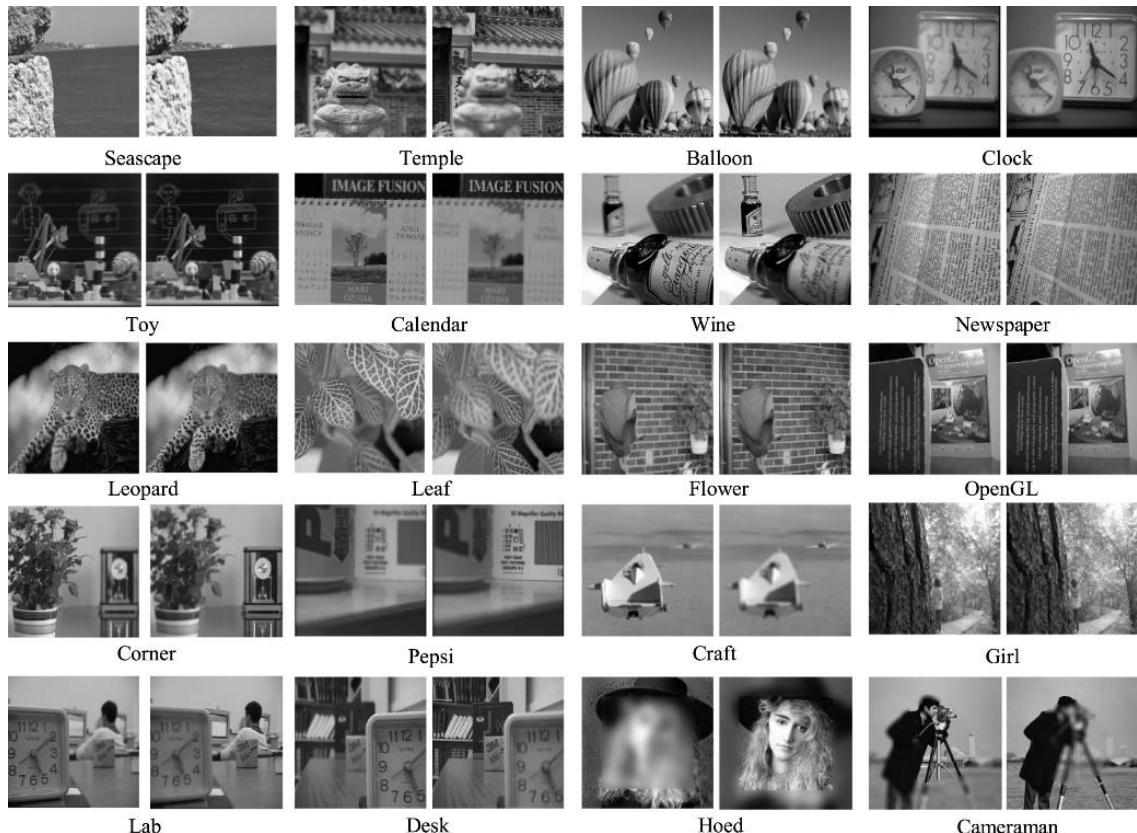


图 3.4 我们实验中使用的多焦点图像数据集

Fig. 3.4 The multi-focus image dataset used in our experiments

3.4.1 评价指标和参数设定

为了比较所提出的融合方法与其它融合方法，提供了主观和客观的比较结果。主观比较是对这些方法融合的图像进行视觉上的比较。实验中使用的客观评价指标包括 Q_{abf} ^[81]、 FMI ^[83]、 Q_w ^[87]、 Q_e ^[87] 和 CE ^[86]。

所提出的方法中有 6 个参数：（1）源图像的大小 $m \times n$; （2）系统的最大迭代次数 t_{max} ; （3）初始阈值 τ_0 ; （4）邻域半径 r ; （5）神经元激发产生的脉冲值 p ; （6）局部权重矩阵 $W_{r \times r}$ 。在实验中所使用的软件平台和版本号是 Matlab 2017b，并在英特尔酷睿 i7-6700 CPU、主频为 3.4 GHz、内存为 16 GB 的主机上运行。对于我们提出的融合方法 $m \times n$ 是根据输入图像所决定的，通过大量的实验对比，运用网格搜索的方法进行了参

数的选择。在实验中我们设置 $r = 3$ 、 $\tau_0 = 1$ 、 $\tau = 1$ 、 $p = 1.1$ 、 $t_{max} = 200$ 并且局部权重矩阵为：

$$W_{3 \times 3} = \begin{pmatrix} 0.7071 & 1 & 0.7071 \\ 1 & 0 & 1 \\ 0.7071 & 1 & 0.7071 \end{pmatrix}$$

实验中给出了 Surfacelet Transform (ST) 的参数：分解层数为 3 层， $pyr_mode = 1$ 、 $lev_array = [-1, 0; 0, -1]$ 、 $HGname = 'ntf'$ 和 $Varargin = 'bo'$ 。特别的说明，所比较的传统方法全部使用原作者论文所设置的参数和原作者所公开的代码在相同平台上实现。对于三种基于深度学习的融合方法的实验平台是 Ubuntu 16.04 操作系统，通过原作者所提供的开源代码在 Inter Xeno Silver 4110 CPU 2.10 GHz 和 NVIDIA Tesla P100 with 16GB RAM 复现。

表 3.1 数据集中的图像大小

Tab. 3.1 The sizes of images in dataset

Images	Seascape	Temple	Balloon	Clock	Toy
Size	580×320	480×516	640×480	512×512	512×512
Images	Calendar	Wine	Newspaper	Leopard	Leaf
Size	260×180	256×256	320×240	480×360	268×204
Images	Flower	OpenGL	Corner	Pepsi	Craft
Size	512×384	512×384	320×240	512×512	160×160
Images	Girl	Lab	Desk	Hoed	Cameraman
Size	636×476	640×480	640×480	256×256	256×256

3.4.2 与传统七种融合方法比较

通过比较所提出的方法和其它方法所融合的图像的视觉质量进行主观评价。图 3.5、图 3.6 分别显示了所提出的方法和其它 7 种方法在 Clock 和 Pepsi 两对多焦点图像上融合的图像。

图 3.5 比较了所提出的方法和其它 7 种方法对尺寸为 512×512 的 Clock 图像生成的融合图像，其中 (a) 和 (b) 为源多焦点图像。从图 3.5 (c) - (i) 可以观察到，CVT、MSVD、LP-SR、GFF、DSIFT、NSCT-PCNN 融合的图像存在一些模糊现象

(见红色矩形框)。我们还可以发现, MSVD 融合的图像存在明显的整体模糊和低对比度; DSIFT 融合的图像中大钟和小钟交界处缺少一些信息。所提出的方法和 BF 具有良好的视觉效果, 包括高清晰度和对比度。

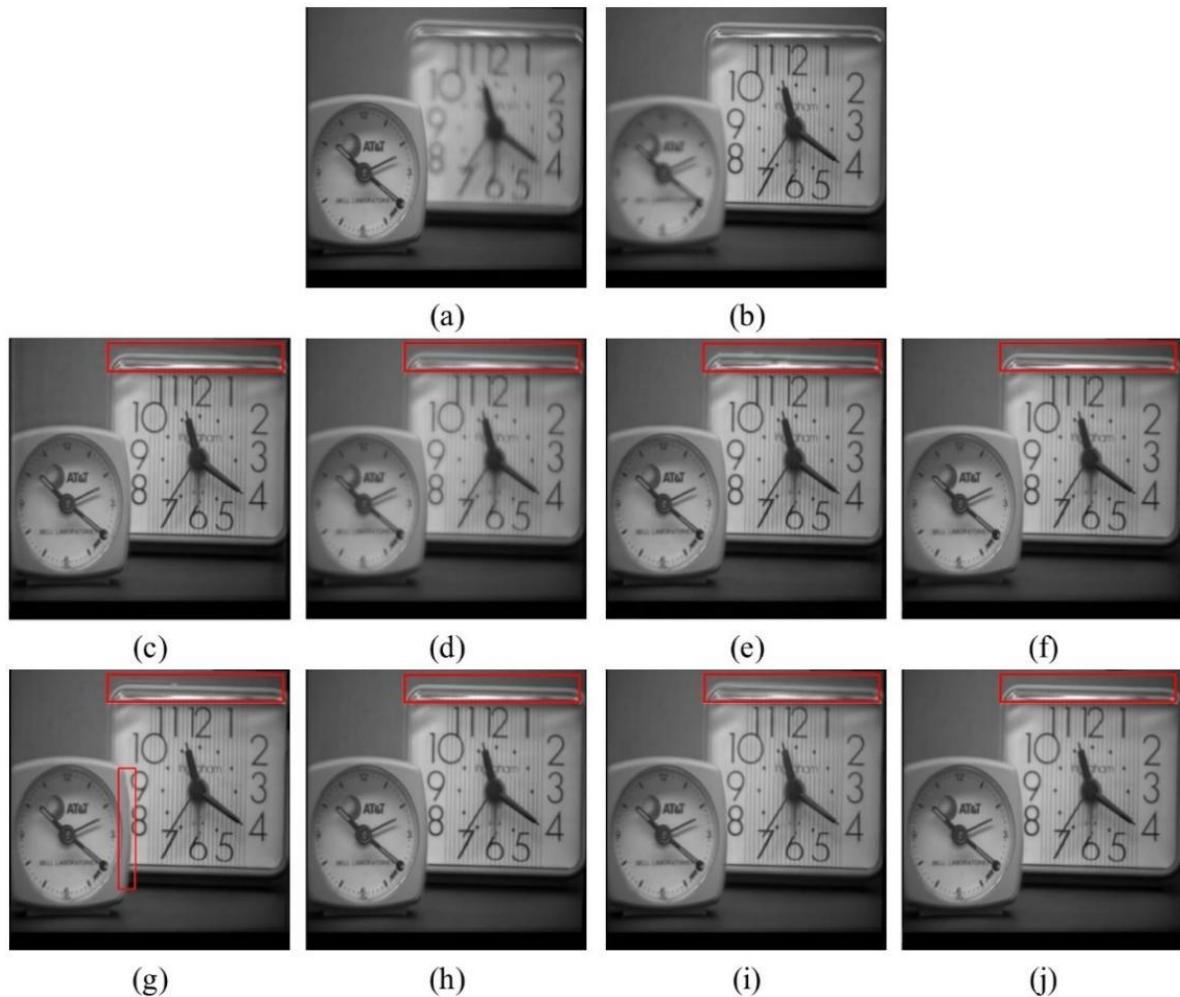


图 3.5 在 Clock 图像上所提出的方法和比较方法的融合结果。(a)和(b)为源图像, (c)-(j)分别为 CVT、MSVD、LP-SR、GFF、DSIFT、BF、NSCT-PCNN 和提出方法的融合结果

Fig. 3.5 Source images and fusion results of the proposed and compared methods on the Clock image: (a) and (b) are source images, and (c)–(j) are the fusion results of CVT, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN and the proposed method, respectively

表 3.2 显示了 8 种融合方法分别对 Clock 和 Pepsi 图像在 5 个指标上的量化比较。在表 3.2 中粗体字标志着对应方法达到了最佳的指标值。

对于 Clock 图像, 从表 3.2 中可以看出:

- (1) 所提出的融合方法在四个指标上获得了最佳值, 即 $Q_{abf} = 0.718$ 、 $FMI = 0.9257$ 、 $Q_e = 0.5782$ 和 $CE = 0.0177$ 。
- (2) GFF 在 Q_w 指标上得到了最好的值, $Q_w = 0.9068$, 而所提出的融合方法达到了较好的值, $Q_w = 0.8873$ 。

表 3.2 所提出的方法与对比的 7 种融合方法在图像 Clock 和 Pepsi 上 5 个指标的对比结果

Tab. 3.2 The comparison results of the proposed method with seven previous fusion methods on images Clock and Pepsi in terms of five metrics

Images	Methods	Q_{abf}	FMI	Q_w	Q_e	CE
Clock	CVT	0.6674	0.9184	0.8848	0.5386	0.0849
	MSVD	0.5695	0.9048	0.8005	0.3436	0.0309
	LP-SR	0.6942	0.9232	0.8175	0.5564	0.1053
	GFF	0.7169	0.9225	0.9068	0.5619	0.0217
	DSIFT	0.7178	0.9249	0.8929	0.5681	0.0168
	BF	0.7174	0.9249	0.8620	0.5611	0.0177
	NSCT-PCNN	0.6685	0.9174	0.8877	0.4949	0.0342
Pepsi	Proposed	0.7178	0.9257	0.8873	0.5782	0.0177
	CVT	0.7657	0.9221	0.9636	0.6020	0.0286
	MSVD	0.6733	0.9089	0.8966	0.4824	0.0273
	LP-SR	0.7695	0.9238	0.9640	0.5585	0.0327
	GFF	0.7833	0.9245	0.9644	0.5318	0.0244
	DSIFT	0.7819	0.9246	0.9627	0.5981	0.0266
	BF	0.7890	0.9247	0.9633	0.6002	0.0280
	NSCT-PCNN	0.7637	0.9194	0.9605	0.5626	0.0287
	Proposed	0.7899	0.9247	0.9634	0.6066	0.0274

对于 Pepsi 图像, 从表 3.2 中我们可以得出以下结论:

- (1) 所提出的融合方法在三个指标上获得了最佳值, 即 $Q_{abf} = 0.7899$ 、 $FMI =$

0.9247 和 $Q_e = 0.6066$ 。

- (2) BF 在 FMI 指标上达到了最佳值, $FMI = 0.9247$ 。
- (3) GFF 在 Q_w 指标上获得了最佳值, $Q_w = 0.9644$, 而提出的融合方法获得了较好的值, $Q_w = 0.9634$ 。
- (4) GFF 在 CE 指标上获得了最好的值, $CE = 0.0244$, 而提出的融合方法获得了较好的值, $CE = 0.0274$ 。

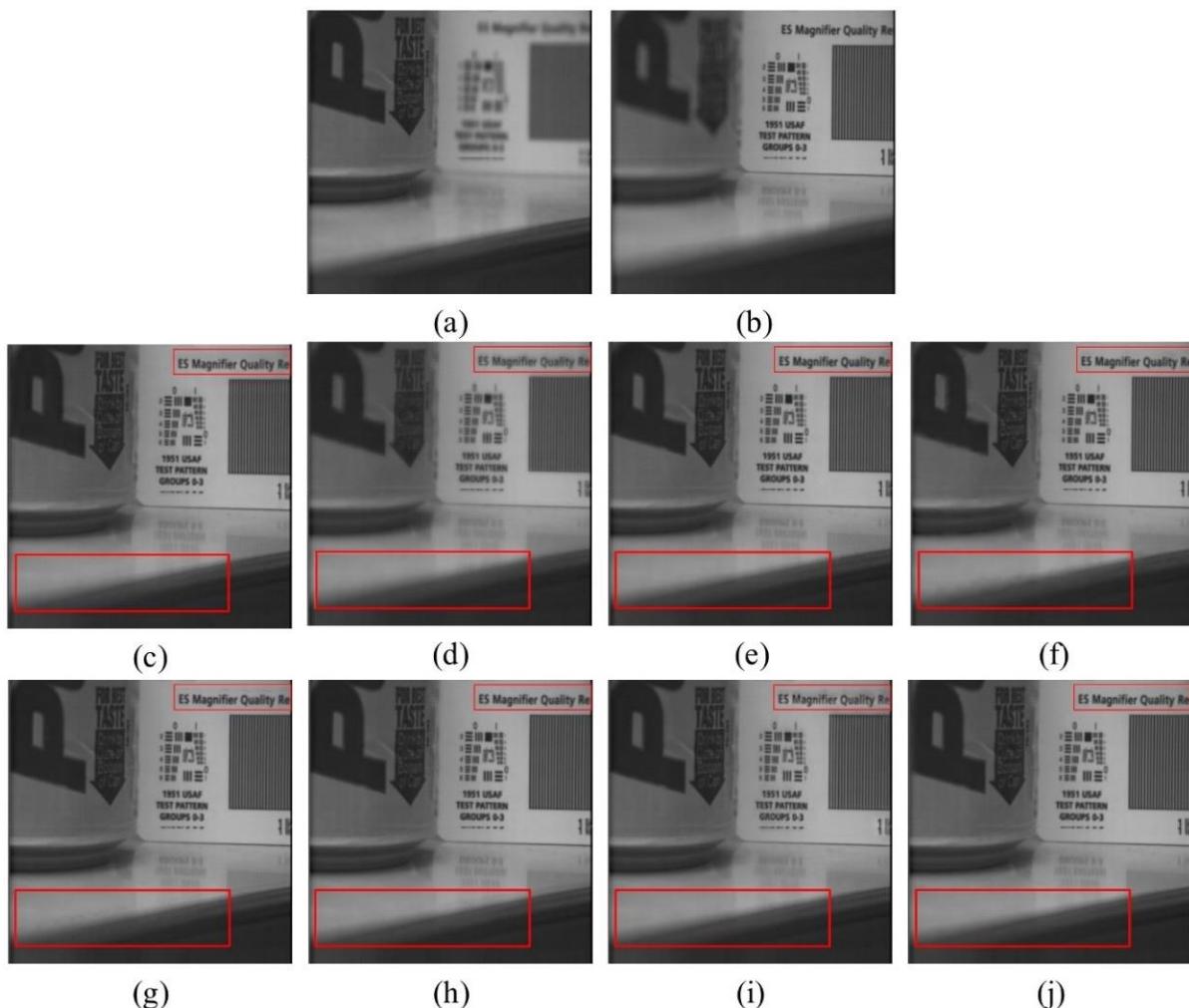


图 3.6 在 Pepsi 图像上所提出的方法和比较方法的融合结果。(a)和(b)为源图像, (c)-(j)分别为 CVT、MSVD、LP-SR、GFF、DSIFT、BF、NSCT-PCNN 和提出方法的融合结果

Fig.3.6 Source images and fusion results of the proposed and compared methods on the Pepsi image: (a) and (b) are source images, and (c)–(j) are the fusion results of CVT, MSVD, LP-SR, GFF, DSIFT, BF, NSCT-PCNN and the proposed method, respectively

图 3.6 比较了所提出的方法和其它 7 种方法对尺寸为 512×512 的 Pepsi 图像所生成的融合图像，其中 (a) 和 (b) 为多焦点源图像。从图 3.6 可以观察到，CVT 融合的图像对比度较低；MSVD 融合的图像有严重的模糊伪影；GFF 和 DSIFT 融合的图像在大框内有很多噪声，DSIFT 融合的图像在小框内有一定的噪声；NSCT-PCNN 融合的图像在小框内有一定的伪影。所提出的方法、LP-SR 和 BF 具有良好的视觉效果。但很难区分哪种方法的视觉效果最好或最差。因此，需要进行客观的评估。

除此之外，图 3.7 还展示了不同融合方法在 20 对图像上对应的 5 个指标，其中横坐标的值对应 20 对图像。如图 3.7 所示，我们提出的融合方法的融合结果在 Q_{abf} 、 FMI 和 CE 指标均达到了最优值， Q_w 和 Q_e 指标也取得了较好的值。

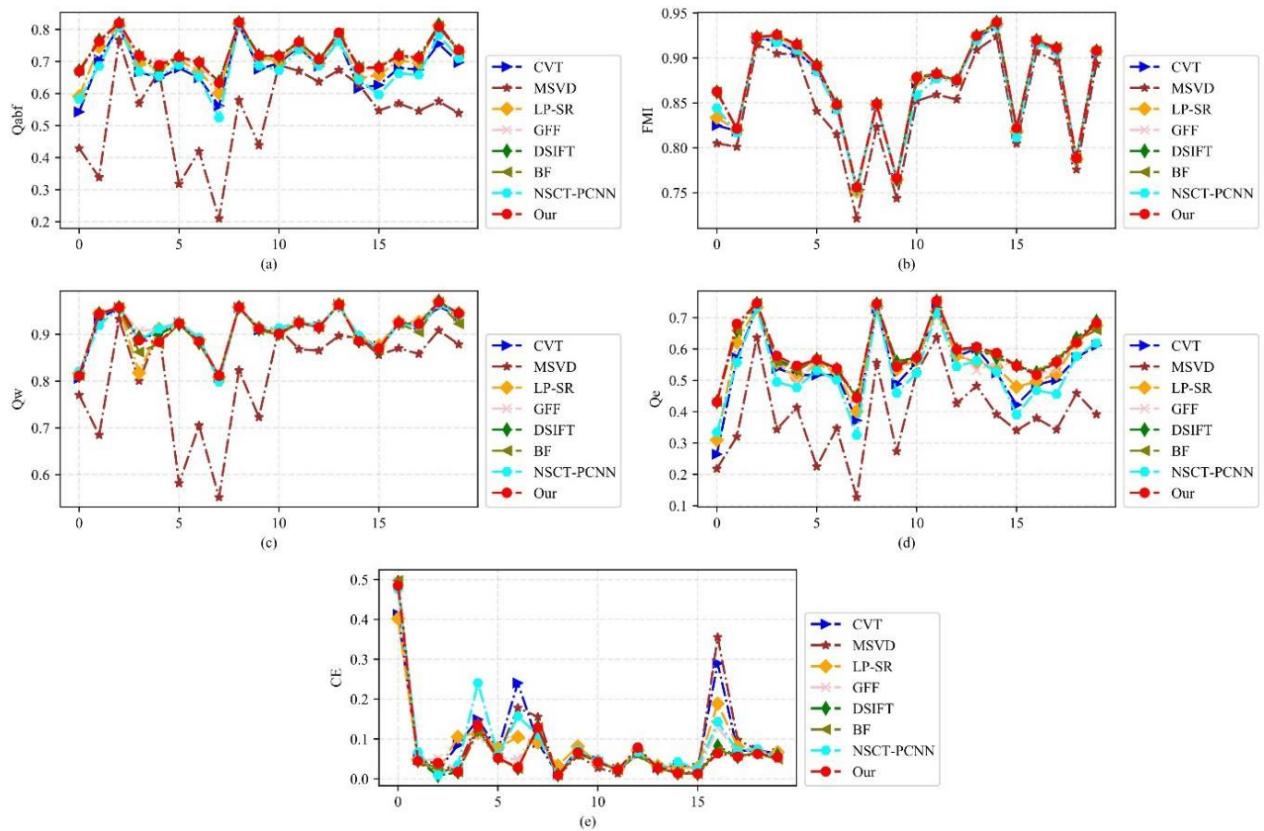


图 3.7 20 对图片在 8 种融合方法上对应的 5 个指标

Fig.3.7 20 pairs of images on 8 fusion methods corresponding to 5 metrics

表 3.3 所提出的方法与 7 种融合方法在数据集上 5 个指标对比结果

Tab. 3.3 The comparison results of the proposed method with seven previous fusion methods on dataset in terms of five metrics

