

基于正交字典学习的多像面相位恢复算法

练秋生 高丽萍 石保顺 陈书贞

(燕山大学信息科学与工程学院 河北 秦皇岛 066004)

摘 要 相位恢复(Phase Retrieval, PR)是指利用信号的非涅尔变换或其它线性变换的强度观测值恢复原始信号。由于相位信息的丢失,相位恢复问题是一个不适定问题。为解决该问题,可利用包含更多图像信息的多像面强度观测值和图像在字典下的稀疏性进行相位恢复。正交字典学习以速度快、效果好的优点得到了成像领域的关注。该文提出利用图像在正交字典下的稀疏表示进行多像面相位恢复的算法。首先,利用多像面强度观测值构造数据保真项,并结合图像在正交字典表示下的稀疏正则项构造多像面相位恢复优化问题。然后,利用分裂 Bregman 方法(Split Bregman Method)对该非线性优化问题进行求解;此外,针对含有数据保真项的非凸子问题,利用极大极小化(Majorization-Minimization, MM)算法将这一子问题转化为易处理的优化问题,并对其进行有效求解;该文提出的算法能够仅利用强度观测值重建图像同时学习字典,得到与之匹配的正交字典。正交字典学习通过阈值处理和奇异值分解两步训练字典。由于自适应正交字典能够捕获图像的结构信息,该文算法在像面个数较少时仍能得到高质量的重建图像。仿真实验表明,该算法无论从客观标准还是从主观视觉都优于现有算法,并对噪声鲁棒。

关键词 相位恢复;多像面;稀疏表示;正交字典;字典学习

中图法分类号 TP391 DOI号 10.11897/SP.J.1016.2018.02509

Multi-Plane Phase Retrieval Algorithm Based on Orthogonal Dictionary Learning

LIAN Qiu-Sheng GAO Li-Ping SHI Bao-Shun CHEN Shu-Zhen

(School of Information Science and Engineering, Yanshan University, Qinhuangdao, Hebei 066004)

Abstract Phase retrieval (PR) refers to recover the original signal from the intensity measurements of its Fresnel transform, or recover the image from the intensities of some other linear transform of the signal. Due to the loss of the phase information, the phase retrieval problem is an ill-posed problem. To address this issue, the intensity measurements at multiple observations can be utilized for phase retrieval. Moreover, the sparsity of the underlying image under an adaptive dictionary is often utilized for phase retrieval. Multiple intensity measurements imply that the measured data contains much known information about the underlying image. Exploiting these multiple intensity measurements enables an improvement of the reconstruction quality. When orthogonal dictionary learning is applied to the image reconstruction field, training an orthogonal dictionary has the advantages of fast speed and good reconstruction quality. Based on this fact, orthogonal dictionary learning is focused by the researchers in the imaging field. In this paper, we propose a multi-plane phase retrieval algorithm. The algorithm exploits the sparse representation model of the underlying image over an adaptive orthogonal dictionary. Firstly, we exploit the intensity measurements of multiple planes to construct a data fidelity term. We combine the data fidelity term and the sparse induced regularization term of the underlying image over an orthogonal dictionary to formulate the phase retrieval optimization problem. Secondly, the split Bregman

收稿日期:2017-04-25;在线出版日期:2017-12-19。本课题得到国家自然科学基金(61471313)和河北省自然科学基金(F2014203076)资助。练秋生,男,1969年生,博士,教授,博士生导师,主要研究领域为非线性压缩感知、深度学习、相位恢复及图像处理。E-mail: lianqs@ysu.edu.cn。高丽萍,女,1991年生,硕士,主要研究方向为相位恢复。石保顺,男,1989年生,博士,讲师,主要研究方向为字典学习及相位恢复。陈书贞,女,1968年生,副教授,主要研究方向为非线性压缩感知及相位恢复。

method is utilized to solve the corresponding nonlinear optimization problem. For the nonconvex optimization sub-problem involving the data fidelity term, we exploit the Majorization-Minimization (MM) method to recast this sub-problem into a tractable optimization problem. Moreover, we propose an efficient method to tackle this problem. MM algorithm constructs a surrogate function for the original complicated optimized problem, the sequence of points generated by solving the surrogate function by basic matrix multiplications is proved to converge to a stationary point of the original problem. The proposed algorithm in this paper can recover the image and learn an orthogonal dictionary simultaneously only from the intensity measurements. The learnt orthogonal dictionary can capture the structure of the underlying image efficiently, and its atoms can match with the structure of the recovered image. The orthogonal dictionary learning problem is often attacked in two steps: the sparse coding step and the dictionary updating step. Since the mutual incoherence constraints for the atoms of the orthogonal dictionary are imposed on the learnt dictionary, the hard thresholding operator can be used for the sparse coding step. For the dictionary updating step, the SVD decomposition is used to update an orthogonal dictionary. Through the two steps, we can train an orthogonal dictionary effectively from the training samples. As the result of the fact that the adaptive orthogonal dictionary can capture the structure of the underlying image, the proposed algorithm can recover the high-quality image only from measurements, even the measured planes are few. Simulated experimental results indicate that the proposed phase retrieval method outperforms the previous multi-plane PR algorithms in terms of both the objective criteria and the subjective visual evaluation. Meanwhile, our algorithm is robust to noise.

Keywords phase retrieval; multiple planes; sparse representation; orthogonal dictionary; dictionary learning

1 引言

相位恢复是指利用信号的非涅尔变换或其它线性变换的强度观测值对原始信号进行重建,目前被广泛应用于显微镜、天文学以及材料分析等领域^[1]. 现有的光学设备只能检测到光场强度,导致相位丢失,然而相位携带了物体结构的重要信息,大约占图像全部信息的 75%. 因此,对相位恢复问题的研究具有重要的意义.

由于相位的丢失,满足强度观测值相同的解不唯一,因此相位恢复问题是一个不适定问题^[2]. 针对该不适定问题, Ivanov 等人^[3]于 1992 年首次提出利用多个强度观测值进行图像重建的多像面相位恢复算法. 由于多个像面的强度观测值携带原始图像更多的信息,与利用单个强度观测值相比,该方法在算法收敛性和抗噪性上有所改善,并扩大了重建范围. 基于这些优点,文献[4]将多像面相位恢复的思想应用到同轴全息中,利用多幅观测干涉图重建原始图像,该方法消除了干涉光的限制.