Images	Methods	Q_{abf}	FMI	Q_w	Q_e	CE
Dataset	CVT	0.6848	0.8656	0.9096	0.5416	0.0960
	MSVD	0.5406	0.8475	0.8142	0.3918	0.0970
	LP-SR	0.7097	0.8675	0.9097	0.5618	0.0856
	GFF	0.7266	0.8697	0.9132	0.5833	0.0761
	DSIFT	0.7274	0.8701	0.9101	0.5935	0.0710
	BF	0.7273	0.8703	0.9052	0.5897	0.0705
	NSCT-PCNN	0.6843	0.8651	0.9107	0.5266	0.0906
	Proposed	0.7277	0.8704	0.9096	0.5933	0.0701

表 3.3 列出了所提出的方法和 7 种比较方法在 20 对图像上的 5 个指标平均融合性能的结果，我们可以从表 3.3 得出以下结论：

- (1) 所提出的融合方法实现了三个指标的最佳值，即 $Q_{abf} = 0.7277$ 、 $FMI = 0.8704$ 和 $CE = 0.0701$ 。
- (2) 对于 Q_w 指标，GFF 实现了 $Q_w = 0.9132$ 的最佳值，而提出的方法实现了较好的值， $Q_w = 0.9096$ 。
- (3) 对于 Q_e 指标，DSIFT 实现了 $Q_e = 0.5935$ 的最佳值，而提出的方法实现了次佳值 $Q_e = 0.5933$ 。

3.4.3 与基于深度学习的融合方法比较

最近，基于深度学习的融合方法表现出良好的融合效果。将提出的方法与 CNN、ECNN 和 IFCNN 三种基于深度学习的方法进行了比较。

在数据集中的 20 对多焦点图像上对所提出的方法和比较的方法进行了评估。图 3.8 和图 3.9 分别展示了所提出的方法和三种基于深度学习的方法在 Wine 和 Girl 这两对多焦点图像上融合的图像。

从图 3.8 (c) - (f) 可以观察到，所提出的方法和三种基于深度学习的方法都能达到视觉上良好的融合效果，从视觉上很难区分哪种方法表现最好或最差。因此，需要计

算源图像与所提出的方法和三种基于深度学习的方法所融合的图像之间的差异图像,如图 3.8 (g) - (j) 所示。从差异图像中可以发现,图 3.8 (i) 所示的 IFCNN 融合图像中含有很多冗余信息。这说明 IFCNN 并没有将源图像 Wine 中聚焦的信息完全融合。

从图 3.9 (c) - (f) 可以观察到,所提出的方法和三种基于深度学习的方法在 Girl 图像上均达到了融合目的,很难用肉眼区分这几种融合结果的好坏。但是,可以从差异图像中可以发现,图 3.9 (i) 所示的 IFCNN 融合图像中依然含有很多冗余信息。这说明 IFCNN 并没有将源图像 Girl 聚焦部分的信息完全融合。

然而,无法直观地判断 CNN、ECNN 和所提出的融合方法中哪种方法表现最好或最差。因此,需要进行客观指标的评估。

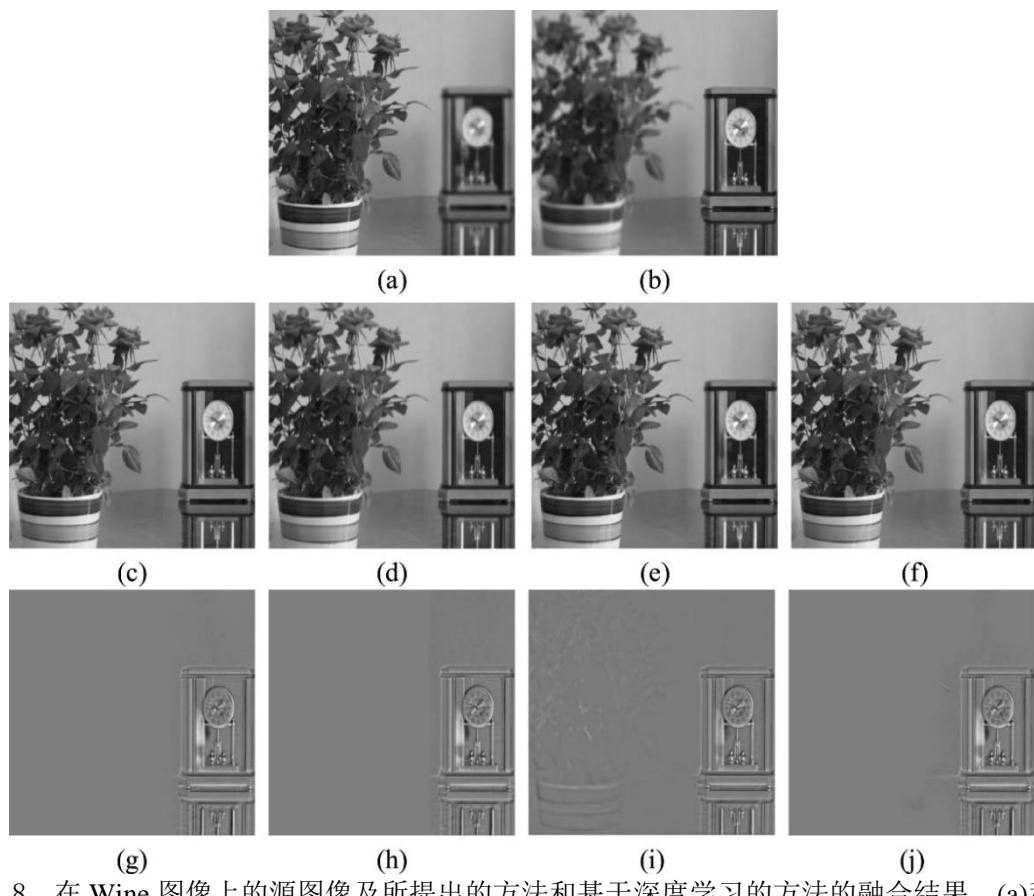


图 3.8 在 Wine 图像上的源图像及所提出的方法和基于深度学习的方法的融合结果。(a)和(b)为多焦点源图像; (c)-(f)分别为 CNN、ECNN、IFCNN 和提出方法的融合结果; (g)-(j)分别为源图像与 CNN、ECNN、IFCNN 和提出方法的差异图

Fig. 3.8 Source images and fusion results of the proposed method and deep-learning-based methods on the Wine image: (a) and (b) are multi-focus source images; (c)–(f) are the fusion results of CNN, ECNN, IFCNN and the proposed method respectively; (g)–(j) are the difference images of source image and CNN, ECNN, IFCNN and the proposed method respectively

表 3.4 显示了四种融合方法分别在 Wine 和 Girl 图像上关于五个指标的客观比较。在表 3.4 中，粗体字标志着相应方法达到了最佳的指标值。对于 Wine 图像从表 3.4 可以看出：

- (1) 所提出的融合方法在三个指标上获得最佳值，即 $Q_{abf} = 0.6979$ 、 $Q_e = 0.5381$ 和 $CE = 0.0280$ 。
- (2) 对于 FMI 指标，ECNN 获得了最佳 $FMI = 0.8480$ ，而所提出的方法获得的次佳值为 $FMI = 0.8478$ 。
- (3) 对于 Q_w 指标，IFCNN-Max 获得了 $Q_w = 0.8995$ 的最佳值，而所提出的方法获得了 $Q_w = 0.8847$ 的次佳值。

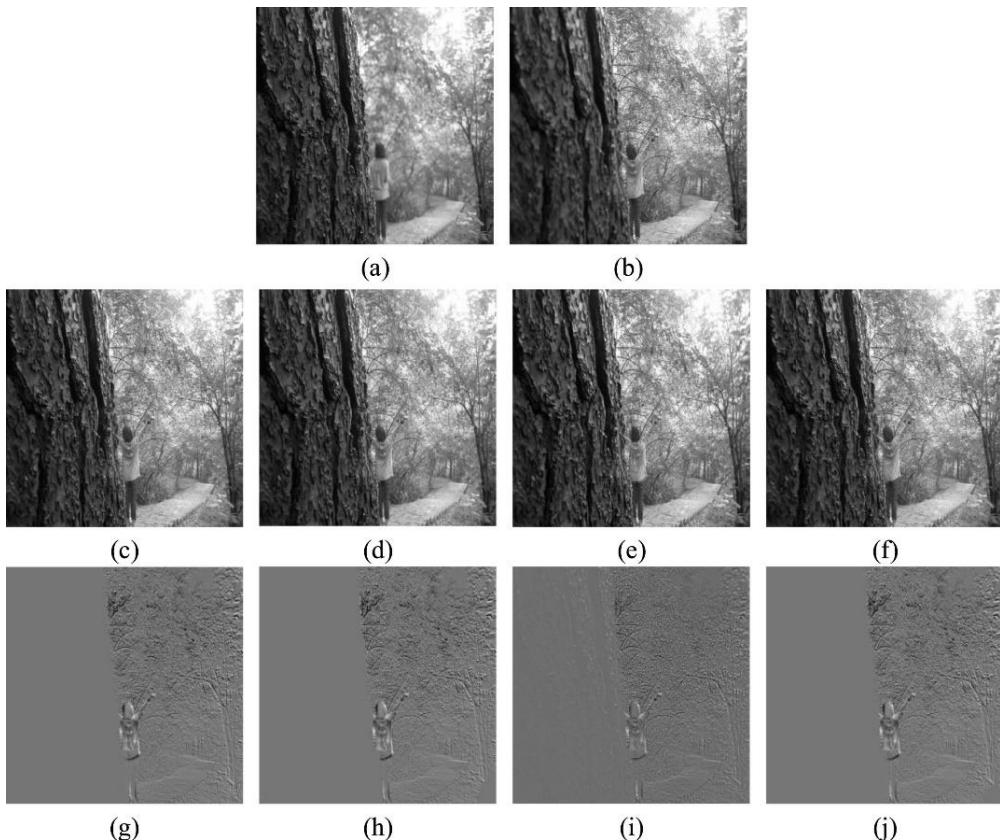


图 3.9 在 Girl 图像上的源图像及所提出的方法和基于深度学习的方法的融合结果。(a)和(b)为多焦点源图像；(c)-(f)分别为 CNN、ECNN、IFCNN 和提出方法的融合结果；(g)-(j)分别为源图像与 CNN、ECNN、IFCNN 和提出方法的差异图

Fig.3.9 Source images and fusion results of the proposed method and deep-learning-based methods on the Girl image: (a) and (b) are multi-focus source images; (c)–(f) are the fusion results of CNN, ECNN, IFCNN and the proposed method respectively; (g)–(j) are the difference images of source image and CNN, ECNN, IFCNN and the proposed method respectively

对于 Girl 图像，从表 3.4 可以看出：

- (1) 所提出的融合方法在三个指标上获得最佳值，即 $Q_{abf} = 0.6812$ 、 $FMI = 0.8219$ 和 $Q_w = 0.8696$ 。
- (2) 从 FMI 指标来看，CNN 获得了最佳 $FMI = 0.8219$ 。
- (3) 对于 Q_e 指标而言，CNN 获得的最佳 $Q_e = 0.5464$ ，而所提出的方法获得了较好的 $Q_e = 0.5461$ 。
- (4) 对于 CE 指标，CNN 和 ECNN 获得了 $CE = 0.0131$ 的最佳值，而所提出的方法获得了 $CE = 0.0132$ 的次佳值。

表 3.4 所提出的方法与三种基于深度学习的融合方法在源图像 Wine 和 Girl 上五个指标的对比结果

Tab. 3.4 The comparison results of the proposed method with three deep-learning-based fusion methods on source images Wine and Girl in terms of five metrics

Images	Methods	Q_{abf}	FMI	Q_w	Q_e	CE
Wine	CNN	0.6966	0.8473	0.8840	0.5365	0.0301
	ECNN	0.6960	0.8480	0.8807	0.5356	0.0287
	IFCNN	0.6444	0.8371	0.8995	0.5050	0.2230
	Proposed	0.6979	0.8478	0.8847	0.5381	0.0280
Girl	CNN	0.6810	0.8219	0.8687	0.5464	0.0131
	ECNN	0.6806	0.8214	0.8662	0.5462	0.0131
	IFCNN	0.6154	0.8099	0.8691	0.4074	0.0187
	Proposed	0.6812	0.8219	0.8696	0.5461	0.0132

表 3.5 列出了所提出的方法和三种比较方法在五个指标方面的平均融合性能的结果。每个客观指标的结果代表数据集中 20 对多焦点图像的平均值。从表 3.5 中可以看出，所提出的融合方法在 4 个指标上取得了最佳值，即 $Q_{abf} = 0.7277$ 、 $FMI = 0.8704$ 、 $Q_e = 0.5933$ 和 $CE = 0.0701$ ，在 Q_w 指标上，CNN 取得了最佳 $Q_w = 0.9099$ ，而所提出的融合方法取得了次佳值 $Q_w = 0.9096$ 。

表 3.5 所提出的方法与 3 种基于深度学习的融合方法 5 个指标在数据集上的对比结果
 Tab. 3.5 The comparison results of the proposed method with three deep-learning-based fusion methods on dataset in terms of five metrics

Images	Methods	Q_{abf}	FMI	Q_w	Q_e	CE
Dataset	CNN	0.7273	0.8703	0.9099	0.5892	0.0724
	ECNN	0.7195	0.8685	0.8996	0.5793	0.0708
	IFCNN	0.6725	0.8595	0.9089	0.5076	0.1085
Proposed		0.7277	0.8704	0.9096	0.5933	0.0701

图 3.10 还展示了在 20 对图像上 CNN、ECNN、IFCNN 和所提出的融合方法关于 5 个指标的结果，其中横坐标的值对应 20 对图像。在图 3.10 中，可以发现所提出的融合方法大部分指标值基本都超过了这三种深度学习融合方法，这表明了我们提出融合方法的在性能上优势。

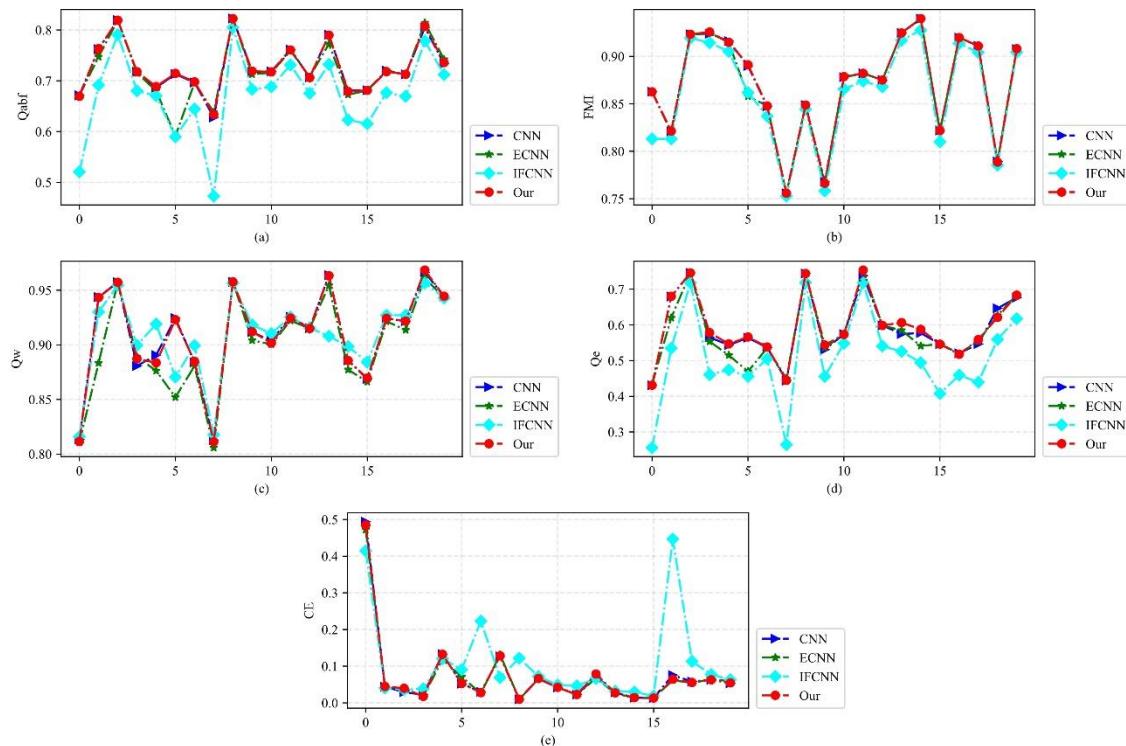


图 3.10 20 对图片在 CNN, ECNN, IFCNN 和我们提出的融合方法上对应的 5 个指标
 Fig.3.10 20 pairs of images on CNN, ECNN, IFCNN and our proposed fusion method corresponding to the 5 metrics

此外，提供了所提出的方法在其余 16 对多焦点图像上的融合结果，包括 Seaside、Temple、Balloon、Toy、Calendar、Newspaper、Leopard、Leaf、Flower、OpenGL、Corner、Craft、Lab、Desk、Hoed 和 Cameraman，如图 3.11 所示。从视觉角度看，这些融合结果表明了所提出的方法有良好的融合效果。



图 3.11 所提出的方法在其他多焦点图像上的融合结果

Fig. 3.11 The fusion results of the proposed method on other multi-focus images

3.4.4 时间效率对比

为了评估计算效率，计算了在 20 对多焦点图像上所提出的和比较的方法的平均执行时间。由于基于深度学习的方法常常需要一个耗时的训练过程，所以 CNN、ECNN 和 IFCNN-Max 没有包括在这个比较中。表 3.6 列出了所提出的方法与其它 7 种方法在平均执行时间上的对比结果。在表 3.6 中，可以观察到以下结果：（1）CVT、MSVD、LP-

SR、GFF、DSIFT 和 BF 的执行时间较低，（2）NSCT-PCNN 的执行时间较高，（3）提出的方法执行时间相对较高。比较结果说明，所提出的方法具有相对较高的计算成本。一般地，与基于深度学习的方法相比，所提出的方法具有高的计算效率。

表 3.6 所提出的方法和其他方法在数据集上的平均执行时间的比较

Tab.3.6 The comparison of average execution time of the proposed approach
and other approaches on dataset

Methods	CVT	MSVD	LP-SR	GFF
Times(s)	0.8918	0.1941	0.0294	0.1874
Methods	DSIFT	BF	NSCT-PCNN	Proposed
Times(s)	4.5603	1.3201	128.5321	23.1298

3.5 本章小结

本章在 Surfacelet Transform 下设计了一种新颖的基于 DTNP 系统的多焦点图像融合方法。提出了一个新的多焦点图像的融合框架。在这个融合框架中，低频和高频 ST 系数的融合规则由 4 个 DTNP 系统控制，分别与两幅多焦点图像中的低频或高频 ST 系数相关联。在实验中，为了证明所提出的融合方法的可用性和有效性，设计了三组比较。

（1）与七种传统融合方法的比较；（2）与三种基于深度学习的融合方法的比较；（3）执行时间的比较。对于前两组比较，考虑了主观评价和定量评价，第三组比较则用于评价计算效率。实验结果表明，所提出的融合方法是一种具有潜力的多焦点图像融合方法。究其原因，主要是由于以下两个方面：（1）ST 可以捕捉到图像更多的细节，提供丰富的互补信息，有利于图像融合；（2）DTNP 系统的灵感来源于神经元的脉冲发射机制和动态阈值机制。在 ST 域中，丰富的互补信息可以触发局部区域中更多的神经元点火，从而使融合图像中的相应区域变得更加清晰。因此，这两种效应的结合使得所提出的融合方法在多焦点图像上表现出良好的融合性能。

4 基于动态阈值神经 P 系统和 NSST 的红外与可见光图像融合

4.1 引言

多传感器数据可以提供有关对象的互补信息，例如，可见光和红外图像。红外线传感器可以获得由红外辐射所捕获的明显的物体信息，如形状、大小和尺寸等。但是，所获得的红外图像存在一些弱点，如分辨率低、场景信息缺失等。可见光成像设备拥有较高的空间分辨率，可以检索到丰富的细节和纹理信息。然而，如果可见光成像设备在恶劣的天气条件下或者物体的空间特征与背景相似的条件下工作，获得的可见光图像在视觉上非常不满意。融合红外和可见光图像是一种有效的技术，它可以为两种传感器提供有用的信息类型，合并后的图像可以有显著的对象信息和丰富的场景信息^[99-101]。如何融合红外和可见光图像一直是图像处理领域的一个活跃的主题，红外与可见光图像融合已经被利用在许多领域，例如，计算机视觉、遥感、视频监控、微观成像和军事任务。本章将探讨如何利用动态阈值神经 P (DTNP) 系统去融合红外与可见光图像。并在一个公开的包含 140 组红外和可见光图像的 TNO 图像数据集上对所提出的方法进行了评估，并与之前的 7 种传统方法和 3 种基于深度学习方法进行比较。