近年来,为进一步减小由于相位恢复问题的不适定性所造成的误差,人们提出了多种利用图像先验知识的多像面相位恢复算法. Gerchberg-Saxton (GS)算法^[5]是一种基于交替投影的相位恢复算法,最初用来解决针对单个像面的相位恢复问题. 利用 GS 算法可以解多像面相位恢复问题,例如并行算法^[6]和 Almoro 等人于 2006 年提出的单波束多强度相位重建(Single-Beam Multiple-Intensity Phase Reconstruction, SBMIR)算法^[7-8]. 并行算法利用多个强度观测值,通过在物面和多个像面之间交替投影并在每个像面上对衍射成像施加幅值约束以重建图像;SBMIR 算法在多个像面之间循环投影,并对每个像面上的衍射成像施加幅值约束以得到重建图像. 这两种算法对像面个数要求高,在像面个数较少时收敛性差,甚至重建失败. 为了克服这些缺点,文献[9]在之前 GS 的基础上进行改进,提出一种基于 GS 相位恢复算法的多平面全息重构算法,该算法通过对每个像面的观测值引进权重因子以实现弱补偿,从而提升算法性能. 文献[10]利用多个像面的观测值针对同轴全息进行图像重建,以克服空光束校

正误差. Tikhonov 正则化约束是解决不适定问题时常用的一种先验知识, Migukin 和 Katkovnik 等人于 2011 年提出的针对多像面相位恢复问题的增广拉格朗日 (Augmented Lagrangian, AL) 算法^[6]将该约束引入到多像面相位恢复问题中, 利用增广拉格朗日方法对所对应的优化问题进行求解, 得到重建图像. 文献[11]提出的 AL-Ph 算法利用 Tikhonov 正则化约束对纯相位图像进行重建. 由于 Tikhonov 正则项的二次性, 导致该优化问题的最优解存在系统误差和过度平滑等缺陷, 不能恢复原始图像的某些重要特征, 例如高阶不连续性. 因此此类算法对像面个数比较敏感, 当像面个数少时重建图像质量差; 此外, 此类算法对噪声敏感, 不利于实际应用.

在图像反问题求解中, 利用图像稀疏性先验大都能够提升重建质量. 基于这一认识, 图像的稀疏性被应用于多像面相位恢复问题中. Katkovnik 等人提出的解耦合增广拉格朗日 (Decoupled Augmented Lagrangian, D-AL) 算法^[12-13], 借助图像在 BM3D (Block-Matching and 3D filtering)^[14]框架下的稀疏先验, 寻找多像面相位恢复问题的最优解. 该算法利用 l_0 范数衡量变换系数的稀疏性, 并对逆成像和 BM3D 滤波这两个过程单独进行建模, 降低了求解的复杂度. 实验证明了该算法的有效性. 文献[15]提出的 4f-SPAR (4f Sparse Phase Amplitude Reconstruction) 算法在 D-AL 算法的基础上, 利用 4f 光学成像系统傅里叶域空间光调制器模拟多像面成像, 在傅里叶域对原始图像进行重建. 该算法能够在像面个数较多、观测值含有少量噪声时得到高质量的重建图像.

上述多像面相位恢复算法在像面个数较少以及观测值含噪声较多时重建图像质量差. 然而, 在实际观测过程中, 像面个数太多会增加操作的复杂度, 如何在像面个数较少时提高重建图像的质量是一个关键的问题.

图像可以表示为少量字典原子的线性组合, 即图像在字典表示下是稀疏的. 由于固定字典的原子基本形态固定, 只能与某些特定的图像结构匹配, 并不能与各类图像的复杂结构最佳匹配^[16], 人们开始根据图像本身来学习字典, 以提高字典性能. 近年来, 图像在自适应字典下的稀疏表示吸引了众多学者的关注. 文献[17]利用 K-SVD^[18]训练字典, 得到与图像本身的结构匹配的过完备字典, 并利用图像在该自适应字典下的稀疏表示进行人脸识别, 实验表明该方法在识别率上有所改善. 训练过完备字典

的缺点是时间复杂度高, 而正交字典学习的提出解决了该问题. 在稀疏编码步骤中, 正交字典学习方法采用简单的阈值法处理稀疏系数, 在保证字典性能的前提下大大降低了时间复杂度, 因此被广泛应用于图像重建^[19-20]等领域. 本文利用图像在自适应正交字典下的稀疏性, 提出了一种在像面个数较少时仍然能够有效重建原始图像的多像面相位恢复算法, 试图在像面个数较少时仍然能够得到高质量的重建图像.

2 多像面相位恢复观测模型

原始图像 u_0 经过前向传播到达第 r 个像面, 得到复值成像 u_r , 该过程可表示为

$$u_r = F_r \cdot u_0 = FFT^{-1} \{ FFT \{ u_0 \} \cdot M_r \}, \quad r=1, \dots, K \quad (1)$$

其中, $u_0 \in \mathbb{C}^M$, $u_r \in \mathbb{C}^M$, K 表示像面个数, $F_r \in \mathbb{C}^{M \times M}$ 表示从物面到第 r 个像面的传递函数. 一般情况下, 近轴光线在空间传播, 其传递函数由 Rayleigh-Sommerfield 积分定义. 考虑到不同的离散化积分模型, 式(1)中的传递函数 F_r 通常为角谱分析、空域离散衍射变换^[21]或傅里叶域离散衍射变换^[22]等. 本文对近场多像面相位恢复问题进行研究, 采用菲涅尔变换作为传递函数, 利用 4f 光学成像系统空间光调制器实现从原始图像 u_0 到复值成像 u_r 的点投影. 空间光调制器相当于傅里叶域的相位调制器, 由文献[15]可知, 其复透射率 M_r 可以表达为以下形式:

$$M_r[l_1, l_2] = \exp(2\pi j \cdot z_r \sqrt{1 - \Delta_v^2(l_1^2 + l_2^2)/f^2} / \lambda) \quad (2)$$

其中, $l = (l_1, l_2)$ 为二维频域离散变量, $z_r = z_1 + (r-1) \cdot \Delta_z$ 表示物面和第 r 个像面之间的距离, Δ_z 是每两个平行像面之间的距离; Δ_v 表示空间光调制器上的像素大小, λ 为光束波长, f 为 4f 光学成像系统中傅里叶透镜的焦距.

根据上述传播模型, 理想多像面成像强度观测模型可表示为:

$$o_r = |u_r|^2, \quad r=1, \dots, K \quad (3)$$

其中, o_r 表示第 r 个观测面上的强度观测值.

多像面相位恢复优化问题可表示为

$$\hat{u}_0 = \arg \min_{u_0} \left\{ \sum_{r=1}^K \beta_r \left\| |F_r u_0| - \sqrt{o_r} \right\|_2^2 \right\} \quad (4)$$

其中, 参数 β_r 用来衡量由于第 r 个像面处强度观测值含噪所造成的误差; 由于上述优化问题只考虑菲

涅尔变换域的幅值约束,并未引入其它先验知识,求解上述优化问题并不能得到高质量的重建图像.因此,将图像在自适应正交字典下的稀疏先验与式(4)中的优化问题相结合,以提高重建图像的质量.

3 基于正交字典学习的多像面相位恢复算法

本文提出的多像面相位恢复算法利用了一种有效的正交字典学习方法训练字典,并构造图像在自适应正交字典下的稀疏正则项.基于正交字典学习的多像面相位恢复算法对应的优化问题可表示为

$$\min_{\mathbf{u}_0, \mathbf{F}, \mathbf{D}} \sum_{r=1}^K \beta_r \|\sqrt{\sigma_r} - |\mathbf{F}_r \mathbf{u}_0|\|_2^2 + \sum_{i=1}^N \|\mathbf{R}_i \mathbf{u}_0 - \mathbf{D} \boldsymbol{\alpha}_i\|_2^2 + \tau \|\boldsymbol{\alpha}_i\|_0 \quad (5)$$

s. t. $\mathbf{D}^H \mathbf{D} = \mathbf{I}$

其中, $\mathbf{R}_i \in \mathbb{R}^{n \times M}$ ($i=1, \dots, N$) 是取块算子, 本文采用最大重叠率对待重建图像进行取块, N 表示图像块的个数; $\mathbf{R}_i \mathbf{u}_0 \in \mathbb{C}^n$ 对应从待重建图像 $\mathbf{u}_0$ 中提取的第 i 个图像块; $\mathbf{D} \in \mathbb{C}^{n \times n}$ 表示正交字典; $\boldsymbol{\alpha}_i \in \mathbb{C}^n$ 是图像块 $\mathbf{R}_i \mathbf{u}_0$ 在正交字典 $\mathbf{D}$ 下的表示系数; $\mathbf{F} = \{\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_N\}$ 是表示系数的集合; $\mathbf{I} \in \mathbb{R}^{n \times n}$ 表示单位阵. 式(5)中第一项为数据保真项, 用来描述观测幅值和衍射成像模值之间存在的误差; 第二项为待重建图像与它在正交字典下的稀疏表示的误差项, 用来刻画稀疏逼近的程度; 第三项为稀疏正则项, 参数 τ 的大小决定了对图像进行稀疏表示的稀疏程度. 由于式(5)利用复图像本身在正交字典下的稀疏性对多像面相位恢复问题进行建模, 在重建过程中, 会导致复图像的幅值和相位之间相互影响, 从而降低重建质量和视觉效果.

利用复图像的某种先验知识进行图像重建, 往往不能够得到理想的重建效果^[23]. 针对这一问题, Fessler 等人提出的用于磁共振成像的图像重建算法分别利用幅值和相位的先验知识构造正则化约束, 减小了图像重建的误差. 另外, Zhao 和 Noll 等人^[24]针对压缩感知领域中图像重构时出现的相位缠绕问题, 分别对图像的幅值和相位构造正则项, 并将边缘保持势函数融合到关于相位的正则项中, 在解决相位缠绕问题的同时保持了图像的边缘, 提高了重建图像的质量.

在上述思想的启发下, 本文分别对复图像的幅值和相位构造正交字典下的稀疏正则项. 令 $\mathbf{m} \in \mathbb{R}^M$ 和 $\boldsymbol{\varphi} \in \mathbb{R}^M$ 分别表示幅值和相位, 则复图像 $\mathbf{u}_0$ 可表示为: $\mathbf{u}_0 = \mathbf{m} e^{j\boldsymbol{\varphi}}$, 优化问题(5)转化为以下形式:

$$\min_{\mathbf{m}, \boldsymbol{\varphi}, \mathbf{F}_1, \mathbf{F}_2, \mathbf{D}_1, \mathbf{D}_2} \sum_{r=1}^K \beta_r \|\sqrt{\sigma_r} - |\mathbf{F}_r(\mathbf{m} e^{j\boldsymbol{\varphi}})|\|_2^2 + \sum_{i=1}^N \|\mathbf{R}_i \mathbf{m} - \mathbf{D}_1 \mathbf{a}_i\|_2^2 + \tau_1 \|\mathbf{a}_i\|_0 + \sum_{i=1}^N \|\mathbf{R}_i \boldsymbol{\varphi} - \mathbf{D}_2 \mathbf{b}_i\|_2^2 + \tau_2 \|\mathbf{b}_i\|_0$$

s. t. $\mathbf{D}_1^T \mathbf{D}_1 = \mathbf{I}, \mathbf{D}_2^T \mathbf{D}_2 = \mathbf{I}$ (6)

其中, $\mathbf{D}_1 \in \mathbb{R}^{n \times n}$ 和 $\mathbf{D}_2 \in \mathbb{R}^{n \times n}$ 分别表示对复图像的幅值和相位自适应的正交字典; $\mathbf{a}_i$ 和 $\mathbf{b}_i$ 分别表示幅值和相位在对应字典下的表示系数; $\mathbf{F}_1 = \{\mathbf{a}_1, \dots, \mathbf{a}_N\}$ 和 $\mathbf{F}_2 = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N\}$ 分别是幅值和相位表示系数的集合; 式(6)中第二项表示幅值在正交字典 $\mathbf{D}_1$ 下进行稀疏表示的误差; 第三项为幅值的稀疏正则项; 第四项表示相位在正交字典 $\mathbf{D}_2$ 下进行稀疏表示的误差; 第五项为相位的稀疏正则项.