4.2 NSST 低频系数与 DTNP 系统的组合

DTNP 系统作为脉冲神经系统 P 系统的一个扩展，结合了神经元的点火机制和动态阈值机制。

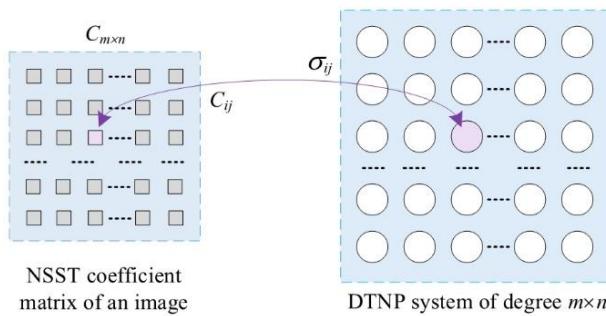


图 4.1 一个 DTNP 系统和相应的 NSST 系数矩阵

Fig. 4.1 A DTNP system Π and a matrix of NSST coefficients

在本章中，为了实现图像融合，将采用具有邻域结构的二维 DTNP 系统来处理 NSST 域的低频系数。具有局部邻域结构的二维 DTNP 系统已经在 3.2.1 小节中详细介绍，但

为了叙述的完整性，这里简要地描述二维 DTNP 系统与 NSST 域的低频系数如何结合。

假设 I 是一个大小为 $m \times n$ 的红外图像或者可见光图像， C 是一个系数矩阵。因此，一个具有邻域结构的 DTNP 系统 Π 被定义为 $m \times n$ 个神经元的矩阵。图 4.1 展示了一个 DTNP 系统 Π 和一个系数矩阵。 $C_{m \times n}$ 表示这个系数矩阵，它包含了 $m \times n$ 个系数， Π 含有 $m \times n$ 个神经元，并且神经元 σ_{ij} 与系数 C_{ij} 相关联， $i = 1, \dots, m, j = 1, \dots, n$ 。对于每一个神经元 σ_{ij} ，它的半径为 r 的邻域被定义为 $\delta_r(\sigma_{ij}) = \{\sigma_{kl} | |l - j| \leq r, |k - i| \leq r, 1 \leq k \leq m, 1 \leq l \leq n\}$ ，其局部连接是与一个权重矩阵 $W_{r \times r}$ 相关联。

神经元 σ_{ij} 有两部分的输入：(1) 从相邻神经元获取的脉冲，(2) 外部输入（对应的 NSST 系数 C_{ij} ）。神经元 σ_{ij} 接收相邻神经元的脉冲的加权值和一个外部输入值，然后根据点火规则将自身产生的脉冲传递给这些周围的神经元。

正如图 4.1 所示，每个神经元 σ_{ij} 包含一个数据单元 u_{ij} 和一个动态阈值单元 τ_{ij} ，还有一个点火规则 $E/(a^u, a^\tau) \rightarrow a^p$ 。该神经元的点火条件为 $E \equiv (u_{ij}(t) \geq \tau_{ij}(t)) \wedge (u_{ij}(t) \geq u) \wedge (\tau_{ij}(t) \geq \tau)$ ，其中 $u_{ij}(t)$ 和 $\tau_{ij}(t)$ 表示神经元 σ_{ij} 在 t 时刻的脉冲值和动态阈值。如果满足这个点火条件，那么神经元 σ_{ij} 将消耗数据单元中的脉冲值 u 和动态阈值单元中的脉冲值 τ ，然后产生一个值为 p 的脉冲并发送给周围的神经元。根据这个点火机制，神经元 σ_{ij} 的状态更新方程为：

$$u_{ij}(t+1) = \begin{cases} u_{ij}(t) - u + C_{ij} + \sum_{\sigma_{kl} \in \delta_r} w_{kl} p_{kl}(t), & \text{if } \sigma_{ij} \text{ fires} \\ u_{ij}(t) + C_{ij} + \sum_{\sigma_{kl} \in \delta_r} w_{kl} p_{kl}(t), & \text{otherwise} \end{cases} \quad (4.1)$$

其中 $u_{ij}(t+1)$ 和 $u_{ij}(t)$ 分别表示神经元 σ_{ij} 在 $t+1$ 和 t 时刻的脉冲值， $p_{kl}(t)$ 表示在 t 时刻从周围神经元 σ_{kl} 收到的脉冲值， w_{kl} 是一个权重， C_{ij} 是一个外部输入。动态阈值的更新公式如下：

$$\tau_{ij}(t+1) = \begin{cases} \tau_{ij}(t) - \tau + p, & \text{if } \sigma_{ij} \text{ fires} \\ \tau_{ij}(t), & \text{otherwise} \end{cases} \quad (4.2)$$

其中 $\tau_{ij}(t+1)$ 和 $\tau_{ij}(t)$ 分别表示神经元 σ_{ij} 在 $t+1$ 和 t 时刻的阈值， p 是神经元 σ_{ij} 点火产生的脉冲。最后，DTNP 算法如下所示。

算法 4.1 DTNP 算法

Algorithm 4.1 The algorithm of DTNP

Algorithm 4.1 DTNP system: $F = \text{DTNP}(C)$

Input:

The NSST coefficient matrix, $C = (C_{ij})_{m \times n}$;

Output:

 算法 4.1 续
 Alg. 4.1 Cont

```

Spiking matrix,  $F = (F_{ij})m \times n$ ;
1: Assign priori parameters:  $t_{max}$ ,  $\tau_0$ ,  $p$ ,  $r$ ,  $W = (w_{ij})r \times r$ ;
   /* Initialization */
2: $U = (u_{ij})_{m \times n} = (0)_{m \times n}$ ;
3: $T = (\tau_{ij})_{m \times n} = (\tau_0)_{m \times n}$ ;
4: $P = (p_{ij})_{m \times n} = (0)_{m \times n}$ ;
5: $F = (F_{ij})_{m \times n} = (0)_{m \times n}$ ;
   /* Working procedure of DTNP systems */
6: for each  $t \in [1, t_{max}]$  do
7:   for each  $i \in [1, m]$  do
8:     for each  $j \in [1, n]$  do
9:       Update  $u_{ij}$  in  $U$  by using Eq. (4.1);
10:      Update  $\tau_{ij}$  in  $T$  by using Eq. (4.2);
11:      if  $u_{ij} \geq \tau_{ij}$  then
12:         $p_{ij} = p$ ;
13:      else:
14:         $p_{ij} = 0$ ;
15:      end if
16:    end for
17:  end for
18:   $F = F + P$ ;
19: end for
   /* Output */
20: return  $F$ ;

```

4.3 红外可见光图像融合框架

本章提出了一个基于 DTNP 系统在 NSST 变换域中的红外和可见光图像融合框架，如图 4.2 所示。整个框架包含 3 部分：（1）NSST 变换，（2）融合规则，（3）NSST 逆变换。

将可见光和红外图像通过 NSST 变换转化为 NSST 系数，然后将这些系数进行融合得到融合的系数。高频系数和低频系数分别采用两种不同的规则进行融合，对于低频系数，使用 DTNP 系统来控制，而高频系数则使用简单的绝对值最大值规则。为此，设计了两个 DTNP 系统 Π_I 和 Π_V 分别来控制红外图像 I_I 和可见光图像 I_V 的低频系数的融合。最后，将融合的 NSST 系数进行逆变换生成一个融合图像 I_F 。

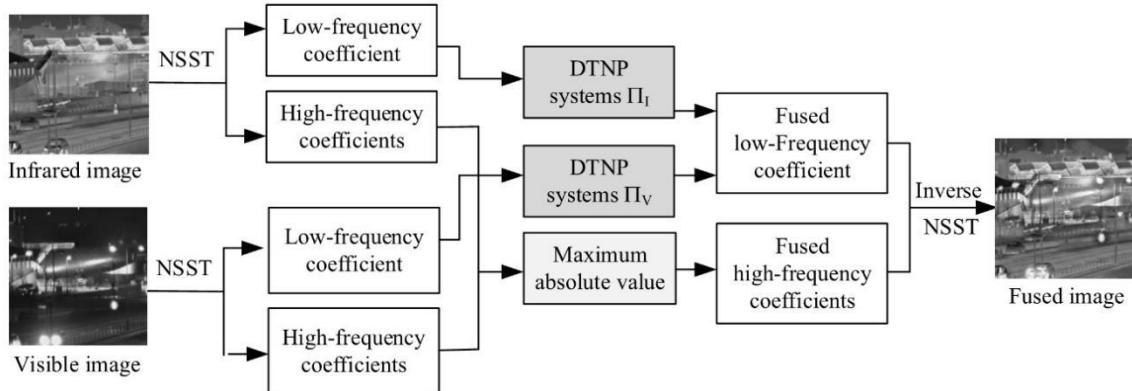


图 4.2 基于 DTNP 系统和 NSST 的可见光和红外图像融合的框架

Fig. 4.2 The presented framework based on DTNP-system and NSST for fusing visible and infrared images

4.3.1 低频 NSST 系数融合规则

通常，低频系数包含了图像的绝大部分能量信息。加权平均法和绝对值最大法是最流行的两种融合方法，但是它们可能会降低融合图像的对比度和损失一些细节信息。为了克服这个缺陷，这个融合框架利用 DTNP 系统来实现了一个低频 NSST 系数的融合规则。

假设 Π_I 和 Π_V 是与红外图像 I_I 和可见光图像 I_V 相关联的两个 DTNP 系统，并且红外图像 I_I 和可见光图像 I_V 的低频 NSST 系数将作为 Π_I 和 Π_V 的外部输入。从初始状态开始，运行这两个 DTNP 系统直到它们达到一个最大迭代次数 t_{max} 为止，系统停机。记 $P_I = (P_{ij}^I)_{m \times n}$, $P_V = (P_{ij}^V)_{m \times n}$, 它们表示 P_I 和 P_V 在执行期间所输出的两个脉冲矩阵。根据这两个脉冲矩阵，可定义低频 NSST 系数的融合规则如下：

$$C_{l_0}^F(i, j) = \begin{cases} C_{l_0}^I(i, j), & \text{if } P_{ij}^I \geq P_{ij}^V \\ C_{l_0}^V(i, j), & \text{if } P_{ij}^I < P_{ij}^V \end{cases} \quad (4.3)$$

其中 $C_{l_0}^I(i, j)$ 和 $C_{l_0}^V(i, j)$ 分别表示红外图像和可见光图像在 (i, j) 处的低频 NSST 系数， $C_{l_0}^F(i, j)$ 表示融合图像在位置 (i, j) 的低频系数， $i = 1, \dots, m$, $j = 1, \dots, n$ 。

4.3.2 高频 NSST 系数融合规则

高频系数表达了图像丰富的轮廓和边缘信息。一般地，具有较大幅度的 NSST 系数是与图像像素亮度突变化有关，并且当高频系数的变化较大时，可以获得丰富的细节信

息。因此，可以采用最大绝对值规则的方法作为高频 NSST 系数的融合规则，这样可以生成包含更多细节信息的融合图像。高频系数的融合规则定义如下：

$$C_{lr}^F(i,j) = \begin{cases} C_{lr}^I(i,j), & \text{if } |C_{lr}^I(i,j)| \geq |C_{lr}^V(i,j)| \\ C_{lr}^V(i,j), & \text{otherwise} \end{cases} \quad (4.4)$$

其中 $C_{lr}^I(i,j)$, $C_{lr}^V(i,j)$ 和 $C_{lr}^F(i,j)$ 分别表示红外图像, 可见光图像和融合图像在 l 层 r 方向位置 (i,j) 的高频系数。

4.4 实验与分析

4.4.1 实验数据集

采用公开的一个包含 140 对红外和可见光图像数据集 TNO 作为测试图像集。由于论文篇幅所限，图 4.3 仅显示了 TNO 图像数据集中的 20 对红外和可见光图像，其中每一对图像左边是红外图像，右边是可见光图像，并且也提供了每对图像的大小。

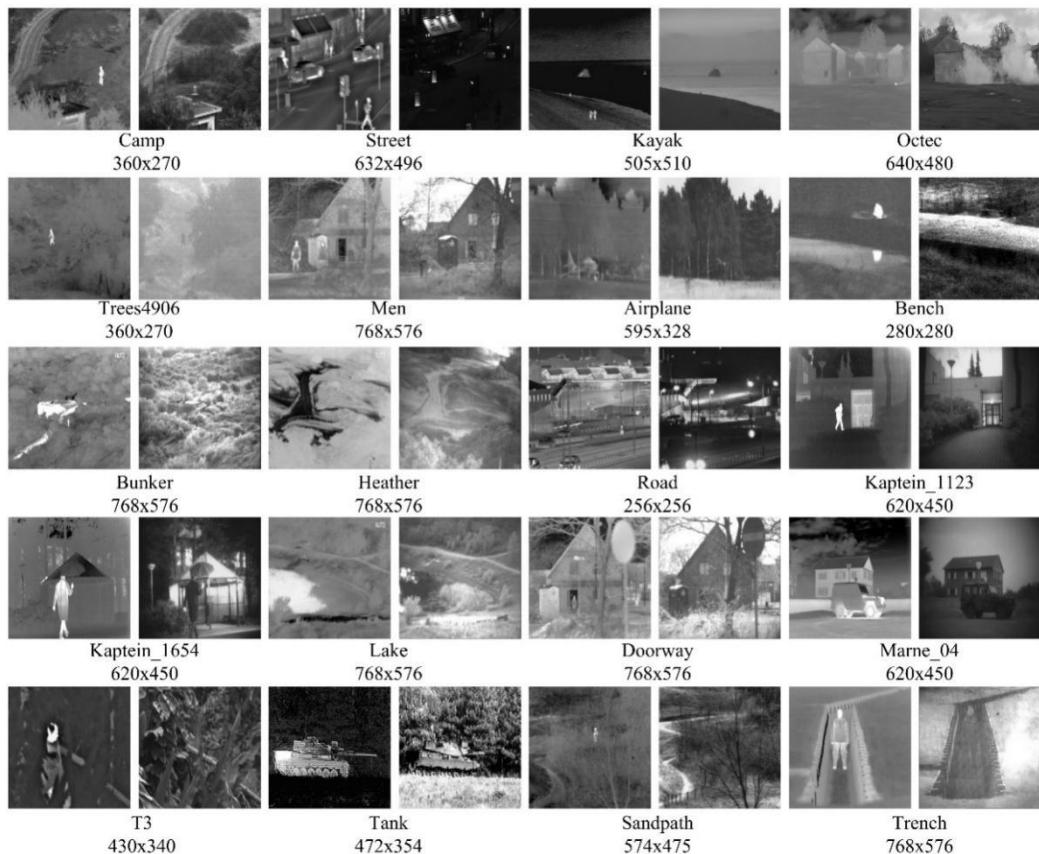


图 4.3 在 TNO 数据中的 20 对红外可见光图像

Fig. 4.3 20 groups of infrared and visible images in TNO image fusion dataset

4.4.2 评价指标

为了定量评估所开发方法和其他方法的融合性能，实验中采用了 5 个评价指标。这些指标分别是特征互信息 (*FMI*)^[83]、信息熵 (*EN*)^[85]、多层次结构相似性 (*MSSSIM*)^[89]、互信息 (*MI*)^[82] 和 Petrovics 度量 (Q_{abf})^[81]。*FMI* 计算图像特征中的相互信息；*EN* 测量的是融合图像中的信息量；*MSSSIM* 计算原始图像和融合图像之间的结构相似度；*MI* 测量从原始图像转移到融合图像中的信息量； Q_{abf} 计算从原始图像转化到融合图像中的边缘信息量。在这 5 项指标中，数值越大表示融合效果越好。

4.4.3 参数设定

实验将对所提出的融合方法在 TNO 图像数据集上进行评估，并与传统的 7 种融合方法和 3 种基于深度学习的融合方法进行比较。这些比较的方法包括：

- (1) 7 种传统的方法：CVT^[93], HMSD^[102], GFF^[16], LRR^[103], DTNP-ST^[104], VSM-WLS^[105], NSCT-SF-PCNN^[97]。
- (2) 3 种深度学习的融合方法：DenseFuse^[50], FusionGAN^[10], IFCNN^[51]。

在实验中所使用的软件平台和版本号是 Matlab 2017b，并在英特尔酷睿 i7-6700 CPU、主频为 3.4 GHz、内存为 16 GB 的主机上运行这些方法。所提出的融合方法的参数设置为： $t_{max} = 110$ 、 $\tau_0 = 1$ 、 $\tau = 1$ 、 $r = 3$ 、 $p = 1.1$ 、权重矩阵 W 如下：

$$W_{3 \times 3} = \begin{pmatrix} 0.7071 & 1 & 0.7071 \\ 1 & 0 & 1 \\ 0.7071 & 1 & 0.7071 \end{pmatrix} \quad (4.5)$$

NSST 采用 '*maxflat*' 金字塔滤波器和 (1 2 3 3) 方向滤波器，7 种传统的方法全部使用原作者提供的代码和原作者设定的实验参数。

三种深度学习的融合方法部署在 Ubuntu 16.04 操作系统，并在 Inter Xeno Silver 4110 2.10GHz CPU 和 16GB RAM 的 NVIDIA Tesla P100 上运行其代码。说明一下，这些代码均是由方法原作者所公开的代码。

对比实验分成 2 组：(1) 与 7 种传统的方法对比；(2) 与 3 种深度学习的方法对比。

4.4.4 定性和定量实验结果比较

在第一组实验中，对所提出的融合方法与 7 种传统的融合方法进行了定量和定性比较。限于篇幅的原因，仅提供了红外图像和可见光图像 Bunker 和 Tank 两对的定性比较结果。

图 4.4 展示了本章提出的融合算法和 7 种传统的方法对 Bunker 图像的融合结果，其中 Bunker 图像的大小为 768×576 。从图 4.4 中可以看出 GFF 融合后的图像没有保留红外图像的信息（见红色矩形框中的区域），因此 GFF 无法对 Bunker 图像得到好的融合结果。观察 CVT 融合的图像，白色矩形框中的物体表现不显著，而 DTNP-ST 融合的图像包含一些模糊的内容（见红色矩形框中的区域）。本章提出的方法，LRR，HMSD，VSM-WLS 和 NSCT-SF-PCNN 可以很好的融合了可见光和红外图像的信息，并且所提出方法的融合图像具有更高的密度和亮度。

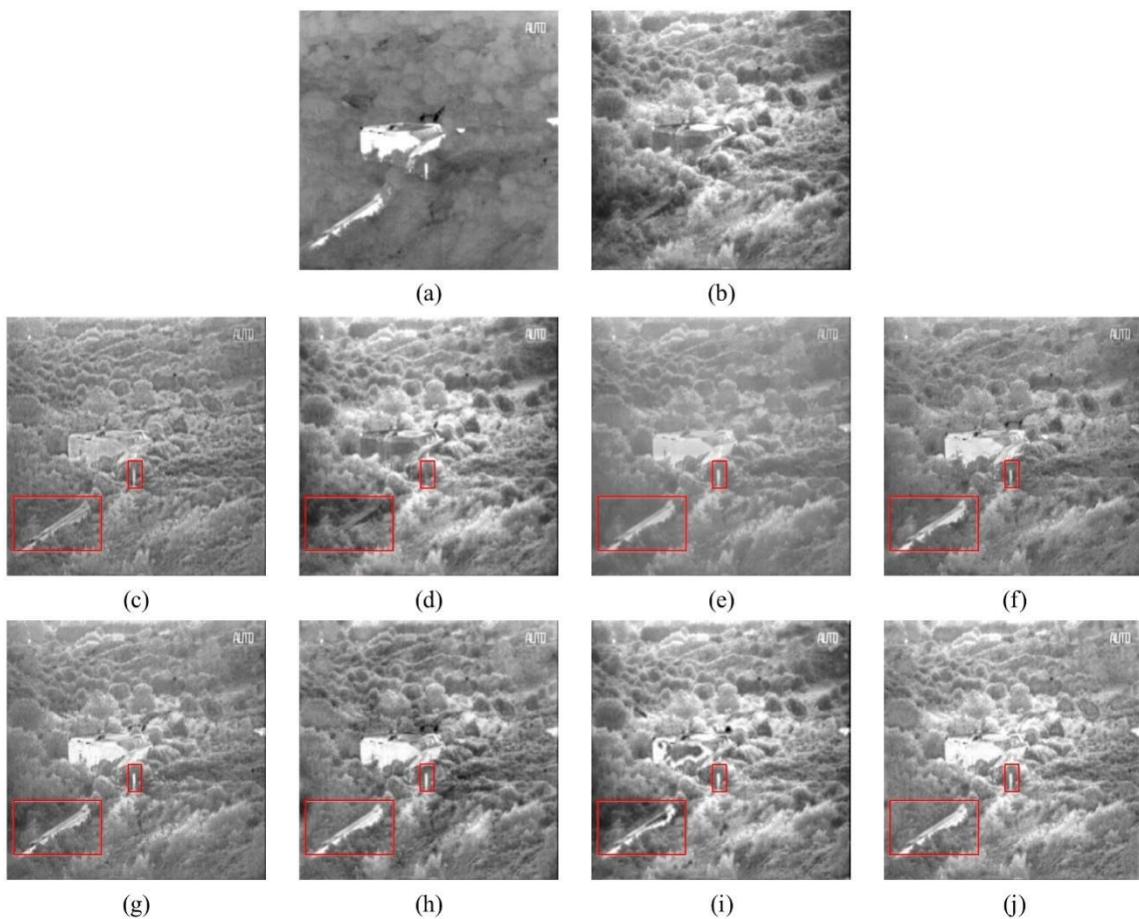


图 4.4 在 Bunker 图像上的融合结果：(a)和(b)为红外和可见光图像；(c)CVT；(d)GFF；(e)LRR；(f)HMSD；(g)VSM-WLS；(h)NSCT-SF-PCNN；(i)DTNP+ST；(j)提出的方法

Fig. 4.4 Fused results on the Bunker image: (a) and (b) are infrared and visible images; (c)CVT; (d) GFF; (e) LRR; (f) HMSD; (g) VSM-WLS; (h) NSCT-SF-PCNN; (i) DTNP+ST;(j) the proposed approach

图 4.5 展示了所提出的方法和 7 种传统方法在 Tank 图像上的融合结果，其中 Tank 图像尺寸为 472×354 。从图 4.5 中可以发现：在由 CVT 融合的图像中，红色矩形框内

的物体表现不显著并且还有一些噪音出现；在由 LRR、HMSD 和 VSM-WLS 融合的图像中，红外图像的内容没有被很好的融合；在由 NSCT-SF-PCNN 融合的图像中包含了大量的黑影，而由 GFF 和 DTNP-ST 合并的图像含有一些黑影（红色小矩形框区域）；所提出的融合方法不仅能突出目标区域的效果，还能保留更多细节和纹理信息。

表 4.1 给出了所提出的方法和其他 7 种方法在 Bunker 图像上关于 5 个指标定量的比较。对于 Bunker 图像，从表 4.1 可以得出以下结论：（1）所提出的方法在 Q_{abf} 、 MI 和 FMI 指标上取得最好的结果；（2）所提出的方法在 $MSSSIM$ 取得第二好的值。

表 4.2 给出了所提出的方法和其他 7 种方法在 Tank 图像上关于 5 个指标定量的比较。对于 Tank 图像，可以得出的结论是：（1）所提出的方法在 MI 、 FMI 、 $MSSSIM$ 和 EN 指标方面的结果最好；（2）所提出的方法在 Q_{abf} 指标上取得了第二好的值。

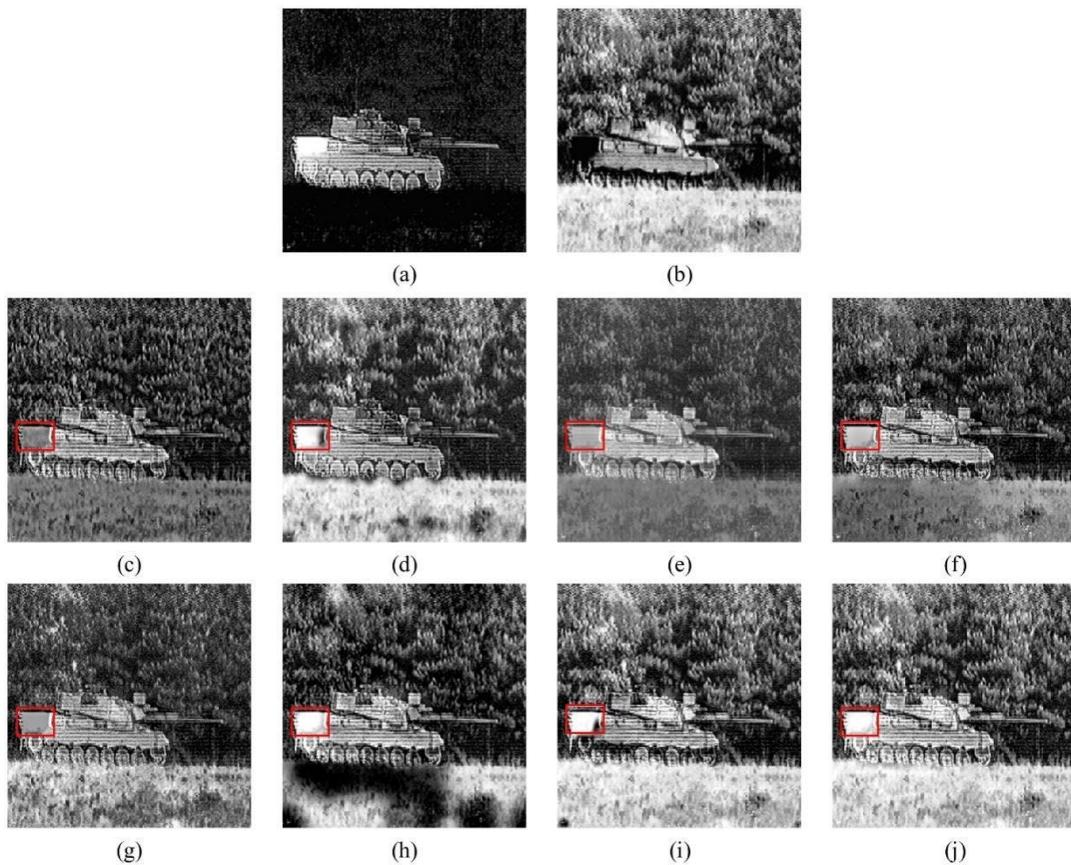


图 4.5 在 Tank 图像上的融合结果：(a)和(b)为红外和可见光图像；(c)CVT；(d)GFF；(e)LRR；(f)HMSD；(g)VSM-WLS；(h)NSCT-SF-PCNN；(i)DTNP+ST；(j)提出的方法。

Fig. 4.5 Fused results on the Tank image: (a) and (b) are infrared and visible images; (c)CVT; (d) GFF; (e) LRR; (f) HMSD; (g) VSM-WLS; (h) NSCT-SF-PCNN; (i) DTNP+ST;(j) the proposed approach

表 4.1 所提出的方法与七种其他方法在 Bunker 图像上的 5 个指标比较

Tab. 4.1 The comparisons of the presented approach with other seven approaches on image Bunker for five metrics

Images	Method	Q_{abf}	MI	FMI	MSSSIM	EN
Bunker	CVT	0.5648	13.8493	0.9274	0.8942	6.9247
	GFF	0.2189	14.0528	0.9302	0.9063	7.4764
	LRR	0.3953	13.6305	0.8936	0.8013	6.8152
	HMSD	0.4775	13.9980	0.9233	0.8643	6.9990
	VSM-WLS	0.5030	14.2152	0.9029	0.9162	7.1076
	NSCT-SF-PCNN	0.5617	14.3368	0.9287	0.9064	7.1684
	DTNP+ST	0.4168	14.3978	0.9303	0.9306	7.4989
	Proposed	0.5781	14.4056	0.9317	0.9232	7.2028

表 4.2 所提出的方法与七种其他方法在 Tank 图像上的 5 个指标比较

Tab. 4.2 The comparisons of the presented approach with other seven approaches on image Tank for five metrics

Images	Method	Q_{abf}	MI	FMI	MSSSIM	EN
Tank	CVT	0.4160	14.8139	0.8177	0.8930	7.4069
	GFF	0.5511	15.8329	0.8141	0.7775	7.9164
	LRR	0.2997	14.6311	0.8043	0.7706	7.3156
	HMSD	0.4464	15.2083	0.8169	0.9037	7.6041
	VSM-WLS	0.3582	15.4866	0.7981	0.8566	7.7433
	NSCT-SF-PCNN	0.3939	14.8062	0.8126	0.8697	7.4031
	DTNP+ST	0.4437	15.7999	0.8203	0.9146	7.8999
	Proposed	0.4513	15.8754	0.8212	0.9525	7.9377

为了评估提出的方法和 7 种传统的方法在整个数据集上的表现，表 4.3 提供了它们在 140 对红外可见光图像上关于 5 个指标的定量比较结果。需要说明一点的是，表 4.3 每一种指标的定量结果是 140 对图像上指标的平均值。从表 4.3 我们可以发现：（1）所提出的方法取得了最好的 Q_{abf} , MI 和 FMI 值；（2）所提出的方法在 EN 指标上取得了较好的值。

表 4.3 所提出的方法与七种其他方法在整个数据集上的 5 个指标比较

Tab. 4.3 The comparisons of the presented approach with other seven approaches on all 140 groups of images in TNO image dataset for five metrics

Images	Method	Q_{abf}	MI	FMI	MSSSIM	EN
Dataset (140 images)	CVT	0.4623	12.8690	0.8817	0.9096	6.4345
	GFF	0.2915	13.2833	0.8810	0.7464	6.6417
	LRR	0.4280	12.8288	0.8716	0.9125	6.4144
	HMSD	0.4765	13.5038	0.8788	0.9234	6.7534
	VSM-WLS	0.4790	13.2071	0.8716	0.9374	6.6035
	NSCT-SF-PCNN	0.4556	13.0391	0.8813	0.8316	7.0196
	DTNP+ST	0.4293	13.1665	0.8747	0.8036	7.0883
Proposed		0.4868	13.5058	0.8844	0.9212	6.7529

4.4.5 与三种基于深度学习的融合方法的比较

第二组实验被用来评估所提出的方法和三种基于深度学习（DL）方法的性能。

图 4.6 是所提出的方法和 3 种基于 DL 的融合方法对 Bunker 图像进行融合的结果图像。从图 4.6 中可以看到，由 FusionGAN 融合的图像亮度低，并且融合图像模糊。所提出的方法、IFCNN 和 DenseFuse 可以很好地融合红外图像和可见光图像的信息，并且从图 4.6 中还可以发现，所提出的方法融合后的图像显示出较高的密度和亮度。

表 4.4 所提出的方法与 3 种深度学习方法在 Bunker 图像上的 5 个指标比较

Tab. 4.4 The comparisons of the presented approach with three fusion approaches based on deep learning on image Bunker for five metrics

Images	Method	Q_{abf}	MI	FMI	MSSSIM	EN	
Bunker	FusionGAN	0.1658	12.8987	0.8945	0.4492	6.4494	
	IFCNN	0.5380	14.1303	0.9050	0.9099	7.0651	
	$\lambda = 1e^0$	0.4197	13.5976	0.9057	0.8351	6.7988	
	DenseFuse	$\lambda = 1e^1$	0.4272	13.6011	0.9033	0.8351	6.8005
		$\lambda = 1e^2$	0.4166	13.5839	0.9039	0.8340	6.7919
		$\lambda = 1e^3$	0.4275	13.5997	0.9018	0.8344	6.7998
Proposed		0.5781	14.4056	0.9317	0.9232	7.2028	

表 4.4 给出了所提出的方法和 3 种基于 DL 方法在 Bunker 图像上关于 5 个指标的定量比较。对于 Bunker 图像，所提出的方法取得了五个指标的最佳结果。

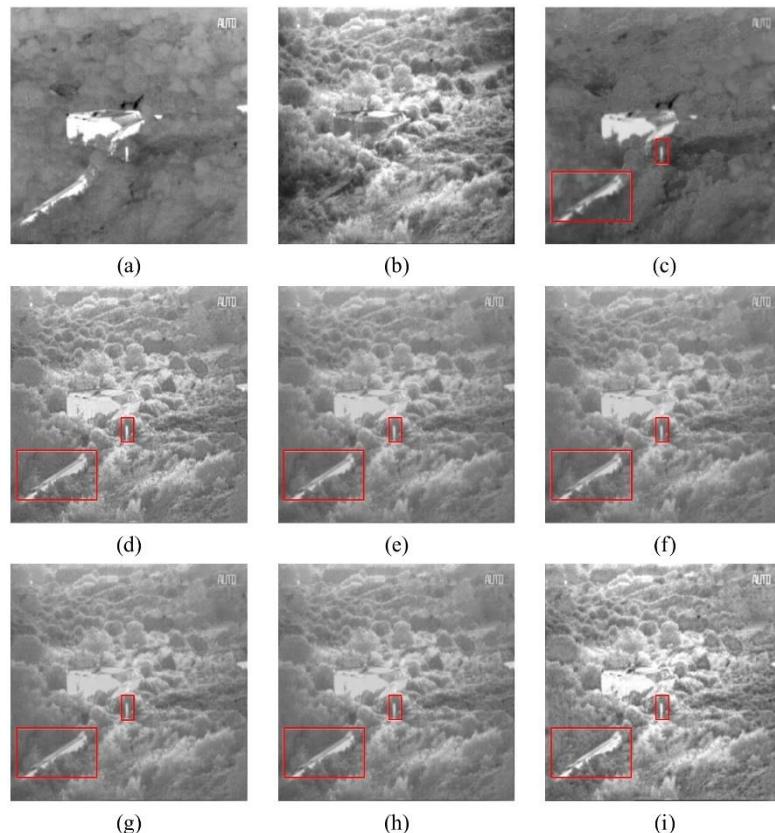


图 4.6 在 Bunker 图像上的融合结果：(a)和(b)是原始红外和可见光图像；(c)FusionGAN；(d)IFCNN；(e)-(h)DenseFuse($\lambda=1e0$, $\lambda=1e1$, $\lambda=1e2$, 和 $\lambda=1e3$)；(i)所提出的方法。

Fig. 4.6 Original images and fused results of the presented and three DL-based approaches on the Bunker image: (a) and (b) are infrared and visible images; (c) FusionGAN; (d) IFCNN; (e)-(h) DenseFuse ($\lambda = 1e0$, $\lambda = 1e1$, $\lambda=1e2$ and $\lambda=1e3$); (i) the proposed approach

图 4.7 是所提出的方法和 3 种基于 DL 的融合方法对 Tank 图像进行融合的结果图像。从图 4.7 中可以看到，通过 FusionGAN 方法融合的图像具有较低的密度和亮度。所提出的方法、IFCNN 和 DenseFuse 所融合的图像有突出目标区域，并保留了更多的细节和纹理信息。

表 4.5 给出了所提出的方法和 3 种基于 DL 方法在 Tank 图像上关于 5 个指标的定量比较。对于 Tank 图像，所提出的方法取得最佳的 *MI*、*FMI*、*MSSSIM*、*EN* 值。

表 4.5 所提出的方法与 3 种深度学习方法在 Tank 图像上的 5 个指标比较

Tab. 4.5 The comparisons of the presented approach with three fusion approaches based on deep learning on image Tank for five metrics

Images	Method	Q_{abf}	MI	FMI	MSSSIM	EN
Tank	FusionGAN	0.4593	12.6339	0.7923	0.3660	6.3169
	IFCNN	0.4249	15.5128	0.8089	0.9297	7.7564
	DenseFuse	$\lambda = 1e^0$	0.4615	15.4956	0.8183	0.8783
	DenseFuse	$\lambda = 1e^1$	0.4754	15.6235	0.8155	0.8749
	DenseFuse	$\lambda = 1e^2$	0.4674	15.5040	0.8172	0.8793
	DenseFuse	$\lambda = 1e^3$	0.4845	15.5719	0.8184	0.8792
	Proposed	0.4513	15.8754	0.8212	0.9525	7.9377

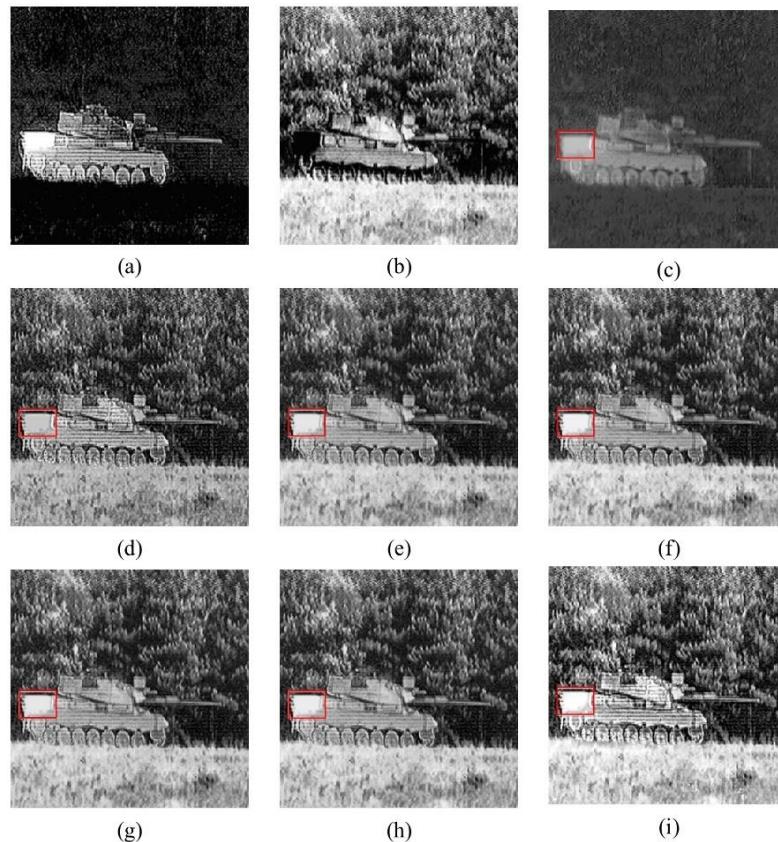


图 4.7 在 Tank 图像上的融合结果: (a)和(b)是原始红外和可见光图像; (c)FusionGAN;

(d)IFCNN; (e)-(h)DenseFuse($\lambda=1e0$, $\lambda=1e1$, $\lambda=1e2$, 和 $\lambda=1e3$); (i)所提出的方法。Fig. 4.7 Original images and fused results of the presented and three DL-based approaches on the Tank image: (a) and (b) are infrared and visible images; (c) FusionGAN; (d) IFCNN; (e)-(h) DenseFuse ($\lambda = 1e0$, $\lambda = 1e1$, $\lambda = 1e2$ and $\lambda = 1e3$); (i) the proposed approach

表 4.6 提供了所提出的方法和 3 种基于深度学习的方法在 140 对红外可见光图像上关于 5 个指标的定量比较结果。从表 4.6 我们可以发现：（1）所提出的方法取得了最佳的 Q_{abf} , MI , FMI 和 EN 值；（2）所推荐的方法在 $MSSSIM$ 指标上取得了次佳值。

表 4.6 所提出的方法与 3 种深度学习方法在整个数据集上的 5 个指标比较

Tab. 4.6 The comparisons of the presented approach with three fusion approaches based on deep learning in TNO image dataset for five metrics

Images	Method	Q_{abf}	MI	FMI	$MSSSIM$	EN
Dataset (140 images)	FusionGAN	0.2353	12.9513	0.8659	0.7351	6.4757
	IFCNN	0.4751	13.0868	0.8741	0.9219	6.5434
	$\lambda = 1e^0$	0.4357	12.6811	0.8803	0.8803	6.3405
	$\lambda = 1e^1$	0.4419	12.7193	0.8804	0.8743	6.3597
	$\lambda = 1e^2$	0.4350	12.6760	0.8797	0.8798	6.3308
	$\lambda = 1e^3$	0.4422	12.7188	0.8800	0.8739	6.3594
Proposed		0.4868	13.5058	0.8844	0.9212	6.7529

我们还提供了所提出的方法在 20 对红外和可见光图像的融合图像结果，如图 4.8 所示。



图 4.8 所提出的方法对 20 对红外和可见光图像进行融合的图像

Fig. 4.8 The images merged by the presented approach on 20 groups of infrared and visible images

4.5 本章小结

本章提出的方法本质上也是一种基于 MST 的方法。基于 MST 的方法是一类流行的融合方法。上面的定性和定量的实验结果表明，所提出的方法明显优于这些传统的基于 MST 的方法。

基于 DL 的方法是一类先进的融合方法，并且它们有超过传统的基于 MST 的融合方法的良好性能。从对比实验结果可以发现：IFCNN 和 DenseFuse 是这些基于深度学习的方法中最好的两种。注意到，与 IFCNN 和 DenseFuse 相比，所提出的融合方法可以接近甚至优于它们的融合性能。这些对比结果表明，DTNP 系统大大地提高基于 MST 的融合性能，凸显了 DTNP 系统在融合红外与可见光图像方面的优势。因此，DTNP 系统可以为红外和可见光图像提供一种有吸引力的融合方法。

5 基于动态阈值神经 P 系统和 NSCT 的多模态医学图像融合

5.1 引言

多模态医学图像融合技术是近年来出现的研究热点^[106-107]，图像融合将多模态传感器的互补信息结合起来，以增强人眼的视觉感知能力或相互补充各图像的局限性。多模态医学影像是多模态影像的重要类型，每种影像方式都有其局限性并显示特定的信息。磁共振成像(MRI)可显示人体内部结构，如腹部、肝脏、胰腺等平滑组织，而计算机断层扫描(CT)可突出显示骨性结构和其他解剖部位，分辨率高。正电子发射断层扫描(PET)和单光子发射计算机断层扫描(SPECT)图像可提供与代谢有关的功能信息。但这些图像常以伪彩色显示，且通常分辨率较低^[107]。这些多模态图像可以提供互补的信息。为了达到更高的诊断准确率，许多研究将同一患者不同模式获得的图像进行综合分析，从而发展了多模式医学图像融合技术。本章提出了一种基于 DTNP 系统的多模态医学图像融合方法，将在 12 对多模态医学图像组成的公开数据集上进行评估，并且与 9 种传统的融合方法和 4 种基于深度学习的融合方法进行比较分析。

5.2 INSML 特征与 DTNP 的组合

在 3.2.1 小节中，已经详细的介绍了具有邻域结构的二维 DTNP 系统。由于 INSML 特征可以更好地触发 DTNP 中神经元点火，所以在医学图像融合中，使用了 INSML 特征作为 DTNP 的外部输入。假设 I 为一个 $m \times n$ 的医学图像， $INSML_{m \times n}$ 是图像的特征矩阵，DTNP 系统 Π 包含 $m \times n$ 个神经元，如图 5.1 所示，其中 $INSML_{ij}$ 是神经元 σ_{ij} 的外部输入，并且每个神经元 σ_{ij} 在邻域 δ_r 中有一组局部连接，其中 $\delta_r(i, j) = \{\delta_{kl} | |k - i| \leq r, |l - j| \leq r\}, 1 \leq i \leq m, 1 \leq j \leq n$ 。

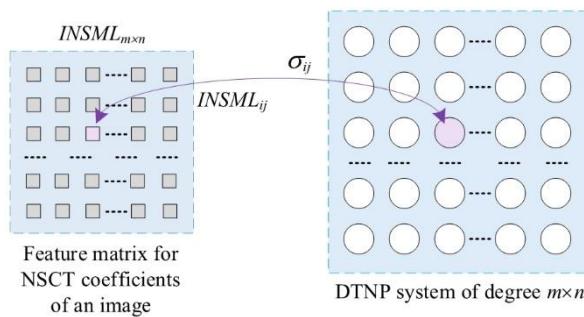


图 5.1 一个 DTNP 系统 Π 和对应的 NSCT 系数特征矩阵

Fig. 5.1 A DTNP system Π and the corresponding feature matrix of NSCT coefficients