本文采用分裂 Bregman 方法^[25]对式(6)中的优化问题进行求解, 令 $\mathbf{d}_1 = \mathbf{m}$, $\mathbf{d}_2 = \boldsymbol{\varphi}$, 并利用 Bregman 迭代求解得到的优化问题, 则式(6)中的优化问题转化为以下形式:

$$\min_{\mathbf{m}, \boldsymbol{\varphi}, \mathbf{F}_1, \mathbf{F}_2, \mathbf{D}_1, \mathbf{D}_2, \mathbf{d}_1, \mathbf{d}_2} D_E^P(\mathbf{m}, \boldsymbol{\varphi}, \mathbf{d}_1, \mathbf{d}_2, \mathbf{m}^t, \boldsymbol{\varphi}^t, \mathbf{d}_1^t, \mathbf{d}_2^t) + \frac{\lambda_2}{2} \|\mathbf{d}_1 - \mathbf{m}\|_2^2 + \frac{\lambda_3}{2} \|\mathbf{d}_2 - \boldsymbol{\varphi}\|_2^2$$

s. t. $\mathbf{D}_1^T \mathbf{D}_1 = \mathbf{I}, \mathbf{D}_2^T \mathbf{D}_2 = \mathbf{I}$ (7)

式中 $D_E^P(\mathbf{m}, \boldsymbol{\varphi}, \mathbf{d}_1, \mathbf{d}_2, \mathbf{m}^t, \boldsymbol{\varphi}^t, \mathbf{d}_1^t, \mathbf{d}_2^t) = E(\mathbf{m}, \boldsymbol{\varphi}, \mathbf{d}_1, \mathbf{d}_2) - \langle \mathbf{p}_{\mathbf{m}}^t, \mathbf{m} - \mathbf{m}^t \rangle + \langle \mathbf{p}_{\mathbf{d}_1}^t, \mathbf{d}_1 - \mathbf{d}_1^t \rangle - \langle \mathbf{p}_{\boldsymbol{\varphi}}^t, \boldsymbol{\varphi} - \boldsymbol{\varphi}^t \rangle + \langle \mathbf{p}_{\mathbf{d}_2}^t, \mathbf{d}_2 - \mathbf{d}_2^t \rangle$ 表示 $E(\mathbf{m}, \boldsymbol{\varphi}, \mathbf{d}_1, \mathbf{d}_2)$ 对应的 Bregman 距离, 其中

$$E(\mathbf{m}, \boldsymbol{\varphi}, \mathbf{d}_1, \mathbf{d}_2) = \sum_{r=1}^K \beta_r \|\sqrt{\sigma_r} - |\mathbf{F}_r(\mathbf{m} e^{j\boldsymbol{\varphi}})|\|_2^2 + \sum_{i=1}^N \|\mathbf{R}_i \mathbf{d}_1 - \mathbf{D}_1 \boldsymbol{\alpha}_{m,i}\|_2^2 + \tau_1 \|\boldsymbol{\alpha}_{m,i}\|_0 + \sum_{i=1}^N \|\mathbf{R}_i \mathbf{d}_2 - \mathbf{D}_2 \boldsymbol{\alpha}_{\varphi,i}\|_2^2 + \tau_2 \|\boldsymbol{\alpha}_{\varphi,i}\|_0,$$

$\mathbf{p}_{(\cdot)}^t$ 表示 $E(\mathbf{m}, \boldsymbol{\varphi}, \mathbf{d}_1, \mathbf{d}_2)$ 对相关变量的偏导数. 交替迭代优化 $\mathbf{m}, \boldsymbol{\varphi}, \mathbf{F}_1, \mathbf{F}_2, \mathbf{D}_1, \mathbf{D}_2, \mathbf{d}_1, \mathbf{d}_2$ 以获得重建图像 $\hat{\mathbf{u}}_0$, 具体求解过程包括以下几个步骤:

(1) 更新幅值 $\mathbf{m}$ 和相位 $\boldsymbol{\varphi}$:

$$\{\mathbf{m}^{t+1}, \boldsymbol{\varphi}^{t+1}\} = \arg \min_{\mathbf{m}, \boldsymbol{\varphi}} \left\{ \sum_{r=1}^K \beta_r \|\sqrt{\sigma_r} - |\mathbf{F}_r(\mathbf{m} e^{j\boldsymbol{\varphi}})|\|_2^2 + \frac{\lambda_2}{2} \|\mathbf{d}_1^t - \mathbf{m}\|_2^2 - \langle \mathbf{p}_{\mathbf{m}}^t, \mathbf{m} - \mathbf{m}^t \rangle + \frac{\lambda_3}{2} \|\mathbf{d}_2^t - \boldsymbol{\varphi}\|_2^2 - \langle \mathbf{p}_{\boldsymbol{\varphi}}^t, \boldsymbol{\varphi} - \boldsymbol{\varphi}^t \rangle \right\} \quad (8)$$

式(8)是一个非凸优化问题, 计算复杂. 针对这类问

题的非凸性和求解的复杂性,文献[26-27]提出利用基于半正定松弛的方法进行求解,此类方法通过矩阵提升将非凸问题转化为一个凸半正定规划问题,通过求解该问题以恢复原始信号.但由于矩阵提升会增大矩阵的维数,此类方法不适用于规模较大的优化问题.MM算法^[28]通过最大化代价函数中的某一项或多项构造辅助函数,将复杂的原始非凸优化问题转化为易处理的优化问题.迭代求解辅助函数,得到的一系列最终收敛于原始优化问题最优解的驻点,并且在每次迭代过程中只需要进行基本的矩阵操作即可得到辅助函数的闭式解,适用于规模较大的优化问题.基于MM算法的上述优点,本文利用MM算法对含有数据保真项的子问题进行优化,以减小由于该项的非凸性所造成的误差.

优化问题(8)的非凸性主要由第一项引起,为这一项构造辅助函数^[29].则该优化问题转化为以下形式:

$$\begin{aligned} \{m^{t+1}, \varphi^{t+1}\} = \arg \min_{m, \varphi} \left\{ \sum_{r=1}^K \beta_r \|m e^{j\varphi} - v_r^t\|_2^2 + \right. \\ \left. \frac{\lambda_2}{2} \|d_1^t - m\|_2^2 - \langle p_m^t, m - m^t \rangle + \right. \\ \left. \frac{\lambda_3}{2} \|d_2^t - \varphi\|_2^2 - \langle p_\varphi^t, \varphi - \varphi^t \rangle \right\} \quad (9) \end{aligned}$$

其中, $v_r^t = F_r^H (P \sqrt{o_r})$, F_r^H 表示非涅尔逆变换, $P = \text{Diag}\{e^{j \arg\{F_r(m^t e^{j\varphi^t})\}}\}$ 是一个对角阵,该对角阵主对角线上的元素是向量 $e^{j \arg\{F_r(m^t e^{j\varphi^t})\}}$ 的元素, $\arg\{\cdot\}$ 表示相位.

固定相位 φ^t , 更新上述优化问题中变量 m 的方法是将其对 m 求偏导,并令偏导数为零,从而得到 m 的最优解:

$$m^{t+1} = \frac{\lambda_2 d_1^t + p_m^t + 2 \sum_{r=1}^K \beta_r E^H v_r^t}{\lambda_2 + 2 \sum_{r=1}^K \beta_r} \quad (10)$$

其中, $E = \text{Diag}(e^{j\varphi^t})$ 是一个关于相位 φ^t 的对角阵.

固定 m^{t+1} , 式(8)中关于相位 φ 的代价函数 $L(\varphi) = \sum_{r=1}^K \beta_r \|m^{t+1} e^{j\varphi} - v_r^t\|_2^2 + \frac{\lambda_3}{2} \|d_2^t - \varphi\|_2^2 - \langle p_\varphi^t, \varphi - \varphi^t \rangle$ 是关于 φ 的非线性函数,本文利用梯度下降法进行求解第 q 次迭代:

$$\varphi^q = \varphi^{q-1} - \alpha \cdot \nabla L(\varphi^{q-1}) \quad (11)$$

其中, α 为步长, $\nabla L(\varphi^{q-1})$ 是代价函数 $L(\varphi)$ 在点 φ^{q-1} 的梯度方向,计算如下:

$$\nabla L(\varphi) = \lambda_3 (\varphi - d_2^t) - p_\varphi^t +$$

$$\sum_{r=1}^K 2 \text{Re}\{j e^{-j\varphi} \odot (\beta_r M^H v_r^t) - j e^{-j\varphi} \odot (\beta_r M^H M e^{j\varphi})\} \quad (12)$$

其中, $\odot$ 表示逐元素相乘, 变量 $M = \text{Diag}(m^{t+1})$ 是一个表示幅值 m^{t+1} 的对角阵, Re 表示取实部.

(2) 更新稀疏系数和正交字典:

固定其它变量,更新稀疏编码和正交字典的过程相当于正交字典学习的过程.为了简便起见,该过程对应的优化问题可以写成以下形式:

$$\begin{aligned} \min_{\{d_l^t, A_l\}} \sum_{l=1}^2 \|X_l - D_l A_l\|_F^2 + \tau_l \|A_l\|_0 \\ \text{s. t. } D_l^T D_l = I, l=1, 2 \end{aligned} \quad (13)$$

其中, $X_1 = [R_1 d_1, \dots, R_N d_1]$ 是图像 d_1 的块矩阵,它的每一列对应一个图像块; $X_2 = [R_1 d_2, \dots, R_N d_2]$ 是图像 d_2 的块矩阵; $A_1 = [\alpha_{m,1}, \dots, \alpha_{m,N}]$ 是图像的幅值在正交字典 D_1 下的表示系数矩阵, $A_2 = [\alpha_{\varphi,1}, \dots, \alpha_{\varphi,N}]$ 是相位在正交字典 D_2 下的表示系数矩阵; $\|\cdot\|_F$ 为矩阵的 Frobenius 范数.利用交替优化方法分两步对上述优化问题进行求解.对第 s 次迭代:

① 稀疏编码:

固定字典 D_l^{s-1} 和图像块 X_l^t , 通过求解以下稀疏编码优化问题得到稀疏系数 A_l 的最优解:

$$A_l^t = \arg \min_{A_l} \{\|X_l^t - D_l^{s-1} A_l\|_F^2 + \tau_l \|A_l\|_0\}, l=1, 2 \quad (14)$$

由于 D_l^{s-1} 的正交性,式(14)中的优化问题可根据硬阈值(hard-thresholding)法求解:

$$A_l^s = H_{\sqrt{\tau_l}}[(D_l^{s-1})^T X_l^t] \quad (15)$$

其中 $H_{\sqrt{\tau}}(\cdot)$ 表示硬阈值操作算子,其定义如下:

$$H_{\sqrt{\tau}}(y) = \begin{cases} y, & |y| > \sqrt{\tau} \\ 0, & |y| \leq \sqrt{\tau} \end{cases} \quad (16)$$

式(16)表示对矩阵 $(D_l^{s-1})^T X_l^t$ 进行逐元素处理.

② 正交字典更新:

固定图像块 X_l^t 和稀疏系数 A_l^s , 通过解以下带约束的优化问题更新正交字典 D_l :

$$\begin{aligned} D_l^s = \arg \min_{D_l} \{\|X_l^t - D_l A_l^s\|_F^2\} \\ \text{s. t. } D_l^T D_l = I, l=1, 2 \end{aligned} \quad (17)$$

式(17)中带约束的优化问题有唯一解,形式如下^[20]:

$$D_l^s = U_l V_l^T \quad (18)$$

其中, U_l 和 V_l 均为正交矩阵,满足下列的奇异值分解过程:

$$X_l^t (A_l^s)^T = U_l \Sigma_l V_l^T \quad (19)$$

其中, Σ_l 是由 $X_l^t (A_l^s)^T$ 的奇异值构成的对角阵.