在 DTNP 系统 Π 中神经元 σ_{ij} 的点火规则形如 $E/(\alpha^u, \alpha^\tau) \rightarrow \alpha^p$, 其中神经元的点火条件为 $E \equiv (u_i(t) \geq \tau_i(t)) \wedge (u_i(t) \geq u) \wedge (\tau_i(t) \geq \tau)$ 。如果满足点火条件, 则神经元 σ_{ij} 点火, 意味着数据单元中脉冲值 u 和动态阈值单元中阈值 τ 将被消耗, 然后产生值为 p 的脉冲并发送给邻近的神经元。根据这个脉冲机制, 神经元 σ_{ij} 的状态方程为:

$$u_{ij}(t+1) = \begin{cases} u_{ij}(t) - u + V_{ij} + \sum_{\sigma_{kl} \in \delta_r} w_{kl} P_{kl}(t), & \text{if } \sigma_{ij} \text{ fires} \\ u_{ij}(t) + V_{ij} + \sum_{\sigma_{kl} \in \delta_r} w_{kl} P_{kl}(t), & \text{otherwise} \end{cases} \quad (5.1)$$

$$\tau_{ij}(t+1) = \begin{cases} \tau_{ij}(t) - \tau + p, & \text{if } \sigma_{ij} \text{ fires} \\ \tau_{ij}(t), & \text{otherwise} \end{cases} \quad (5.2)$$

其中, $p_{kl}(t)$ 是神经元 σ_{ij} 从周围神经元 σ_{kl} 接收到的脉冲值, $w_{kl}(t)$ 是局部权重, $V_{ij} = INSML_{ij}$ 是外部输入, p 是神经元 σ_{ij} 点火产生的脉冲值。

在初始状态, 设置每一个神经元 σ_{ij} 的 $u_{ij}(0) = v_{ij}(0) = 0$, $\tau_{ij}(0) = \tau_0$, 并且分配一个局部权重矩阵 $W_{r \times r}$ 。然后, DTNP 系统 Π 从初始状态和一个外部输入 $INSML_{m \times n}$ 开始执行, 直到达到最大迭代次数 t_{max} 而停止。在迭代过程中将每个神经元 σ_{ij} 产生的脉冲之和视为其输出, 并将它作为多模态医学图像融合的控制信号。为了简单起见, 在应用点火规则时采用最大消耗策略, 即设置 $u = u_{ij}$ 和 $\tau = \tau_{ij}$ 。

5.3 多模态医学图像的融合框架

本章提出了一种基于 DTNP 系统在 NSCT 变换域的多模态医学图像融合框架, 如图 5.2 所示。所提出的融合框架包括三个部分, 即 (1) NSCT 变换, (2) 融合规则和 (3) 逆 NSCT 变换。在图 5.2 中, 源图像 A 和 B 为多模态医学图像。

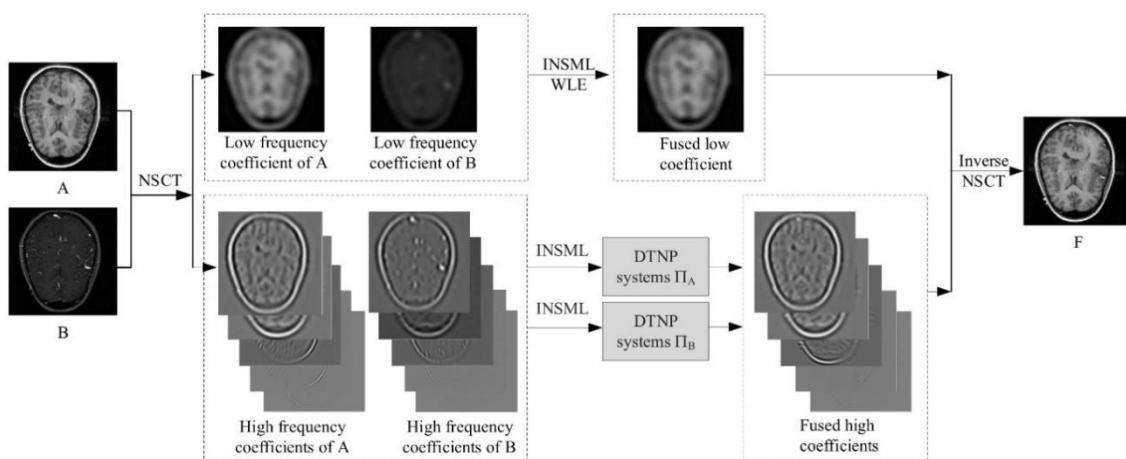


图 5.2 基于 DTNP 系统在 NSCT 变换域中的融合框架

Fig. 5.2 Proposed fusion framework based on DTNP systems in the NSCT domain

使用 NSCT 变换将多模态医学图像分解成低频系数和高频系数，然后将低频和高频系数进行融合生成融合后的 NSCT 系数，其中低频和高频系数分别使用了不同的融合规则进行融合。高频 NSCT 系数采用基于 DTNP 系统的融合规则进行融合，而低频 NSCT 系数采用 INSML-WLE 规则。在整个融合框架里，DTNP 系统是融合框架的关键组成部分，两个 DTNP 系统 Π_A 和 Π_B 分别对多模态医学图像 A 和多模态医学图像 B 的高频系数进行融合，然后将融合后的 NSCT 系数进行转换。最后，利用逆 NSCT 变换重建融合图像。基于这个融合框架提出的多模态医学图像融合方法包括以下步骤：

- (1) NSCT 变换：使用 NSCT 变换将多模态医学图像分解为 NSCT 系数，于是可以得到多模态医学图像的低频 NSCT 系数和高频的 NSCT 系数。
- (2) 低频 NSCT 系数的融合：利用 INSML-WLE（公式（5.8））对多模态医学图像的低频 NSCT 系数进行融合。
- (3) 高频 NSCT 系数的融合：将多模态医学图像高频 NSCT 系数的 INSML 特征矩阵分别作为 DTNP 系统 Π_A 和 Π_B 的外部输入。这两个 DTNP 系统从初始状态开始，然后连续执行直到达到最大迭代次数 t_{max} 停止。最后两个 DTNP 系统的输出作为高频融合规则的控制信号（公式（5.9））。
- (4) 逆 NSCT 变换：将融合的 NSCT 系数进行逆 NSCT 变换重建出融合图像。

在下面的小节中，我们将详细解释用于低频和高频 NSCT 系数的两种融合规则。

5.3.1 低频 NSCT 系数融合规则

一般来说，对低频系数采取的融合策略对最终的融合图像有很大影响。在所提出的融合方法中，考虑了两个重要因素，即能量保存和细节提取。众所周知，图像的大部分能量都包含在其低频系数中，由于源图像是通过不同的医学成像机制获得的，同一位置的不同源图像的能量强度可能有较大的差异。因此，常用的基于加权平均的融合规则往往会导致融合后的医学图像出现能量损失，亮度会急剧下降，导致视觉效果不佳。为了克服这一局限性，针对低频系数引入了一种称为 WLE 的活动水平测量方法。WLE 特征是基于局部邻域的低频系数计算出来的，它用来表达医学图像的大部分能量。WLE 的定义为：

$$WLE_{l0}(i, j) = \sum_a \sum_b W'(a, b) C_{l0}(i + a, j + b)^2 \quad (5.3)$$

其中 $C_{l0}(i, j)$ 表示在位置 (i, j) 的低频 NSCT 系数， W' 是一个权重矩阵，它的定义为：

$$W' = \frac{1}{16} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{vmatrix} \quad (5.4)$$

通常低频系数还可以包含少量的细节信息，这对多模态医学图像的融合很有帮助。为了提取细节信息考虑采用 INSML 特征，其中 INSML 被定义如下：

$$INSML_{l0}(i, j) = \sum_a \sum_b W''(a, b) IML_{l0}(i + a, j + b) \quad (5.5)$$

其中 W'' 是一个权重矩阵，其定义为：

$$W'' = \frac{1}{15} \begin{vmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix} \quad (5.6)$$

IML 的定义为：

$$\begin{aligned} IML_{l0}(i, j) = & |2C_{l0}(i, j) - C_{l0}(i - 1, j) - C_{l0}(i + 1, j)| \\ & + |2C_{l0}(i, j) - C_{l0}(i, j - 1) - C_{l0}(i, j + 1)| \\ & + \frac{1}{\sqrt{2}} |2C_{l0}(i, j) - C_{l0}(i - 1, j - 1) - C_{l0}(i + 1, j + 1)| \\ & + \frac{1}{\sqrt{2}} |2C_{l0}(i, j) - C_{l0}(i - 1, j + 1) - C_{l0}(i + 1, j - 1)| \end{aligned} \quad (5.7)$$

在 $IML_{l0}(i, j)$ 中考虑了对角线系数和相邻系数，在本章的实验中设置对角系数为 $\frac{1}{\sqrt{2}}$ 。

INSML 和 WLE 特征分别表达了多模态医学图像的主要能量和部分细节。因此，将 WLE 和 INSML 结合起来称为 WLE-INSML 特征，它将参与构建多模态医学图像的低频系数的融合规则。假设 $\alpha_{l0}^A(i, j) = INSML_{l0}^A(i, j) \times WLE_{l0}^A(i, j)$ 、 $\alpha_{l0}^B(i, j) = INSML_{l0}^B(i, j) \times WLE_{l0}^B(i, j)$ ，它们分别对应于与源图像 A 和 B。基于 WLE-INSML 的特征的低频系数融合规则被定义如下：

$$C_{l0}^F(i, j) = \begin{cases} C_{l0}^A(i, j), & \text{if } \alpha_{l0}^A(i, j) \geq \alpha_{l0}^B(i, j) \\ C_{l0}^B(i, j), & \text{otherwise} \end{cases} \quad (5.8)$$

其中 $C_{l0}^A(i, j)$ 和 $C_{l0}^B(i, j)$ 表示两张多模态医学图像在位置 (i, j) 的低频 NSCT 系数， $C_{l0}^F(i, j)$ 代表融合图像在位置 (i, j) 的低频系数。低频 NSCT 系数的融合由算法 5.1 所实现。

算法 5.1 低频 NSCT 系数的融合规则

Algorithm 5.1 Fusion rule for low frequency NSCT coefficients

Input:

Low-frequency NSCT coefficients of I_A , $C_0^A = (C_0^A(i, j))$;

Low-frequency NSCT coefficients of I_B , $C_0^B = (C_0^B(i, j))$;

Output:

Fused low-frequency NSCT coefficients, $C_0^F = (C_0^F(i, j))$;

1: for each $i \in (1, m)$ do

2: for each $j \in (1, n)$ do

算法 5.1 续

Alg. 5.1 Cont

```

3: Compute  $WLE_{l0}^A(i,j)$  and  $WLE_{l0}^B(i,j)$  based on Eq. (5.3)
4: Compute  $INSML_{l0}^A(i,j)$  and  $INSML_{l0}^B(i,j)$  based on Eq. (5.5)
5:  $\alpha_{l0}^A(i,j) = INSML_{l0}^A(i,j) \times WLE_{l0}^A(i,j)$ 
6:  $\alpha_{l0}^B(i,j) = INSML_{l0}^B(i,j) \times WLE_{l0}^B(i,j)$ 
7: if  $\alpha_{l0}^A(i,j) \geq \alpha_{l0}^B(i,j)$  then
8:    $C_0^F(i,j) = C_0^A(i,j);$ 
9: else
10:   $C_0^F(i,j) = C_0^B(i,j);$ 
11: end if
12: end for
13: end for
14: return  $C_0^F;$ 

```

5.3.2 高频 NSCT 系数融合规则

高频系数包含了图像的边缘和轮廓，它们代表了相应位置处信息的丰富程度。在 NSCT 分解中绝对值较大的系数对应的是亮度突然变化的像素，即图像中反差较大的边缘特征，如区域的边界、亮线和轮廓等。当高频系数的绝对值较大时，往往可以得到丰富的细节信息。为了生成具有丰富细节信息的高频融合图像，采用高频系数的 INSML 特征来表示图像的边缘特征，用 $INSML_{lr}(i,j)$ 来表示。INSML 特征作为 DTNP 系统的外部输入，DTNP 系统有一个特点就是相邻神经元协同点火。INSML 特征是基于局部邻域的相邻高频系数计算出来的，因此相邻高频系数的 INSML 特征可以触发相邻神经元的协同点火。

Π_A 和 Π_B 为两个 DTNP 系统，它们分别对应于多模态医学图像 A 和 B，将多模态医学图像的高频系数特征分别作为 Π_A 和 Π_B 的外部输入。两个 DTNP 系统从初始状态开始执行，直到达到最大迭代次数 t_{max} 停止。 P_A 和 P_B 表示与 Π_A 和 Π_B 相关的脉冲累加值矩阵。记 $P_A = (p_{ij}^A)_{m \times n}$, $P_B = (p_{ij}^B)_{m \times n}$, 其中 p_{ij}^A 和 p_{ij}^B 是神经元 σ_{ij} 在 Π_A 和 Π_B 的脉冲累积值。根据这两个脉冲矩阵，高频 NSCT 系数的融合规则定义如下：

$$C_{lr}^F(i,j) = \begin{cases} C_{lr}^A(i,j), & \text{if } p_{ij}^A \geq p_{ij}^B \\ C_{lr}^B(i,j), & \text{if } p_{ij}^A < p_{ij}^B \end{cases} \quad (5.9)$$

其中 $C_{lr}^A(i,j)$ 和 $C_{lr}^B(i,j)$ 表示多模态医学图像在 l 层 r 方向位置为 (i,j) 的高频 NSCT 系数， $C_{lr}^F(i,j)$ 表示融合图像 F 在 l 层 r 方向位置为 (i,j) 的高频 NSCT 系数，并且 $1 \leq i \leq m$ 、 $1 \leq$

$j \leq m$ 。多模态医学图像高频 NSCT 系数的融合规则在算法 5.2 中实现。在算法 5.2 中 N 表示所有层和方向的高频分解平面数。

图 5.3 展示了 DTNP 系统 Π_A 和 Π_B 的输出 P_A 和 P_B 分别在一对 CT 和 MRI 图像的高频系数平面上的活动水平的例子。在参数 [1 2 3 3] 下, NSCT 变换可以得到 22 个高频系数平面, 这些图像都是二值图, 其中白色表示 P_A 具有比 P_B 更高的活动水平, 反之亦然。可以看出, 这些活动图像对于 22 个高频系数平面来说是不同的。

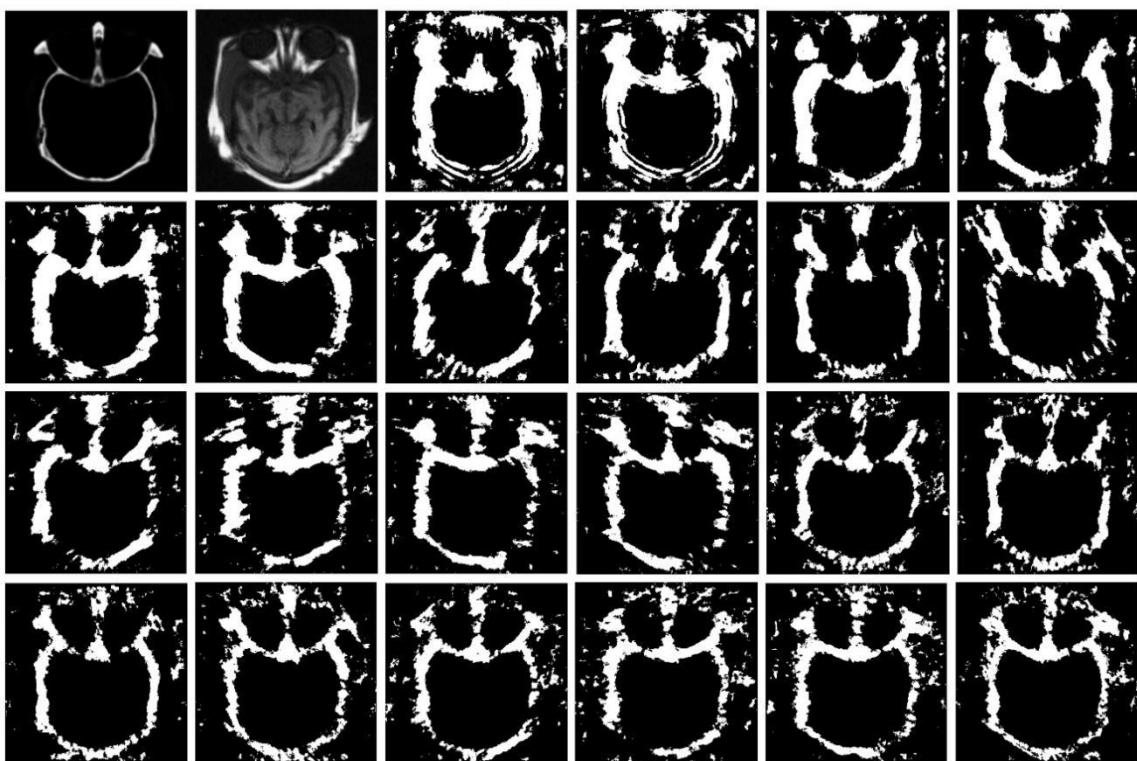


图 5.3 DTNP 系统 Π_A 和 Π_B 的输出 P_A 和 P_B 在高频系数平面上的活动水平

Fig. 5.3 Activity levels of outputs P_A and P_B of the DTNP systems Π_A and Π_B on the high-frequency coefficient planes

算法 5.2 高频 NSCT 系数的融合规则

Algorithm. 5.2 Fusion rule for high frequency NSCT coefficients

Input:

High-frequency NSCT coefficients of I_A , $C_l^A = (C_l^A(i,j))$;
High-frequency NSCT coefficients of I_B , $C_l^B = (C_l^B(i,j))$;

Output:

Fused high-frequency NSCT coefficients, $C_l^F = (C_l^F(i,j))$;
/* Compute the INSML feature matrixes of C_l^A and C_l^B : $V_A = V_{ij}^A(i,j)$ and $V_B = V_{ij}^B(i,j)$ */

算法 5.2 续

Alg. 5.2 Cont

```

1: for each  $i \in [1, m]$  do
2:   for each  $j \in [1, n]$  do
3:     Compute  $V_{ij}^A(i, j) = INSML_{l_l}^A(i, j)$  of  $C_l^A$  based on Eq. (5.5)
4:     Compute  $V_{ij}^B(i, j) = INSML_{l_l}^B(i, j)$  of  $C_l^B$  based on Eq. (5.5)
5:   end for
6: end for
/* Compute the spike sum matrixes of  $V_A$  and  $V_B$ :  $P_A = P_{ij}^A(i, j)$  and  $P_B = P_{ij}^B(i, j)$  */
7:  $P_A = DTNP(V_A);$ 
8:  $P_B = DTNP(V_B);$ 
9: for each  $i \in [1, m]$  do
10:  for each  $j \in [1, n]$  do
11:    if  $P_{ij}^A(i, j) \geq P_{ij}^B(i, j)$  then
12:       $C_l^F(i, j) = C_l^A(i, j);$ 
13:    else
14:       $C_l^F(i, j) = C_l^B(i, j);$ 
15:    end if
16:  end for
17: end for
18: return  $C_l^F;$ 

```

5.3.3 融合算法

我们提出的融合方法是基于 DTNP 系统在 NSCT 变换域中进行多模态医学图像融合，由算法 5.3 实现。算法 5.3 有两个输入图像 I_A 和 I_B 还有一个输出 I_F ，其中 $NSCT(\cdot)$ 表示 NSCT 变换， $INSCT(\cdot)$ 表示 NSCT 的逆变换； C_0^A 和 C_0^B 表示图像 A 和 B 的低频系数； C_l^A 和 C_l^B 分别表示图像 A 和 B 的高频系数； $1 \leq l \leq N$ ， N 表示高频系数分解的平面。

算法 5.3 融合算法

Algorithm. 5.3 Fusion algorithm

Input:Multi-modality medical images, I_A and I_B ;

Output:

Fused image I_F ;/* NSCT decomposition of image I_A */1: $[C_0^A, C_1^A, \dots, C_m^A] = NSCT(I_A);$ /* NSCT decomposition of image I_B */

算法 5.3 续

Alg. 5.3 Cont

```

2:  $[C_0^B, C_1^B, \dots, C_m^B] = NSCT(I_B);$ 
   /* fusion of low-frequency coefficients */
3:  $C_0^F = LowFrequencyFusion(C_0^A, C_0^B);$ 
   /* fusion of high-frequency coefficients */
4: for each  $l \in (1, N)$  do
5:    $C_l^F = HighFrequencyFusion(C_l^A, C_l^B);$ 
6: end for
   /* Inverse NSCT transform */
7:  $I_F = INSCT(C_0^F, C_1^F, \dots, C_N^F);$ 
8: return  $I_F;$ 

```

5.4 实验结果

为了评估所提出的融合方法的有效性，实验中采用了 12 对多模态医学图像，如图 5.4 所示，包括 MRI_T1 和 MRI_T2、CT 和 MRI、MRI 和 PET、MRI 和 SPECT 类型的图像。这些图像来源于开源数据库，并且它们的大小均为 256×256 。开源数据库链接 [http://www.med.harvard.edu/aanlib/。](http://www.med.harvard.edu/aanlib/)