③ 更新中间变量 d_1, d_2 :

$$\{d_1^{t+1}, d_2^{t+1}\} = \arg \min_{d_1, d_2} \left\{ \frac{\lambda_2}{2} \|d_1 - m^{t+1}\|_2^2 + \frac{\lambda_3}{2} \|d_2 - \phi^{t+1}\|_2^2 + \sum_{i=1}^N \|R_i d_1 - D_1^{t+1} \alpha_{m,i}^{t+1}\|_2^2 + \sum_{i=1}^N \|R_i d_2 - D_2^{t+1} \alpha_{\phi,i}^{t+1}\|_2^2 + \langle p_{d_1}^t, d_1 - d_1^t \rangle + \langle p_{d_2}^t, d_2 - d_2^t \rangle \right\} \quad (20)$$

上述优化问题含有关于 d_1 和 d_2 的简单二次项, 其闭式解分别为

$$d_1^{t+1} = \frac{\lambda_2 m^{t+1} - p_{d_1}^t + 2 \sum_{i=1}^N R_i^T D_1^{t+1} \alpha_{m,i}^{t+1}}{\lambda_2 I + 2 \sum_{i=1}^N R_i^T R_i} \quad (21)$$

$$d_2^{t+1} = \frac{\lambda_3 \phi^{t+1} - p_{d_2}^t + 2 \sum_{i=1}^N R_i^T D_2^{t+1} \alpha_{\phi,i}^{t+1}}{\lambda_3 I + 2 \sum_{i=1}^N R_i^T R_i} \quad (22)$$

综上所述, 本文提出的基于正交字典学习的多像面相位恢复算法 (Multi-plane Phase Retrieval Based on Orthogonal Dictionary Learning, MPR-ODL) 的具体实现过程归纳如下:

MPR-ODL 算法.

输入: 观测强度值 $o_r, r=1, \dots, K$, 图像的初始估计值 u_0^0 , 初始字典 D^0

输出: 重建图像 $\hat{u}_0$

1. 初始化: $m^0 = \text{abs}(u_0^0)$, $\phi^0 = \text{angle}(u_0^0)$, $d_1^0 = m^0$, $d_2^0 = \phi^0$, $D_1^0 = D_2^0 = D^0$, $b_1^0 = b_2^0 = 0$, 相位更新的最大迭代次数 $Q=5$, 正交字典学习的最大迭代次数 $S=100$, 算法整体迭代次数 $iter$

2. FOR $t=0$ TO $iter-1$

(1) 根据式(10)更新图像的幅值 m^{t+1} ;

(2) FOR $q=1$ TO Q

$$\phi^q = \phi^{q-1} - \gamma \cdot \nabla L(\phi^{q-1});$$

END FOR

$$\phi^{t+1} = \phi^Q;$$

(3) FOR $s=1$ TO S

1) 根据式(15)更新稀疏编码 $A_l^s, l=1, 2$;

2) 根据式(18)更新正交字典 $D_l^s, l=1, 2$;

END FOR

$$A_l^{t+1} = A_l^S, D_l^{t+1} = D_l^S, l=1, 2;$$

(4) 根据式(21)、(22)更新中间变量 d_1^{t+1}, d_2^{t+1} ;
END FOR

3. 输出: 重建图像 $\hat{u}_0 = m^{iter} e^{j\phi^{iter}}$.

本文将多像面相位恢复过程分为两个阶段, 第一阶段不利用稀疏正则项, 仅通过迭代求解式(4)对

原始图像进行重建, 得到原始图像的粗略估计; 第二阶段将第一阶段的重建结果作为初始值, 利用图像在正交字典下的稀疏性对原始图像进行精确重建. 由于正交字典学习时间复杂度相对较高, 为提高图像重建速度, 本文在实现过程中先利用正交 DCT 字典对图像进行稀疏表示, 由大量实验获得以下经验值: 从第 $(iter \times 4/5)$ 次迭代开始, 每隔 $(iter/50)$ 次进行一次正交字典学习, 既能保证重建结果具有较高的精确度, 又能避免大量的时间损耗.

4 实验结果

为了测试本文算法的性能, 本文针对不同像面个数以及不同噪声分别进行相位恢复实验, 并将实验结果与 AL 算法和 4f-SPAR 算法的实验结果进行对比和分析.

为了便于和 AL 算法以及 4f-SPAR 算法进行比较, 本文中复图像的构造方法和文献[15]一致, 即 $u_0 = m e^{j\phi}$, $m = w_1 + 1$, $\phi = \pi(w_2 - 1/2)/5$, 其中 w_1 和 w_2 分别为构成复图像幅值和相位的测试图像, 图 1 给出了本文用于仿真实验的标准灰度图像. 另外, 本文提出的算法利用 4f 光学成像系统空间光调制器模拟多像面成像, 该模型的参数设置和文献[15]保持一致, 即物面和像面上的像素大小为 $\Delta_x = 3.45 \mu\text{m}$, 空间光调制器上的像素大小为 $\Delta_v = 8.0 \mu\text{m}$, $\lambda = 0.532 \mu\text{m}$, $z_1 = f$, f 表示傅里叶透镜的焦距, $\Delta_z = 20 \text{ mm}$. 采用均方根误差 (Root-Mean-Square Error, RMSE) 作为客观评价标准, 采用结构相似性^[30] (Structural Similarity, SSIM) 作为视觉效果评价标准.

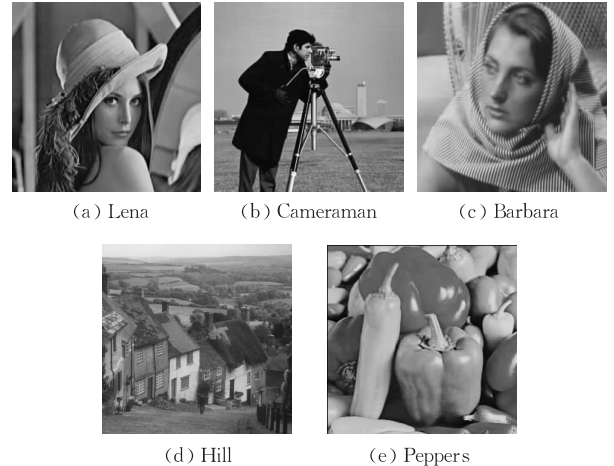


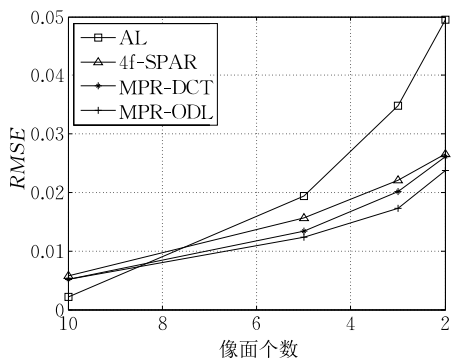
图 1 用于测试的标准图像

本文使用正交 DCT (Discrete Cosine Transform) 字典作为正交字典学习的初始字典 D^0 . 为了验证正

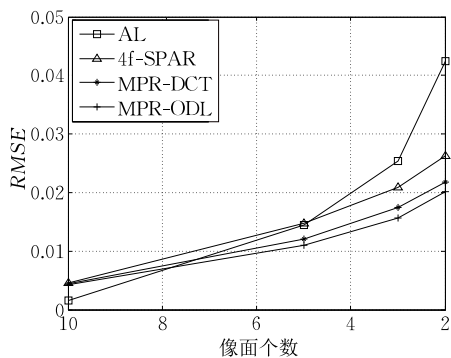
交字典学习的优越性,在实验仿真中,用固定正交 DCT 字典替换本文算法中的自适应正交字典进行图像重建,为便于描述,将该算法称之为基于正交 DCT 字典的多像面相位恢复算法 (Multi-plane Phase Retrieval Based on DCT, MPR-DCT),并将实验结果和 MPR-ODL 算法进行对比和分析。

4.1 无噪声情况

为验证算法在不同像面个数时图像重建的性能,本文在不含噪声情况下,首先将图像 Lena 作为幅值, Cameraman 作为相位构造复图像,针对像面个数 $K=2, 3, 5, 10$ 分别利用 AL 算法、4f-SPAR 算法、MPR-DCT 算法以及 MPR-ODL 算法对该复图像进行相位恢复仿真实验,为保证算法比较的公平性,对每种算法独立运行 3 次,选取 3 次实验结果的平均值作为重建结果进行对比和分析。图 2 给出了上述 4 种算法的幅值重建图像和相位重建图像对应的 RMSE 数值随像面个数变化的曲线图。其中,当像面个数为 10 时,4 种算法的迭代次数均设置为 500 次。由于像面个数减少,观测值减少,相位恢复的难度增大,迭代次数应逐渐增加才能保证算法的收敛性。通过大量实验,获得以下经验值:当像面个数 K 依次取 5、3、2 时,4 种算法的迭代次数依次设



(a) 幅值图像Lena



(b) 相位图像Cameraman

图2 不同像面个数下4种算法对复图像进行相位恢复实验结果 RMSE 曲线图

置为 1000 次、2000 次、5000 次。另外,当像面个数为 10 时,本文 MPR-ODL 算法中的其它参数在以下区间内取值: $\beta_r \in [1.1, 2.5]$ ($r=1, \dots, K$), $\lambda_2 \in [0.1, 0.64]$, $\lambda_3 \in [0.18, 1]$, $\tau_1 \in [0.006, 0.17]$, $\tau_2 \in [0.01, 0.17]$, 步长 $\gamma \in [0.015, 0.064]$ 。随着像面个数减少,进行多像面相位恢复的已知信息减少,因此,用于描述测量值和衍射成像模值之间误差的数据保真项所占的比重应逐渐增大,以确保重建图像的可靠性;利用图像稀疏先验知识的正则项所占的比重应逐渐增大,以提高重建图像的质量;所以参数 β_r 、 τ_1 和 τ_2 随像面个数的减少而增大。当观测值含噪声时,测量值和衍射成像模值之间的误差主要由噪声导致,随着观测值噪声含量增大,用于衡量噪声大小的数据保真项所占的比重应逐渐减小,参数 β_r 逐渐减小;用于去噪的稀疏正则项所占的比例应逐渐增大,参数 λ_2 和 λ_3 、 τ_1 和 τ_2 应逐渐增大,以确保算法的抗噪性。

由图 2 可知,4 种算法的 RMSE 数值在像面个数为 10 时都较低,说明 4 种算法均能得到高质量的重建图像。随着像面个数减少,相位恢复难度增大,4 种算法的重建结果所对应的 RMSE 值逐渐上升,SSIM 值逐渐下降。其中,AL 算法对像面个数最敏感,当像面个数减小时,该算法对应的 RMSE 上升最为明显;4f-SPAR 算法和 MPR-DCT 算法对应的 RMSE 上升趋势比 AL 算法平缓;MPR-ODL 算法对应的 RMSE 随像面个数减小上升最为平缓,并且在像面个数 $K=5, 3, 2$ 时均低于另外 3 种算法。总体上来看,本文提出的 MPR-ODL 算法重建效果优于其它 3 种算法。

为了更加直观地显示 MPR-ODL 算法的优越性,图 3 和图 4 分别给出了 2 个像面时,AL 算法、4f-SPAR 算法、MPR-DCT 算法以及 MPR-ODL 算法对幅值 Lena 和相位 Cameraman 的重建图像和部分重建细节。

由图 3 和图 4 可以看出,当像面个数为 2 时,AL 算法的重建图像质量最差,图像中存在大量伪迹,由于该算法对复图像整体应用 Tikhonov 正则化约束,导致重建幅值和相位之间互相干扰严重;相比于 AL 算法,4f-SPAR 算法的重建效果有较大提升,由于该算法对图像的幅值和相位分别构造 BM3D 框架下的稀疏正则项,幅值和相位之间的相互干扰减小,但仍然存在较多的伪迹;此外,重建图像丢失了部分细节,例如相位中的草地区域被过度平滑;MPR-DCT 算法的重建图像中的伪迹进一步减少,并保留了部分细节,如图 4(g) 所示。



图 3 4 种算法对幅值和相位的重建结果(RMSE,SSIM)($K=2$)