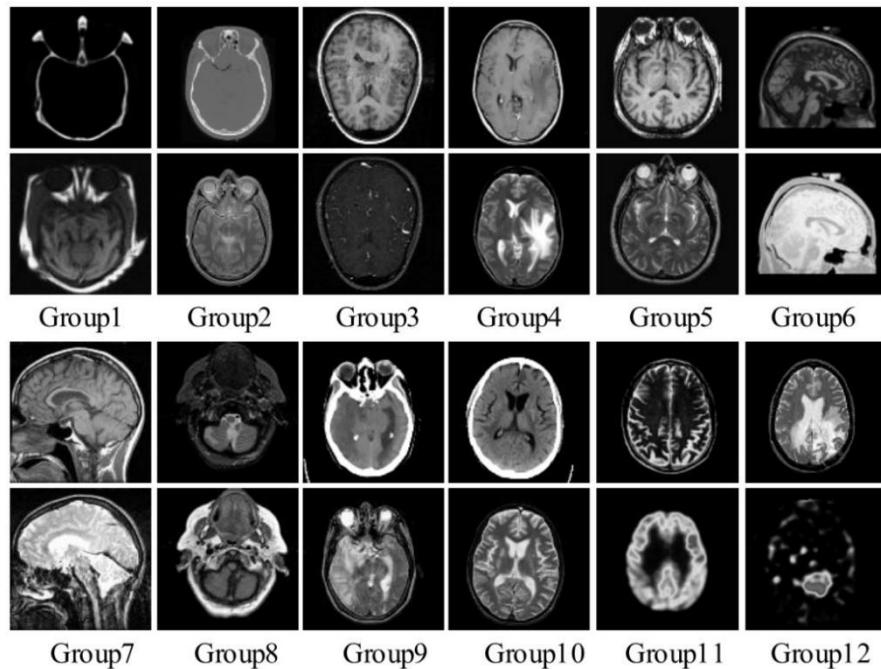


图 5.4 医学图像数据集

Fig. 5.4 Medical image dataset

在实验中，将所提出的融合方法与 9 种传统方法（Wavelet^[27]、DWT^[108]、CVT^[93]、DTCWT^[31]、ASR^[109]、NSCT^[35]、NSCT-SR^[95]、NSCT-SF-PCNN^[97]、NSST-PAPCNN^[110]）和 4 种基于深度学习的融合方法（DenseFuse^[50]、IFCNN^[51]、DDcGAN^[52]、MDLatLRR^[53]）进行了定性和定量的比较分析。

5.4.1 参数设定和客观评价指标

所提出的融合方法一种是基于 DTNP 系统在 NSCT 变换域的方法。NSCT 采用金字塔滤波器，实验中所使用的方向滤波器参数为'PKVA'、'9/7'和[1 2 3 3]。DTNP 系统的参数设置为 $t_{max} = 110$ 、 $\tau_0 = 1$ 、 $\tau = 1$ 、 $r = 3$ 和 $p = 1.1$ 。除此之外， $W_{3 \times 3} = \{w_{ij}\}_{3 \times 3}$ 被定义如下：(1) $w_{22} = 0$ ；(2) $w_{ij} = 1/sqrt((i - 2)^2 + (j - 2)^2)$ 其中 $i \neq 2$ 、 $j \neq 2$ 、 $1 \leq i \leq 3$ 和 $1 \leq j \leq 3$ 。

所提出的融合方法在 Matlab 2017b 中实现，并在英特尔酷睿 i7-6700 CPU 主频为 3.4 GHz，内存为 16 GB 的主机上运行。除基于深度学习的 4 种融合方法外，其他比较的传统方法都是在同一平台上使用相应作者提供的代码实现的。IFCNN、DenseFuse、DDcGAN 和 MDLatLRR 是在 Ubuntu 16.04 操作系统和 Inter Xeno Silver 4110 2.10GHz CPU 和 16GB RAM 的 NVIDIA Tesla P100 上实现的。

在实验中，采用了 6 个融合质量度量指标来客观评价所提出和比较的融合方法，包括 Q_{abf} ^[81]、互信息 (MI)^[82]、特征互信息 (FMI)^[83]、差异相关性之和 (SCD)^[88]、多层次结构相似性指数度量 (MSSSIM)^[89] 和标准差 (SD)^[84]。对于这 6 个度量指标，度量值越大，对应方法的融合性能越好。

5.4.2 实验结果分析

在第一组实验中，将所提出的方法与 9 种传统的融合方法进行比较，包括 Wavelet、DWT、CVT、DTCWT、ASR、NSCT、NSCT-SR、NSCT-SF-PCNN 和 NSST-PAPCNN。由于篇幅所限，仅对两组不同类型的医学图像作为例子来进行定量与定性对比分析，分别是 Group 1 (CT/MRI) 和 Group 2 (MRI_T1/MRI_T2)。这些方法融合的图像如图 5.5 和图 5.6 所示。从图 5.5 可以看出：(1) Wavelet 融合的图像不能很好的保留源图像的细节信息，整体亮度较低；(2) DWT 融合的图像存在一定的模糊，导致融合结果不理想；(3) 用 CVT、DTCWT、ASR、NSCT 和 NSCT-SF-PCNN 得到的融合图像也可以观察到上述问题，这些图像表现出较低的对比度和亮度。相比之下，所提出的融合方法、NSCT-SR 和 NSST-PAPCNN 融合的图像不仅保留了源图像的细节信息，而且表现出较高的对比度。为了更好地区分融合图像之间的差异，我们可以从图(i)、(k)和(l)所示的

局部区域细节放大图中观察到，所提出的方法融合了多模态图像中更多的细节信息，而与 NSCT-SR 和 NSST-PAPCNN 融合的图像细节相对较少。

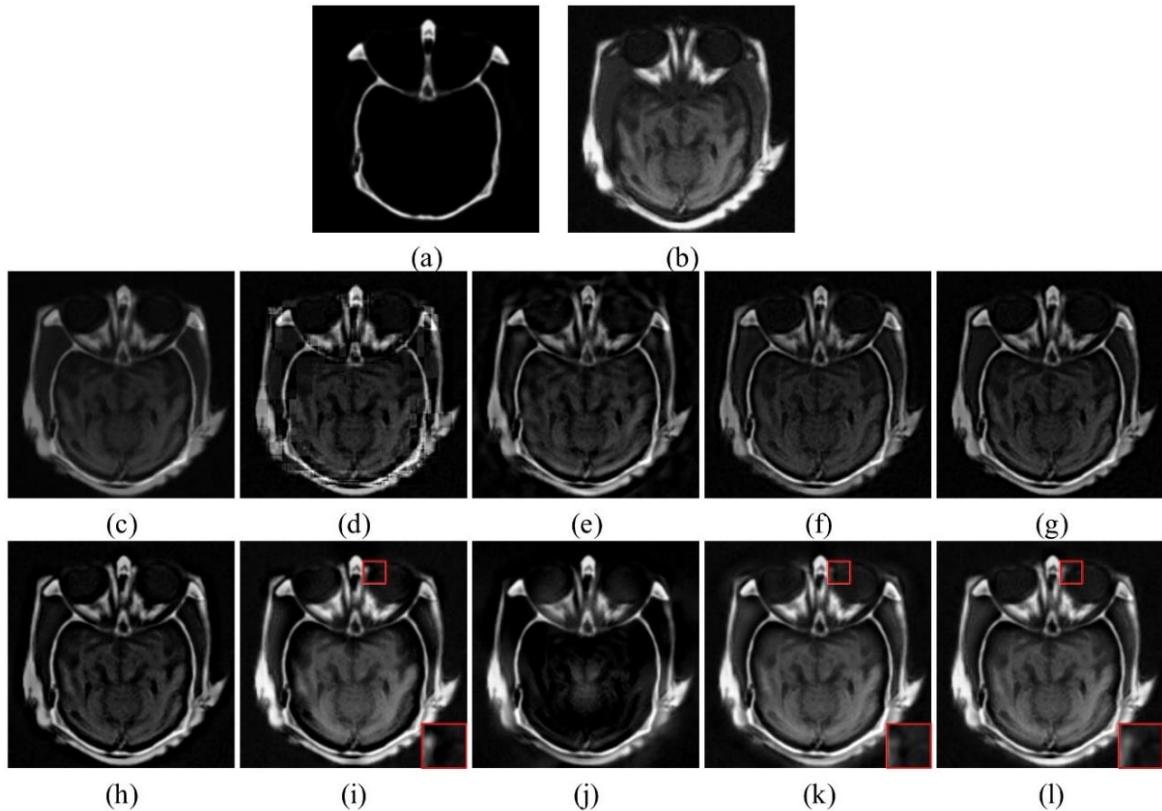


图 5.5 Group 1 图像的融合结果。(a)CT 图像, (b)MRI 图像, (c) Wavelet, (d)DWT, (e)CVT, (f)DTCWT, (g)ASR, (h)NSCT, (i)NSCT-SR, (j)NSCT-SF-PCNN, (k)NSST-PAPCNN 和 (l)提出方法。

Fig. 5.5 Fusion results obtained for the Group 1 images: (a) CT image, (b) MRI image, (c) Wavelet, (d) DWT, (e)CVT, (f) DTCWT, (g) ASR, (h) NSCT, (i) NSCT-SR, (j) NSCT-SF-PCNN, (k) NSST-PAPCNN and (l) the proposed method

从图 5.6 中可以看出：（1）对于 MRI 类型的医学图像，DWT 融合的图像非常的模糊；（2）Wavelet 融合的结果不能较好地保存源图像的细节；（3）Wavelet、DWT、CVT、DTCWT、ASR 和 NSCT 融合后的图像对比度和亮度较低；（4）NSCT-SR 和 NSCT-SF-PCNN 融合的图像含有较多的阴影。所提出的方法和 NSST-PAPCNN 可以保留源图像更多的细节信息，并且具有较高的对比度。从放大的方框中可以看出，与 NSST-PAPCNN 相比，所提出的方法融合了更多的多模态图像的细节信息。

由于单凭主观视觉效果很难判断这些方法融合的图像质量，因此有必要进行客观评估。表 5.1 展示了 10 种融合方法在两对医学图像上关于 6 个指标的定量比较结果。在表 5.1 中，粗体字表示对应的方法达到了最佳的指标值。对于 Group 1 (CT/MRI) 图像，从

表 5.1 中可以看出：（1）所提出的融合方法在 Q_{abf} 、 SCD 和 $MSSSIM$ 三个指标上达到了最佳值；（2）所提出的融合方法在 MI 指标上达到了次佳值；（3）所提出的融合方法在 FMI 和 SD 指标上达到了一个较高的值。

对于 Group 6 (MRI_T1/MRI_T2 图像) 图像，从表 5.1 中可以看出：（1）所提出的融合方法在 FMI 、 SCD 和 SD 三个指标上达到了最佳值；（2）所提出的融合方法在 Q_{abf} 、 MI 和 $MSSSIM$ 三个指标上达到了一个相对较高值。

表 5.1 所提出的方法与传统方法在两对多模态医学图像上的 6 个指标评价的结果比较

Tab.5.1 Comparison of the results of the proposed method with the conventional method for the evaluation of six metrics on two pairs of multi-modality medical images

Images	Methods	Q_{abf}	MI	FMI	SCD	$MS_S SIM$	SD
Group 1 CT/MR1	Wavelet	0.4038	11.8110	0.9019	1.3350	0.8638	33.5205
	DWT	0.5570	12.2950	0.8766	1.4341	0.8543	45.4366
	CVT	0.5593	12.5179	0.8961	1.4118	0.8445	41.2239
	DTCWT	0.6041	12.5481	0.9038	1.4386	0.8965	37.7423
	ASR	0.7068	12.3557	0.9028	1.5144	0.9355	40.0486
	NSCT	0.7087	12.2556	0.9050	1.6033	0.9465	45.0023
	NSCT-SR	0.7254	13.1784	0.9056	1.6626	0.9377	56.8796
	NSCT-SF-PCNN	0.5782	11.8549	0.9090	1.4566	0.8214	53.4049
	NSST-PAPCNN	0.6873	13.9104	0.9016	1.7223	0.9249	56.0239
	Proposed	0.7371	13.6377	0.9047	1.7269	0.9478	55.9198
Group 6 MR1_T1/ MR2_T2	Wavelet	0.4197	8.6867	0.8949	0.4425	0.8318	65.3143
	DWT	0.5269	8.8097	0.8803	0.4904	0.8364	69.6428
	CVT	0.5309	10.4741	0.8898	0.4723	0.8608	67.5608
	DTCWT	0.5384	9.5217	0.8943	0.4526	?0.8637	66.6649
	ASR	0.5586	8.4815?	0.8940	0.4663	0.8786	67.2142
	NSCT	0.5905	9.9105	0.8902	0.5404	0.8877	68.5398
	NSCT-SR	0.5885	10.4861	0.9036	0.1723	0.7951	72.5683
	NSCT-SF-PCNN	0.5393	12.1960	0.8952	0.3262	0.7613	75.2603
	NSST-PAPCNN	0.5016	11.5344	0.8819	1.1824	0.8271	98.0668
	Proposed	0.5801	10.4887	0.9039	1.3197	0.8716	100.2606

表 5.2 还列出了所提出的方法和 9 种传统方法在整个数据集上关于 6 个指标的平均融合性能。从表 5.2 可以看出：(1) 所提出的融合方法在 FMI 、 SCD 和 SD 三个指标上达到了最佳值；(2) 所提出的融合方法在 Q_{abf} 和 $MSSSIM$ 上达到了次佳值；(3) 所提出的融合方法在 MI 上达到了较高值。

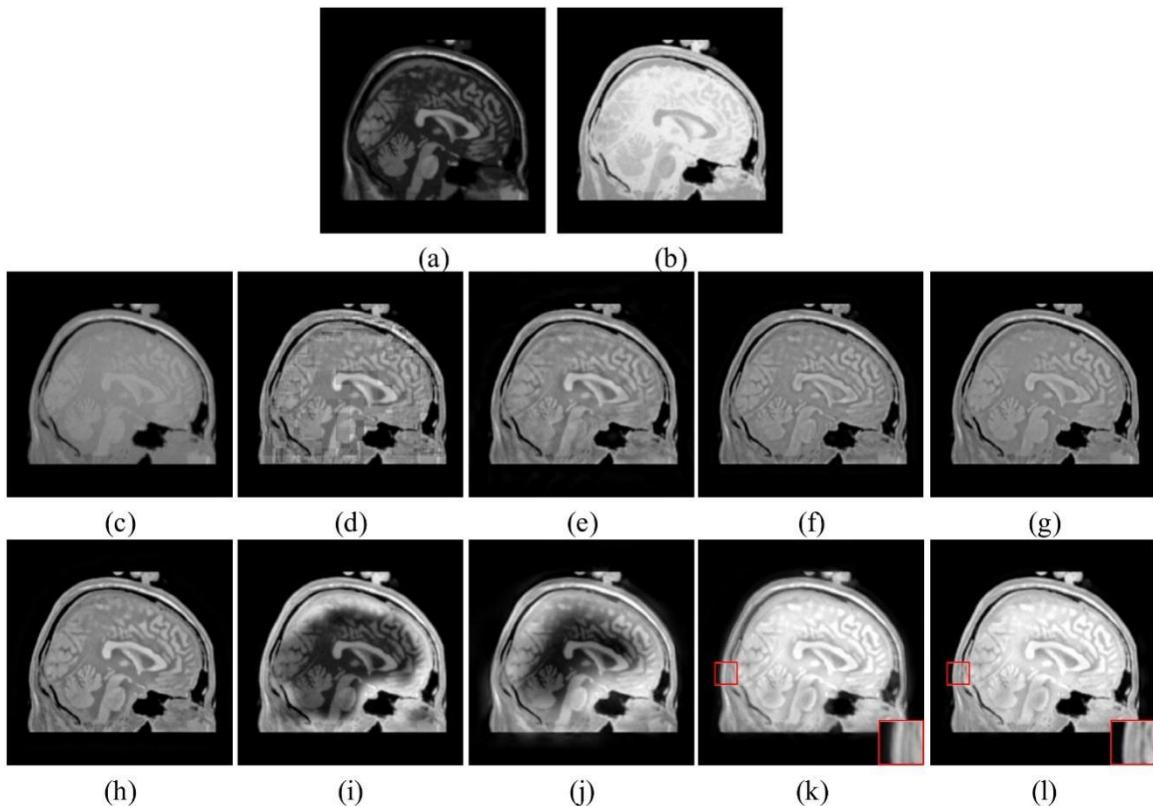


图 5.6 Group 2 图像的融合结果。(a)CT 图像, (b)MRI 图像, (c) Wavelet, (d)DWT, (e)CVT, (f)DTCWT, (g)ASR, (h)NSCT, (i)NSCT-SR, (j)NSCT-SF-PCNN, (k)NSST-PAPCNN 和 (l)提出的方法。

Fig. 5.6 Fusion results obtained for the Group 2 images: (a) CT image, (b) MRI image, (c) Wavelet, (d) DWT, (e) CVT, (f) DTCWT, (g) ASR, (h) NSCT, (i) NSCT-SR, (j) NSCT-SF-PCNN, (k) NSST-PAPCNN and (l) the proposed method

表 5.2 所提出的方法与传统方法在数据集上的 6 个指标评价的结果比较

Tab. 5.2 Comparison between the results of the proposed method and previous methods for the dataset in terms of six metrics

Images	Methods	Q_{abf}	MI	FMI	SCD	$MSSSIM$	SD
Dataset	Wavelet	0.3897	10.0751	0.8690	0.8521	0.8705	53.6580
	DWT	0.5079	10.3557	0.8625	0.8765	0.8749	61.7084
	CVT	0.5135	11.5817	0.8699	0.8802	0.8817	59.2285

表 5.2 续
Tab. 5.2 Cont

DTCWT	0.5361	10.7497	0.8748	0.8815	0.8960	57.0032
ASR	0.5784	10.1581	0.8785	0.9105	0.9086	57.4626
NSCT	0.5798	10.7822	0.8763	0.9753	0.9369	60.8543
NSCT-SR	0.6033	11.0311	0.8704	1.1197	0.9265	69.4709
NSCT-SF-PCNN	0.5335	11.1061	0.8741	1.0384	0.8642	69.3907
NSST-PAPCNN	0.5530	11.5741	0.8726	1.4069	0.9224	75.0621
Proposed	0.5805	11.3653	0.8787	1.4144	0.9324	75.1189

除此之外，我们还展示了在 12 对多模态医学图像上不同融合方法的 6 个指标值，其中横坐标的值对应 12 对图像。如图 5.7 所示，从图中的 (b) (d) (f) 中可以观察到，我们提出的融合方法的融合结果在 *FMI*、*SCD* 和 *SD* 指标均达到了最优值。

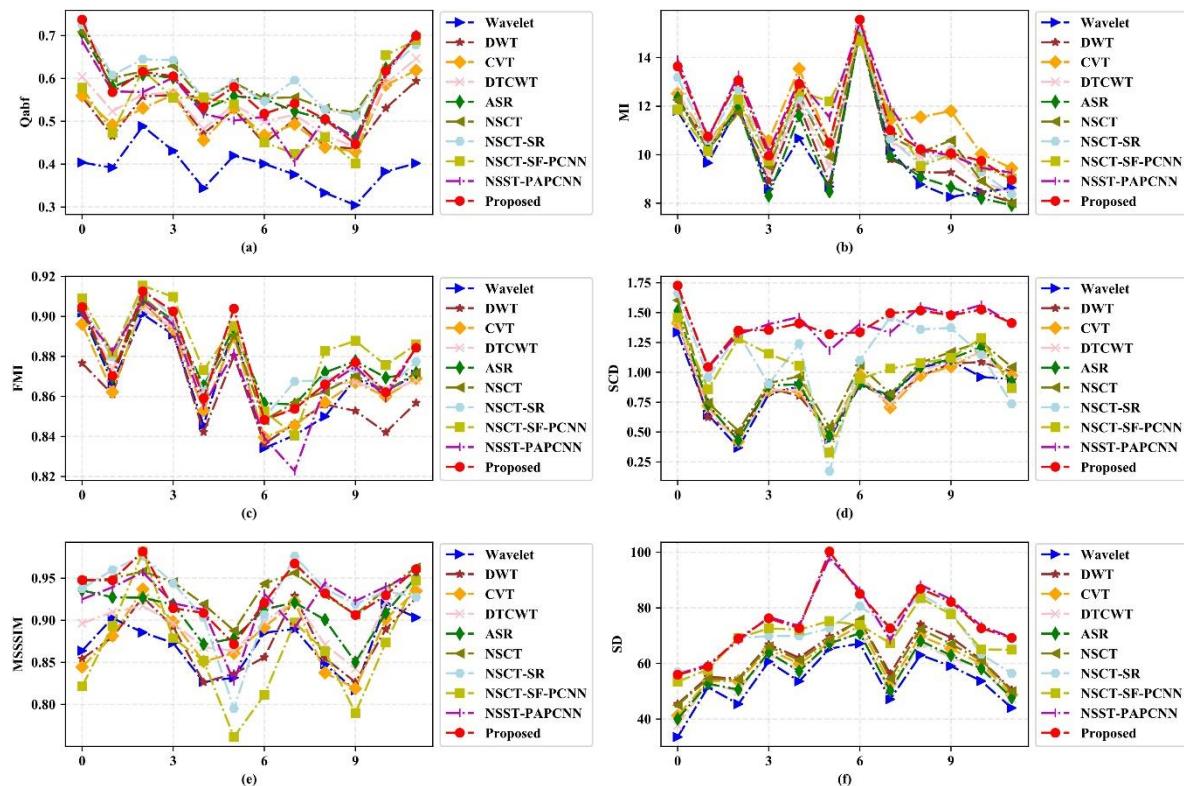


图 5.7 12 对图片在 10 种融合方法上对应的 6 个指标

Fig.5.7 12 pairs of images on ten fusion methods corresponding to six metrics

最近几年，基于深度学习（DL）的图像融合方法变得越来越流行，并且这些方法表现出良好的融合性能。将所提出的方法与基于 DL 的 4 种方法进行了比较，它们分别是 IFCNN、DenseFuse、DDcGAN 和 MDLatLRR。同样，仅使用了两对典型的多模态医学图像 Group 11（MRI/PET）和 Group 12（MRI/SPECT）来进行对比分析，其它的图像有类似的结果。图 5.8 和图 5.9 分别展示了所提出的方法和 4 种基于 DL 在这两对医学图像上的融合结果，其中（a）和（b）分别对应不同模态的两幅源医学图像。

观察图 5.8 的 MRI/PET 图像，所提出的方法、MDLatLRR、IFCNN 和 DenseFuse 在视觉上均表现出了良好的融合性能，但是 DDcGAN 融合结果显得有些模糊。在图 5.9 中，可以观察到，所提出的融合方法和 MDLatLRR（level 3 和 level 4）都保留了更多的细节、对比度更好。但是，无法直观地判断这些融合方法中哪种方法最好或最差，所以我们进行了客观的比较。

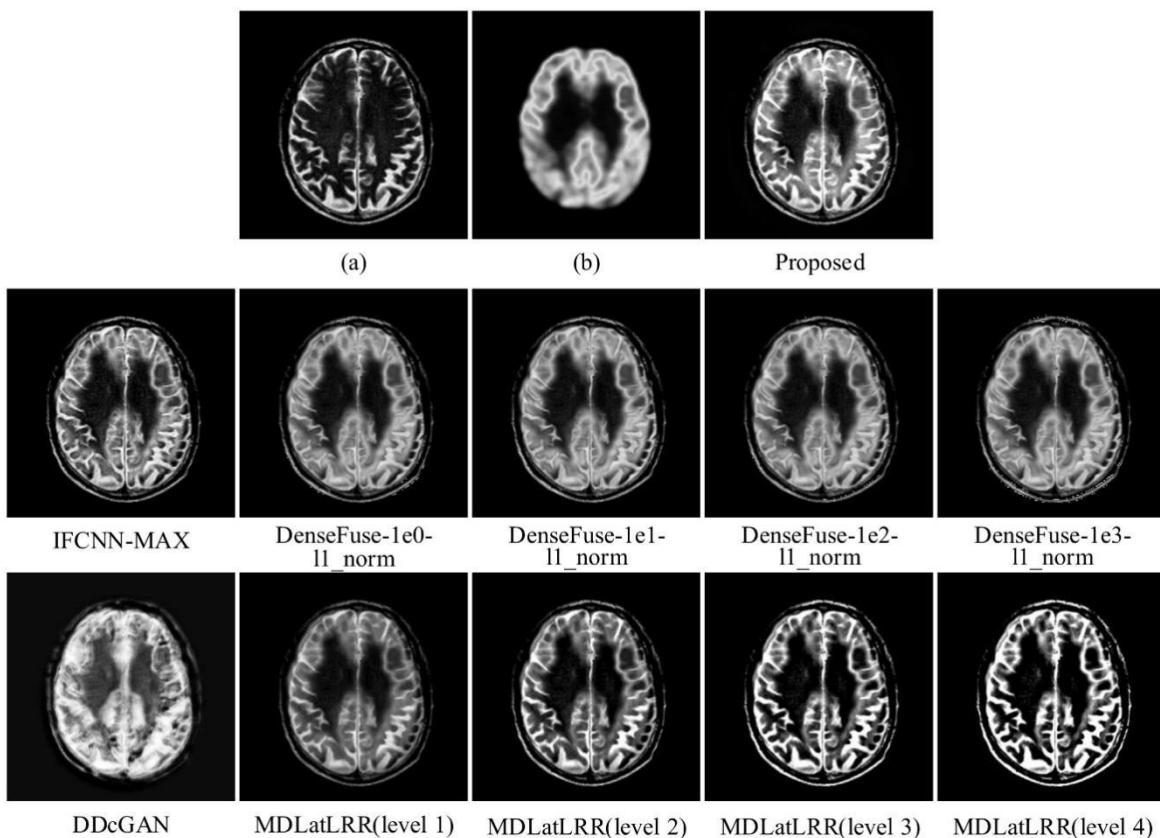


图 5.8 Group 11 图像的融合结果，(a)MRI 图像和(b) PET 图像

Fig. 5.8 Fusion results on the Group 11 images, (a) MRI image and (b) PET image

表 5.3 列出了所提出的方法和 4 种基于 DL 方法在两对医学图像上关于六个指标的定量比较。粗体对应于每个指标达到最佳的方法。对于 Group 11 (MRI/SPECT 图像)，从

表 5.3 中可以看出：（1）所提出的融合方法在 FMI 和 SCD 这两个指标上达到了最佳值；（2）所提出的融合方法在 Q_{abf} 和 MI 指标上达到了次佳值；（3）所提出的融合方法在 $MSSSIM$ 和 SD 指标上达到了一个相对较高值。

表 5.3 所提出的方法与基于 DL 在两对多模态医学图像上的 6 个指标评价的结果比较

Tab. 5.3 Comparison between the results of the proposed method and fusion methods based on DL on two multi-modality medical images in terms of six metrics

Images	Method	Q_{abf}	MI	FMI	SCD	$MSSSIM$	SD
Group 11 MRI/PET	IFCNN	0.5845	8.4333	0.8513	1.3514	0.9395	65.0031
		$\lambda = 1e^0$	0.5147	8.4868	0.8490	1.1832	0.9108
		DenseFuse	$\lambda = 1e^1$	0.5223	8.5148	0.8507	1.2212
		l_1 -norm	$\lambda = 1e^2$	0.5140	8.4969	0.8501	1.2012
	DDcGAN	$\lambda = 1e^3$	0.4973	8.5058	0.8440	1.1741	0.8998
		0.1556	10.0066	0.7863	1.2004	0.7514	77.2907
		level 1	0.5233	8.3320	0.8602	1.0607	0.9363
		MDLatLRR	level 2	0.6185	7.8842	0.8584	1.3508
	Proposed	level 3	0.5413	6.9550	0.8395	1.3437	0.9035
		level 4	0.4235	5.9628	0.8233	1.2698	0.8634
		0.6176	9.7471	0.8621	1.5261	0.9297	72.6617
		IFCNN-Max	0.6757	8.3005	0.8603	1.1393	0.9569
Group 12 MRI/SPECT	IFCNN	$\lambda = 1e^0$	0.6751	8.5757	0.8711	0.9672	0.9288
		DenseFuse	$\lambda = 1e^1$	0.6865	8.6748	0.8737	0.9777
		l_1 -norm	$\lambda = 1e^2$	0.6685	8.6550	0.8692	0.9673
		$\lambda = 1e^3$	0.6637	8.6017	0.8669	0.9688	0.9233
	DDcGAN	0.1003	11.1360	0.8308	0.7138	0.5482	76.3385
		level 1	0.6172	8.0978	0.8759	0.9497	0.9418
		MDLatLRR	level 2	0.6966	7.6814	0.8650	1.0503
		level 3	0.5767	7.1560	0.8510	1.0832	0.9276
	Proposed	level 4	0.4073	6.5555	0.8354	1.0673	0.8718
		0.6992	8.9643	0.8843	1.4128	0.9607	69.1982

对于 Group 12 (MRI/PET 图像)，从表 5.3 中可以看出：(1) 所提出的融合方法在 Q_{abf} 、 FMI 和 SCD 三个指标上达到了最佳值；(2) 我们提出的融合方法达到了次佳值 $MSSSIM = 0.9607$ 。

表 5.4 还比较了所提出的方法和 4 种基于 DL 方法在整个数据集上的平均融合性能。从表 5.4 可以看出：(1) 所提出的融合方法在 Q_{abf} 、 FMI 和 SCD 三个指标上达到了最佳值；(2) 所提出的融合方法在 MI 和 SD 上达到了次佳值；(3) 所提出的融合方法在 $MSSSIM$ 上达到了较高值。

图 5.10 还展示了 IFCNN、DenseFuse、DDcGAN、MDLatLRR 和我们提出的融合方法在每对图像上的 6 个指标值，其中横坐标的值对应 12 对图像。如图 5.10 所示，可以发现所提出的融合方法在大部分指标上都超过了这 4 种基于 DL 方法，这表明了所提出融合方法的优势。

表 5.4 所提出的方法和基于 DL 的方法在数据集上的 6 个指标比较结果

Tab. 5.4 Comparison between the results of the proposed method and DL-based methods on the dataset in terms of six metrics

Images	Method	Q_{abf}	MI	FMI	SCD	$MSSSIM$	SD
IFCNN		0.5695	10.4741	0.8651	1.1999	0.9342	67.6995
	$\lambda = 1e^0$	0.5467	10.3444	0.8744	1.0146	0.8933	64.7813
	$\lambda = 1e^1$	0.5566	10.4553	0.8756	1.0307	0.8911	66.1080
	$\lambda = 1e^2$	0.5465	10.4082	0.8746	1.0162	0.8930	64.9892
Dataset	$\lambda = 1e^3$	0.5488	10.3720	0.8737	1.0237	0.8895	66.1107
	DDcGAN	0.2719	12.0227	0.8255	1.2149	0.6916	82.8303
	MDLatLRR	level 1	0.4799	10.1511	0.8728	0.8861	0.8999
		level 2	0.5704	9.9505	0.8703	1.0820	0.9476
		level 3	0.5136	9.5233	0.8601	1.1752	0.9274
Proposed		level 4	0.4031	8.9957	0.8453	1.1984	0.8824
			0.5805	11.3653	0.8787	1.4144	0.9324
							75.1189

需要指出的是，所提出的融合方法是一种基于 NSCT 的方法，其中采用 DTNP 系统来提高其融合性能。实验比较了所提出的 NSCT+DTNP 方法和 3 种基于 NSCT 的方法 (NSCT、NSCT-SR、NSCT-SF-PCNN)。图 5.5 和图 5.6 显示，所提出的方法取得了较好的融合性能。从表 5.1 和表 5.2 中可以看出，对于 6 个评估指标，所提出的方法在大部分情况下都优于 NSCT、NSCT-SR 和 NSCT-SF-PCNN。

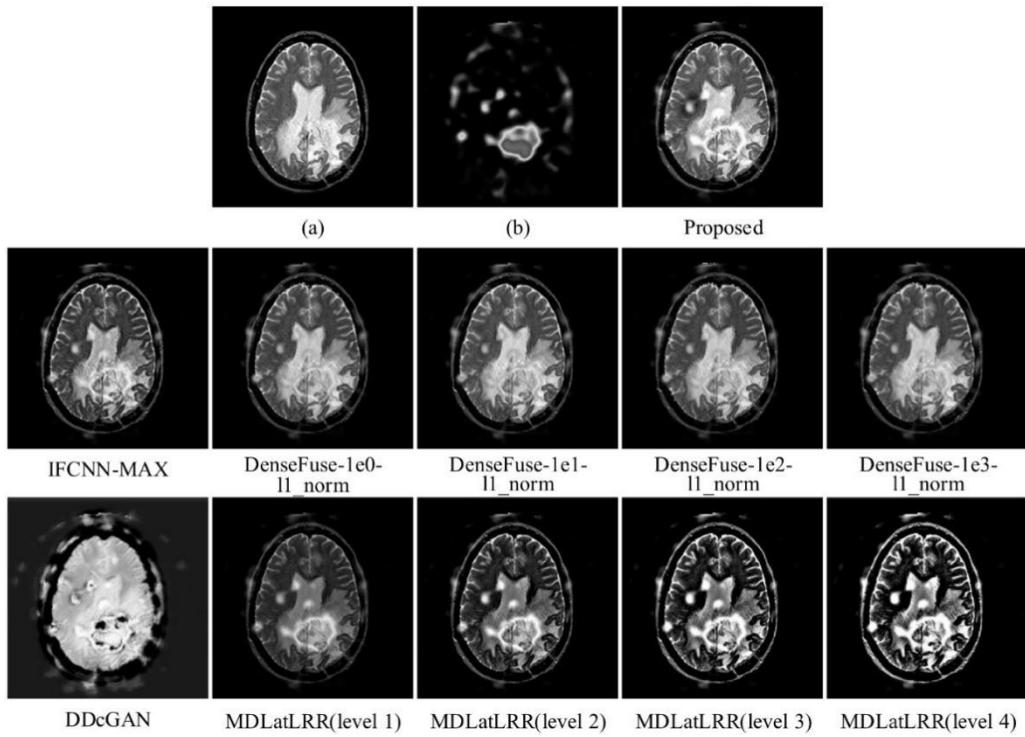


图 5.9 Group 12 图像的融合结果, (a) MRI 图像和 (b) SPECT 图像
Fig. 5.9 Fusion results on the Group 12 images: (a) MRI image and (b) SPECT image

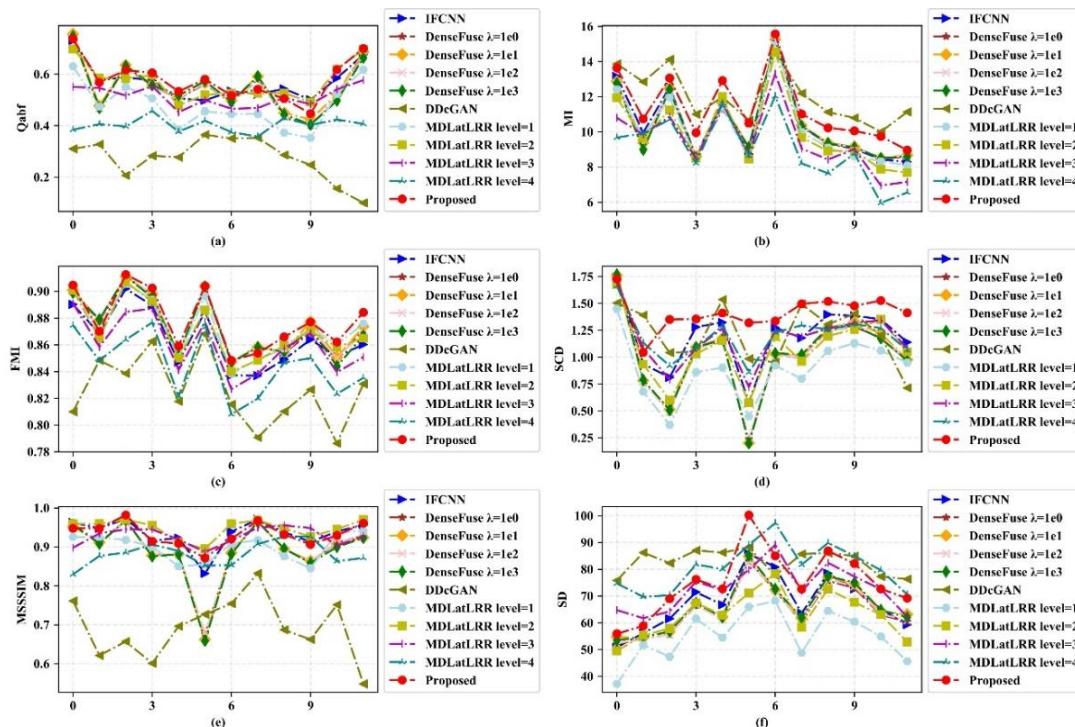


图 5.10 12 对图片在 DL 和我们提出的融合方法上对应的 6 个指标
Fig.5.10 12 pairs of images on DL and our proposed fusion method corresponding to the 6 metrics

此外，表 5.3 和表 5.4 中的结果也表明，所提出的融合方法在大部分指标上超过 4 种基于 DL 的融合方法。因此，这些主观和客观的对比结果表明，DTNP 系统在提高基于 NSCT 的融合方法的性能方面具有一定的优势。

在本章中，我们将提出的融合算法在整个多模态医学图像的融合结果展示出来，如图 5.11 所示，从主观角度看，这些融合结果表明所提出的方法对多模态医学图像具有良好的融合效果。

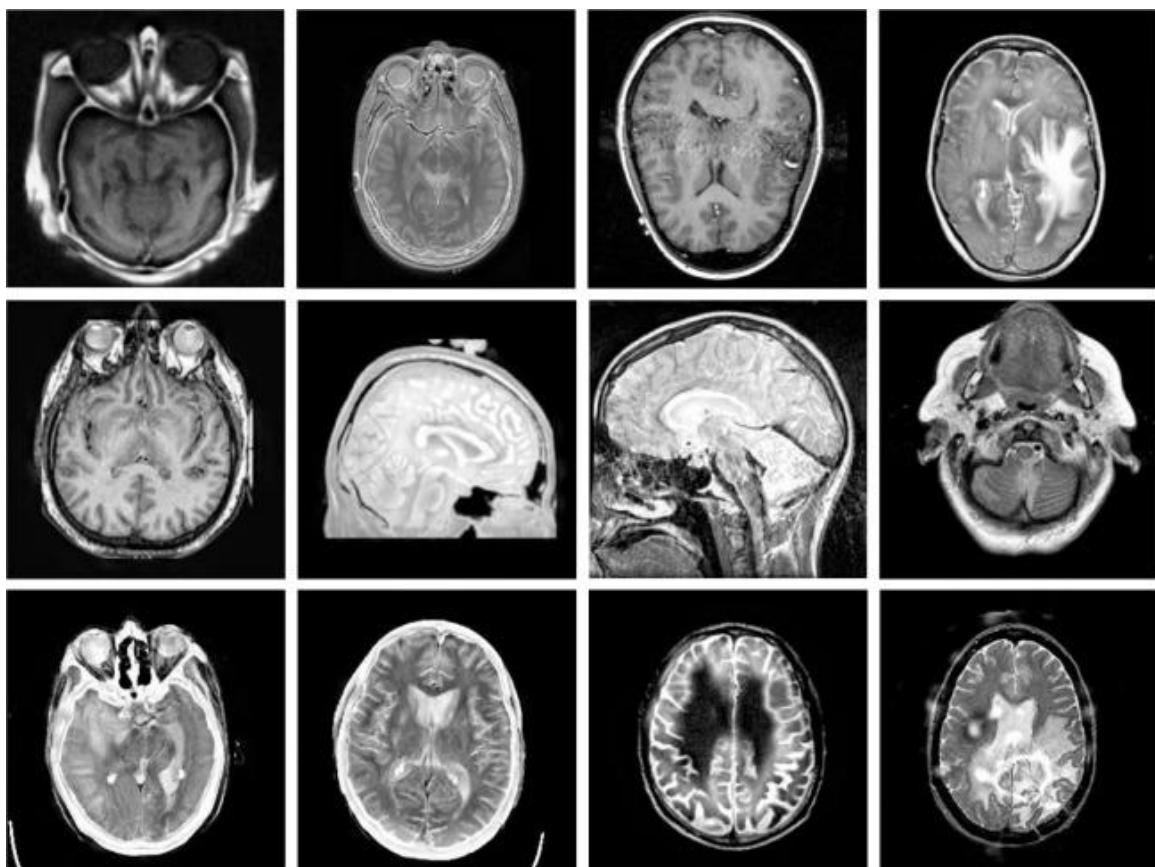


图 5.11 所提出的方法在 12 对多模态医学图像上的融合结果

Fig. 5.11 Fusion results of the proposed method on twelve pairs of multi-modality medical images

5.5 本章小结

本章研究了一种基于 DTNP 系统在 NSCT 变换域的多模态医学图像融合框架。对于低频 NSCT 系数，结合 INSML 和 WLE 特征来表达多模态医学图像的主要能量和局部细节。对于高频 NSCT 系数，利用 INSML 特征提取医学图像的边缘、轮廓等细节信息，将其作为 DTNP 系统的外部输入。由于 DTNP 系统中神经元的协同点火机制，INSML

特征可以更好地触发神经元点火。实验使用了 12 对多模态医学图像对所提出的融合方法和与 9 种传统的方法和 4 种基于深度学习的融合方法在 6 个指标上进行了对比。实验的主观和客观比较结果验证了所提出的融合方法在多模态医学图像融合上的优势。

结 论

本文首先介绍了图像融合的背景、意义和研究现状，简要的介绍了本文使用的 3 种多尺度变换方法，然后详细的介绍了 DTNP 系统的定义以及具有局部邻域结构的二维 DTNP 系统，最后在此基础上介绍了本文的主要工作：

(1) 结合多焦点图像成像的原理，本文提出了一个基于 DTNP 系统的多焦点图像融合框架。开发出一种基于 DTNP 系统和 ST 的多焦点图像融合算法。在 20 幅多焦点图像上对比实验结果验证了所提出的融合算法在多焦点图像融合上的可用性和融合性能等方面的优势通过实验表明。与目前的深度学习相比，所提出的算法具有高的计算效率。

(2) 提出了一种基于 DTNP 系统和非下采样剪切波变换 (NSST) 的红外和可见光图像融合框架。开发出一种基于 DTNP 系统和 NSST 变换的红外和可见图像融合算法。该算法的主要思想是通过脉冲累加值来选取融合对象。通过实验表明，在该框架下融合的红外和可见光图像包含源图像丰富细节信息的同时还具有良好的一致性，同目前基于深度学习的融合方法相比具有竞争优势，甚至超过部分深度学习融合方法。

(3) 提出了一种基于 DTNP 系统的多模态医学图像融合框架，并开发出相应的多模态医学图像融合算法。实验使用了 12 对多模态医学图像对所提出的融合方法和与 9 种传统的方法和 4 种基于深度学习的融合方法在 6 个指标上进行了对比。实验的主观和客观比较结果验证了所提出的融合方法在多模态医学图像融合上的优势。

近年来，彩色图像和三维医学图像受到了极大的关注。然而，所提出的融合框架是相对于灰度图像而设计的，因此，融合框架不能直接应用于彩色图像和三维医学图像融合，未来的研究将重点开发基于 DTNP 系统的彩色图像和三维医学图像融合框架。

参 考 文 献

- [1] 侯昕廷. 变分法在图像融合中的应用研究 [D]. 武汉大学, 2017.
- [2] Liu Y, Wang L, Cheng J, et al. Multi-focus image fusion: A survey of the state of the art [J]. Information Fusion, 2020, 64: 71–91.
- [3] 王玉, 王明泉. 基于小波变换的红外与可见光图像融合技术研究 [J]. 红外, 2013, 34(03): 12–14.
- [4] Li S, Yin H, Fang L. Group-sparse representation with dictionary learning for medical image denoising and fusion [J]. IEEE Transactions on Biomedical Engineering, 2012, 59(12): 3450–3459.
- [5] Reinhard E, Adhikhmin M, Gooch B, et al. Color transfer between images [J]. IEEE Computer Graphics and Applications, 2001, 21(5): 34–41.
- [6] Simone G, Farina A, Morabito F C, et al. Image fusion techniques for remote sensing applications [J]. Information Fusion, 2002, 3(1): 3–15.
- [7] Peng H, Wang J, Pérez-Jiménez M J, et al. Dynamic threshold neural P systems [J]. Knowledge-Based Systems, 2019, 163: 875–884.
- [8] Daily M, Farr T, Elachi C, et al. Geologic interpretation from composited radar and Landsat imagery [J]. Photogrammetric Engineering and Remote Sensing, 1979, 45(8): 1109–1116.
- [9] 韩潇. 视觉物联网环境下的图像融合算法研究 [D]. 江南大学, 2014.
- [10] Ma J, Yu W, Liang P, et al. FusionGAN: A generative adversarial network for infrared and visible image fusion [J]. Information Fusion, 2019, 48: 11–26.
- [11] De I, Chanda B. Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure [J]. Information Fusion, 2013, 14(2): 136–146.
- [12] Li M, Cai W, Tan Z. A region-based multi-sensor image fusion scheme using pulse-coupled neural network [J]. Pattern Recognition Letters, 2006, 27(16): 1948–1956.
- [13] Li S, Yang B. Multifocus image fusion using region segmentation and spatial frequency [J]. Image and Vision Computing, 2008, 26(7): 971–979.
- [14] 刘贤文. 多尺度几何分析和稀疏表示的多光谱图像融合方法 [D]. 南京理工大学, 2018.
- [15] Wan T, Zhu C, Qin Z. Multi-focus image fusion based on robust principal component analysis [J]. Pattern Recognition Letters, 2013, 34(9): 1001–1008.
- [16] Li S, Kang X, Hu J. Image fusion with guided filtering [J]. IEEE Transactions on Image Processing, 2013, 22(7): 2864–2875.
- [17] Jiang Y, Wang M. Image fusion with morphological component analysis [J]. Information Fusion, 2014, 18: 107–118.

- [18] Huang W, Jing Z. Multi-focus image fusion using pulse coupled neural network [J]. Pattern Recognition Letters, 2007, 28(9): 1123–1132.
- [19] Wang Z, Ma Y, Gu J. Multi-focus image fusion using PCNN [J]. Pattern Recognition, 2010, 43(6): 2003–2016.
- [20] Zhang Y, Chen L, Zhao Z, et al. Multi-focus image fusion based on robust principal component analysis and pulse-coupled neural network [J]. Optik, 2014, 125(17): 5002–5006.
- [21] 刘智嘉, 贾鹏, 夏寅辉, 等. 基于红外与可见光图像融合技术发展与性能评价 [J]. 激光与红外, 2019, 49(05): 633–640.
- [22] Peter J. The Laplacian pyramid as a compact image code[J]. Fundamental Papers in Wavelet Theory, 2006, 31(4): 28.
- [23] Toet A. Image fusion by a ratio of low-pass pyramid [J]. Pattern Recognition Letters, 1989, 9(4): 245–253.
- [24] He D-X, Meng Y, Wang C-Y. Contrast pyramid based image fusion scheme for infrared image and visible image[C]. 2011 IEEE International Geoscience and Remote Sensing Symposium, 2011: 597–600.
- [25] Petrovic V S, Xydeas C S. Gradient-based multiresolution image fusion [J]. IEEE Transactions on Image Processing, 2004, 13(2): 228–237.
- [26] Zhou Z, Li S, Wang B. Multi-scale weighted gradient-based fusion for multi-focus images [J]. Information Fusion, 2014, 20: 60–72.
- [27] Chipman L J, Orr T M, Graham L N. Wavelets and image fusion[C]. Proceedings, International Conference on Image Processing, 1995: 248–251.
- [28] Tian J, Chen L. Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure [J]. Signal Processing, 2012, 92(9): 2137–2146.
- [29] Redondo R, Šroubek F, Fischer S, et al. Multifocus image fusion using the log-Gabor transform and a multisize windows technique [J]. Information Fusion, 2009, 10(2): 163–171.
- [30] Pajares G, De La Cruz J M. A wavelet-based image fusion tutorial [J]. Pattern Recognition, 2004, 37(9): 1855–1872.
- [31] Lewis J J, O' Callaghan R J, Nikolov S G, et al. Pixel-and region-based image fusion with complex wavelets [J]. Information Fusion, 2007, 8(2): 119–130.
- [32] Li S, Yang B. Multifocus image fusion by combining curvelet and wavelet transform [J]. Pattern Recognition Letters, 2008, 29(9): 1295–1301.
- [33] Zhang B, Zhang C, Liu Y, et al. Multi-focus image fusion algorithm based on compound PCNN in Surfacelet domain [J]. Optik, 2014, 125(1): 296–300.

- [34] Jin X, Chen G, Hou J, et al. Multimodal sensor medical image fusion based on nonsubsampled shearlet transform and S-PCNNs in HSV space [J]. *Signal Processing*, 2018, 153: 379–395.
- [35] Zhang Q, Guo B-1. Multifocus image fusion using the nonsubsampled contourlet transform [J]. *Signal Processing*, 2009, 89(7): 1334–1346.
- [36] Yang L, Guo B, Ni W. Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform [J]. *Neurocomputing*, 2008, 72(1–3): 203–211.
- [37] Zhu Z, Yin H, Chai Y, et al. A novel multi-modality image fusion method based on image decomposition and sparse representation [J]. *Information Sciences*, 2018, 432: 516–529.
- [38] Zhang Q, Shi T, Wang F, et al. Robust sparse representation based multi-focus image fusion with dictionary construction and local spatial consistency [J]. *Pattern Recognition*, 2018, 83: 299–313.
- [39] Ma X, Hu S, Liu S, et al. Multi-focus image fusion based on joint sparse representation and optimum theory [J]. *Signal Processing: Image Communication*, 2019, 78: 125–134.
- [40] Zhang M, Li S, Yu F, et al. Image fusion employing adaptive spectral-spatial gradient sparse regularization in UAV remote sensing [J]. *Signal Processing*, 2020, 170: 107434.
- [41] Zhang Y, Yang M, Li N, et al. Analysis-synthesis dictionary pair learning and patch saliency measure for image fusion [J]. *Signal Processing*, 2020, 167: 107327.
- [42] Hu Q, Hu S, Zhang F. Multi-modality medical image fusion based on separable dictionary learning and Gabor filtering [J]. *Signal Processing: Image Communication*, 2020, 83: 115758.
- [43] Li H, Wang Y, Yang Z, et al. Discriminative dictionary learning-based multiple component decomposition for detail-preserving noisy image fusion [J]. *IEEE Transactions on Instrumentation and Measurement*, 2019, 69(4): 1082–1102.
- [44] Zhu Z, Chai Y, Yin H, et al. A novel dictionary learning approach for multi-modality medical image fusion [J]. *Neurocomputing*, 2016, 214: 471–482.
- [45] Li H, He X, Tao D, et al. Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning [J]. *Pattern Recognition*, 2018, 79: 130–146.
- [46] Liu Y, Chen X, Ward R K, et al. Medical image fusion via convolutional sparsity based morphological component analysis [J]. *IEEE Signal Processing Letters*, 2019, 26(3): 485–489.

- [47] Tang H, Xiao B, Li W, et al. Pixel convolutional neural network for multi-focus image fusion [J]. *Information Sciences*, 2018, 433: 125–141.
- [48] Guo X, Nie R, Cao J, et al. Fully convolutional network-based multifocus image fusion [J]. *Neural Computation*, 2018, 30(7): 1775–1800.
- [49] Amin-Naji M, Aghagolzadeh A, Ezoji M. Ensemble of CNN for multi-focus image fusion [J]. *Information Fusion*, 2019, 51: 201–214.
- [50] Li H, Wu X-J. Densefuse: A fusion approach to infrared and visible images [J]. *IEEE Transactions on Image Processing*, 2018, 28(5): 2614–2623.
- [51] Zhang Y, Liu Y, Sun P, et al. IFCNN: A general image fusion framework based on convolutional neural network [J]. *Information Fusion*, 2020, 54: 99–118.
- [52] Ma J, Xu H, Jiang J, et al. DDcGAN: A dual-discriminator conditional generative adversarial network for multi-resolution image fusion [J]. *IEEE Transactions on Image Processing*, 2020, 29: 4980–4995.
- [53] Li H, Wu X-J, Kittler J. MDLatLRR: A novel decomposition method for infrared and visible image fusion [J]. *IEEE Transactions on Image Processing*, 2020, 29: 4733–4746.
- [54] Ionescu M, Păun G, Yokomori T. Spiking neural P systems [J]. *Fundamenta Informaticae*, 2006, 71(2, 3): 279–308.
- [55] Peng H, Yang J, Wang J, et al. Spiking neural P systems with multiple channels [J]. *Neural Networks*, 2017, 95: 66–71.
- [56] Peng H, Wang J. Coupled neural P systems [J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, 30(6): 1672–1682.
- [57] Peng H, Li B, Wang J, et al. Spiking neural P systems with inhibitory rules [J]. *Knowledge-Based Systems*, 2020, 188: 105064.
- [58] Peng H, Lv Z, Li B, et al. Nonlinear spiking neural P systems [J]. *International Journal of Neural Systems*, 2020, 30(10): 2050008.
- [59] Peng H, Bao T, Luo X, et al. Dendrite P systems [J]. *Neural Networks*, 2020, 127: 110–120.
- [60] Lu Y M, Do M N. Multidimensional directional filter banks and surfacelets [J]. *IEEE Transactions on Image Processing*, 2007, 16(4): 918–931.
- [61] Bamberger R H, Smith M J. A filter bank for the directional decomposition of images: theory and design [J]. *IEEE Transactions on Signal Processing*, 1992, 40(4): 882–893.
- [62] 王建. 基于改进的稀疏表示和神经网络的图像融合研究 [D]. 江南大学, 2019.
- [63] Ganasala P, Kumar V. Multimodality medical image fusion based on new features in NSST domain [J]. *Biomedical Engineering Letters*, 2014, 4(4): 414–424.
- [64] Liu X, Mei W, Du H. Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion [J]. *Neurocomputing*, 2017, 235: 131–139.

- [65] 楼建强, 李俊峰, 戴文战. 非下采样剪切波变换的医学图像融合 [J]. 中国图象图形学报, 2017, 22(11): 1574–1583.
- [66] 欧阳宁, 邹宁, 张彤, 等. 基于非下采样 Shearlet 变换与聚焦区域检测的多聚焦图像融合算法 [J]. 计算机应用, 2015, 35(02): 490–494.
- [67] Da Cunha A L, Zhou J, Do M N. The nonsubsampled contourlet transform: theory, design, and applications [J]. IEEE Transactions on Image Processing, 2006, 15(10): 3089–3101.
- [68] 沈洁. 基于多分辨率分析的图像融合技术研究 [D]. 扬州大学, 2009.
- [69] Bhatnagar G, Wu Q J, Liu Z. Directive contrast based multimodal medical image fusion in NSCT domain [J]. IEEE Transactions on Multimedia, 2013, 15(5): 1014–1024.
- [70] 鲍程辉, 贺新光, 蒋卫国. 非下采样 Contourlet 变换和脉冲耦合神经网络相结合的遥感图像融合方法 [J]. 遥感信息, 2015, 30(02): 50–56.
- [71] 张晶, 梁艳梅, 常胜江, 等. 基于能量最小化的肾脏计算断层扫描图像分割方法 [J]. 物理学报, 2013, 62(20): 1–6.
- [72] 赵建奇. 基于词袋模型的 MR 图像脑肿瘤分割方法的研究 [D]. 南方医科大学, 2014.
- [73] 吕芳芳. 基于几何代数的多模态医学图像配准研究 [D]. 深圳大学, 2018.
- [74] Herman G T. Fundamentals of computerized tomography: image reconstruction from projections [M]. Springer Science & Business Media, 2009.
- [75] McRobbie D W, Moore E A, Graves M J, et al. MRI from Picture to Proton [M]. Cambridge university press, 2017.
- [76] 雷元义, 冯霞. 医学影像技术与设备的过去, 现在与未来 [J]. 医疗设备信息, 2004, (08): 30–32.
- [77] Shaw L J, Iskandrian A E. Prognostic value of gated myocardial perfusion SPECT [J]. Journal of nuclear cardiology, 2004, 11(2): 171–185.
- [78] 杨国城. 基于多分辨率分析的医学图像融合关键技术研究 [D]. 电子科技大学, 2016.
- [79] 余佶. 多焦面纺织纤维图像融合的研究 [D]. 东华大学, 2011.
- [80] 杨艳春, 李娇, 王阳萍. 图像融合质量评价方法研究综述 [J]. 计算机科学与探索, 2018, 12(7): 1021–1035.
- [81] Xydeas C a, Petrovic V. Objective image fusion performance measure [J]. Electronics Letters, 2000, 36(4): 308–309.
- [82] Qu G, Zhang D, Yan P. Information measure for performance of image fusion [J]. Electronics Letters, 2002, 38(7): 313–315.
- [83] Haghigat M B A, Aghagolzadeh A, Seyedarabi H. A non-reference image fusion metric based on mutual information of image features [J]. Computers & Electrical Engineering, 2011, 37(5): 744–756.
- [84] Ma J, Ma Y, Li C. Infrared and visible image fusion methods and applications: A survey [J]. Information Fusion, 2019, 45: 153–178.

- [85] Roberts J W, Van Aardt J A, Ahmed F B. Assessment of image fusion procedures using entropy, image quality, and multispectral classification [J]. Journal of Applied Remote Sensing, 2008, 2(1): 023522.
- [86] Cover T M. Elements of information theory [M]. John Wiley & Sons, 1999.
- [87] Piella G, Heijmans H. A new quality metric for image fusion[C]. International Conference on Image Processing, 2003: 173–176.
- [88] Aslantas V, Bendes E. A new image quality metric for image fusion: The sum of the correlations of differences [J]. Aeu—international Journal of Electronics and Communications, 2015, 69(12): 1890–1896.
- [89] Ma K, Zeng K, Wang Z. Perceptual quality assessment for multi-exposure image fusion [J]. IEEE Transactions on Image Processing, 2015, 24(11): 3345–3356.
- [90] Meher B, Agrawal S, Panda R, et al. A survey on region based image fusion methods [J]. Information Fusion, 2019, 48: 119–132.
- [91] Zhang Y, Bai X, Wang T. Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure [J]. Information Fusion, 2017, 35: 81–101.
- [92] Nejati M, Samavi S, Shirani S. Multi-focus image fusion using dictionary-based sparse representation [J]. Information Fusion, 2015, 25: 72–84.
- [93] Nencini F, Garzelli A, Baronti S, et al. Remote sensing image fusion using the curvelet transform [J]. Information Fusion, 2007, 8(2): 143–156.
- [94] Naidu V P S. Image fusion technique using multi-resolution singular Value decomposition [J]. Defence Science Journal, 2011, 61(5): 479–484.
- [95] Liu Y, Liu S, Wang Z. A general framework for image fusion based on multi-scale transform and sparse representation [J]. Information Fusion, 2015, 24: 147–164.
- [96] Liu Y, Liu S, Wang Z. Multi-focus image fusion with dense SIFT [J]. Information Fusion, 2015, 23: 139–155.
- [97] Qu X B, Yan J W, Xiao H Z, et al. Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain [J]. Acta Automatica Sinica, 2008, 34(12): 1508–1514.
- [98] Liu Y, Chen X, Peng H, et al. Multi-focus image fusion with a deep convolutional neural network [J]. Information Fusion, 2017, 36: 191–207.
- [99] Zhao H, Shang Z, Tang Y Y, et al. Multi-focus image fusion based on the neighbor distance [J]. Pattern Recognition, 2013, 46(3): 1002–1011.
- [100] Zhang X, Ye P, Peng S, et al. SiamFT: An RGB-infrared fusion tracking method via fully convolutional siamese networks [J]. IEEE Access, 2019, 7: 122122–122133.
- [101] Zhang X, Ye P, Peng S, et al. DSiamMFT: An RGB-T fusion tracking method via dynamic Siamese networks using multi-layer feature fusion [J]. Signal Processing: Image Communication, 2020, 84: 115756.

- [102] Zhou Z, Wang B, Li S, et al. Perceptual fusion of infrared and visible images through a hybrid multi-scale decomposition with Gaussian and bilateral filters [J]. *Information Fusion*, 2016, 30: 15–26.
- [103] Li H, Wu X-J. Multi-focus image fusion using dictionary learning and low-rank representation[C]. International Conference on Image and Graphics, 2017: 675–686.
- [104] Li B, Peng H, Wang J, et al. Multi-focus image fusion based on dynamic threshold neural P systems and surfacelet transform [J]. *Knowledge-Based Systems*, 2020, 196: 105794.
- [105] Ma J, Zhou Z, Wang B, et al. Infrared and visible image fusion based on visual saliency map and weighted least square optimization [J]. *Infrared Physics & Technology*, 2017, 82: 8–17.
- [106] Li S, Kang X, Fang L, et al. Pixel-level image fusion: A survey of the state of the art [J]. *Information Fusion*, 2017, 33: 100–112.
- [107] James A P, Dasarathy B V. Medical image fusion: A survey of the state of the art [J]. *Information Fusion*, 2014, 19: 4–19.
- [108] Li X, Tian X, Sun Y, et al. Medical image fusion by multi-resolution analysis of wavelets transform[J]. *Wavelet Analysis and Applications*, 2007: 389–396.
- [109] Liu Y, Wang Z. Simultaneous image fusion and denoising with adaptive sparse representation [J]. *IET Image Processing*, 2014, 9(5): 347–357.
- [110] Yin M, Liu X, Liu Y, et al. Medical image fusion with parameter-adaptive pulse coupled neural network in nonsubsampled shearlet transform domain [J]. *IEEE Transactions on Instrumentation and Measurement*, 2018, 68(1): 49–64.

攻读硕士学位期间发表论文及科研成果

一、发表的论文：

- 1 **Li Bo**, Peng Hong, Wang Jun, et al. Multi-focus image fusion based on dynamic threshold neural P systems and surfacelet transform[J]. Knowledge-Based Systems, 2020, 196, 105794. (SCI 一区, 本硕士学位论文第三章)
- 2 **Li Bo**, Peng Hong, Wang Jun. A novel fusion method based on dynamic threshold neural P systems and nonsubsampled contourlet transform for multi-modality medical images[J]. Signal Processing, 2021, 178, 107793. (SCI 二区, 本硕士学位论文第五章)
- 3 **Li Bo**, Peng Hong, Luo Xiaohui, et al. Medical image fusion method based on coupled neural P systems in nonsubsampled shearlet transform domain[J]. International Journal of Neural Systems, 2021, 31(01): 2050050: 1-16. (SCI 一区)
- 4 Peng Hong, **Li Bo**, Wang Jun, et al. Spiking neural P systems with inhibitory rules[J]. Knowledge-Based Systems, 2020, 188, 105064. (SCI 一区)
- 5 Yang Qian, **Li Bo**, Huang Yue, et al. Spiking neural P systems with structural plasticity and anti-spikes[J]. Theoretical Computer Science, 2020, 801: 143-156. (SCI 四区)
- 6 Huang Yue, **Li Bo**, Lv Zeqiong, et al. On string languages generated by spiking neural p systems with multiple channels[J]. International Journal of Unconventional Computing, 2019, 14(3-4), 243-266. (SCI 四区)
- 7 Peng Hong, Lv Zeqiong, **Li Bo**, et al. Nonlinear spiking neural P systems[J]. International Journal of Neural Systems, 2020, 30(10): 2050008:1-17 (SCI 一区)

二.负责和参与的项目：

- 1 西华大学研究生创新基金项目：基于脉冲神经 P 系统的图像融合（项目编号 ycjj2019020），2019-2020，负责人。
- 2 西华大学西华杯创新创业项目：基于耦合神经 P 系统的图像融合(项目编号 2020129)，2020，负责人。
- 3 国家自然科学基金项目：大脑启发的膜计算模型及学习机理构建(项目编号 201605)，2021，参研。

致 谢

春去秋来，时光荏苒，转眼间，三年研究生求学生涯即将结束，回想过去的酸甜苦辣，我心里全是深深的感激，从本科阶段算起，我已经在西华大学呆了整整七年，感谢母校七年来来的关怀，让我经历知识的洗礼。感谢我身边的每一位老师、同学和朋友的支持与陪伴，让我在面临各种挑战和困难时不畏艰险、奋力前行。

首先，要感谢我的研究生导师彭宏教授，恩师不管从学业上还是生活中都给予我莫大的帮助。在学业上，恩师勤奋严谨的治学风格和锲而不舍的学术造诣，潜移默化的影响着我对学术的态度。尽管恩师已年过半百，但是他仍然坚持在科研一线。恩师总是给予我们很多的机会和平台，使我记忆犹新的是，恩师带我去参加膜计算会议，这是我研究生阶段第一次参加的会议，在会议上我结识了很多的朋友也提高了我的眼界。恩师总是在百忙之中对我的课题和实验进行指导，在实验阶段经常遇到很多困难，在晚上 10 点恩师还与我讨论实验方案。在生活上，我也由衷的感谢恩师和师母给予无微不至的关怀。

其次，我由衷感谢我的父母，感谢他们给予我学业上的支持，有了他们的支持自己才能安心学习。我特别要感谢我的母亲，没有她的支持和关心，我不可能有今天的成绩。我还要感谢我们实验室的罗晓辉老师，罗老师在图像领域有非常高的造诣，非常感谢罗老师在图像领域给予我的指导。感谢实验室的杨倩师姐、钱江师兄对我科研实验以及论文写作的帮助。我还要感谢包婷婷同学，正是有了她的鼓励和陪伴，使我的研究生生活充满了快乐。

最后，感谢给予该论文宝贵建议的各位老师和专家，祝您们身体健康、工作顺利！