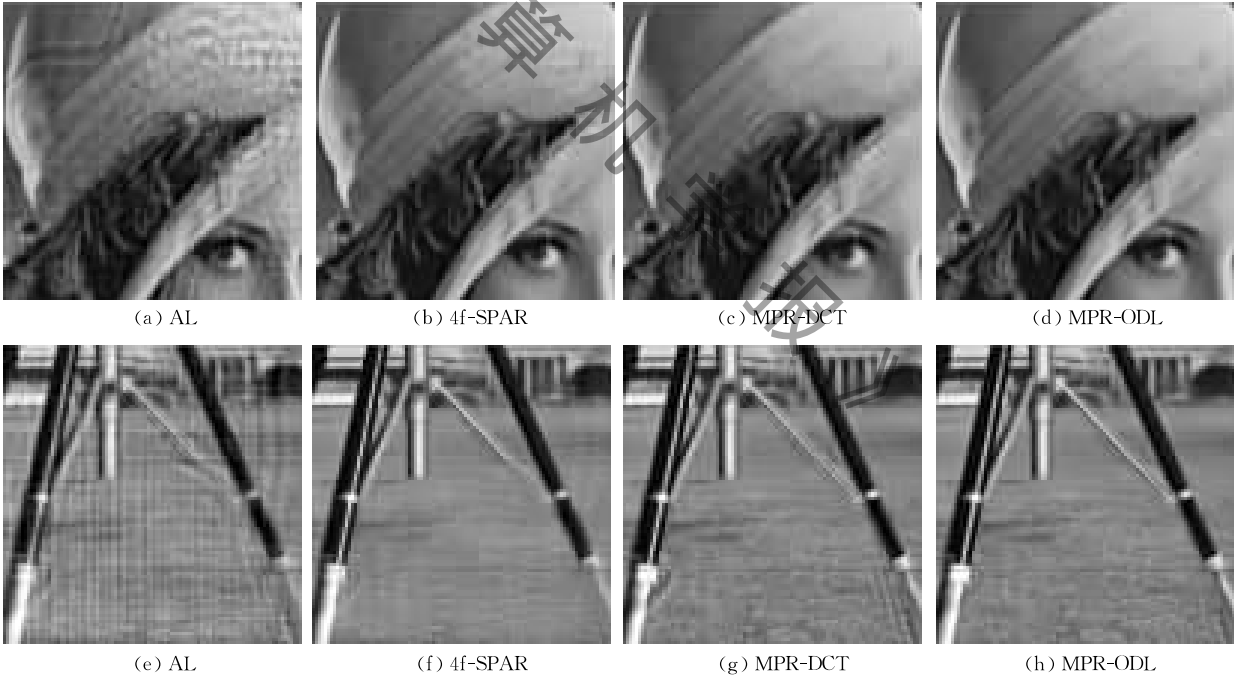
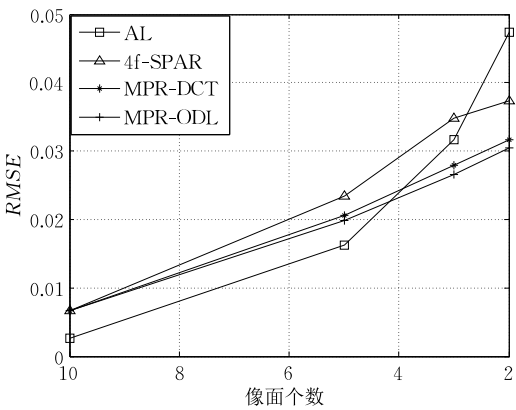


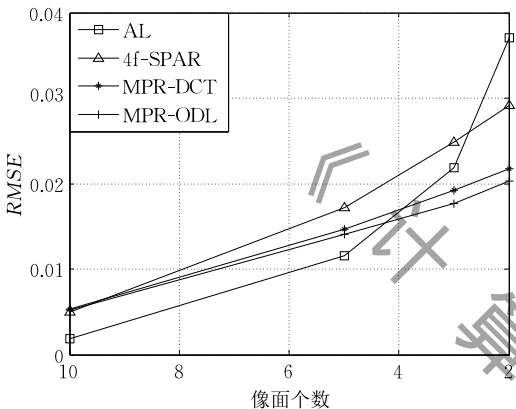
图 4 4 种算法对幅值和相位的重建细节($K=2$)

在 MPR-DCT 算法的基础上,MPR-ODL 算法进行正交字典学习,利用与待重建图像匹配的正交字典对幅值和相位分别进行稀疏表示.与 MPR-DCT 算法相比,该算法的重建图像中含有更少的伪迹,在图中的线特征边缘区域尤为明显;与 4f-SPAR 算法相比,本文算法能够保留图像的部分细节;因此,在像面个数 $K=2$ 时,本文提出的 MPR-ODL 算法仍然能够实现有效重建.

为验证本文算法对不同类型的复图像均具有良好的重建效果,选取 2 幅自然图像构造复图像并利用上述 4 种算法进行仿真实验.将标准灰度图像 Hill 截取大小为 256×256 的一部分作为幅值,标准灰度图像 Peppers 作为相位构造复图像,针对像面个数 $K=2,3,5,10$ 分别利用 AL 算法、4f-SPAR 算法、MPR-DCT 算法以及 MPR-ODL 算法对该复图像进行相位恢复仿真实验,并将实验结果进行对比和分析.图 5



(a) 幅值图像Lena



(b) 相位图像Cameraman

图 5 不同像面个数下,4 种算法对复图像进行相位恢复实验结果的 RMSE 曲线图

给出了上述 4 种算法的幅值重建图像和相位重建图像对应的 $RMSE$ 数值随像面个数变化的曲线图。

由图 5 可以看出,该图中 4 种算法的重建结果对应的 $RMSE$ 数值变化曲线与图 2 的基本保持一致,即 AL 算法对像面个数最敏感,对应的 $RMSE$ 值随着像面个数减少急剧上升;MPR-ODL 算法对应的 $RMSE$ 值随像面个数减少变化最平缓,并且在像面个数 ≤ 4 时该算法对应的 $RMSE$ 值最小,算法性能最好,从而验证了本文算法对不同类型的复图像均具有优于其它 3 种算法的重建效果。

为了直观地显示算法的重建效果,图 6 和图 7 分别给出了 2 个像面时,4 种算法对幅值 Hill 和相位 Peppers 的重建图像和部分重建细节。

由上述两组仿真实验可知,相比于 AL 算法和 4f-SPAR 算法,MPR-DCT 算法和 MPR-ODL 算法在重建效果上存在比较明显的优势。由于自适应字典能够捕获图像的结构信息,因此,相比于利用固定正交 DCT 字典的 MPR-DCT 算法,利用自适应正交字典的 MPR-ODL 算法能够更好地恢复图像细节,为进一步验证自适应正交字典的这一优越性,将含有丰富纹理的标准灰度图像 Barbara 像截取大小为 256×256 的一部分作为幅值,标准灰度图像 Cameraman 作为相位,在像面个数 $K=2$ 时对上述 4 种算法进行仿真实验。图 8 给出了 4 种算法的重建图像,图 9 给出了 MPR-DCT 算法和 MPR-ODL 算法的部分重建细节。



图 6 4 种算法对幅值和相位的重建结果($RMSE$, $SSIM$) ($K=2$)

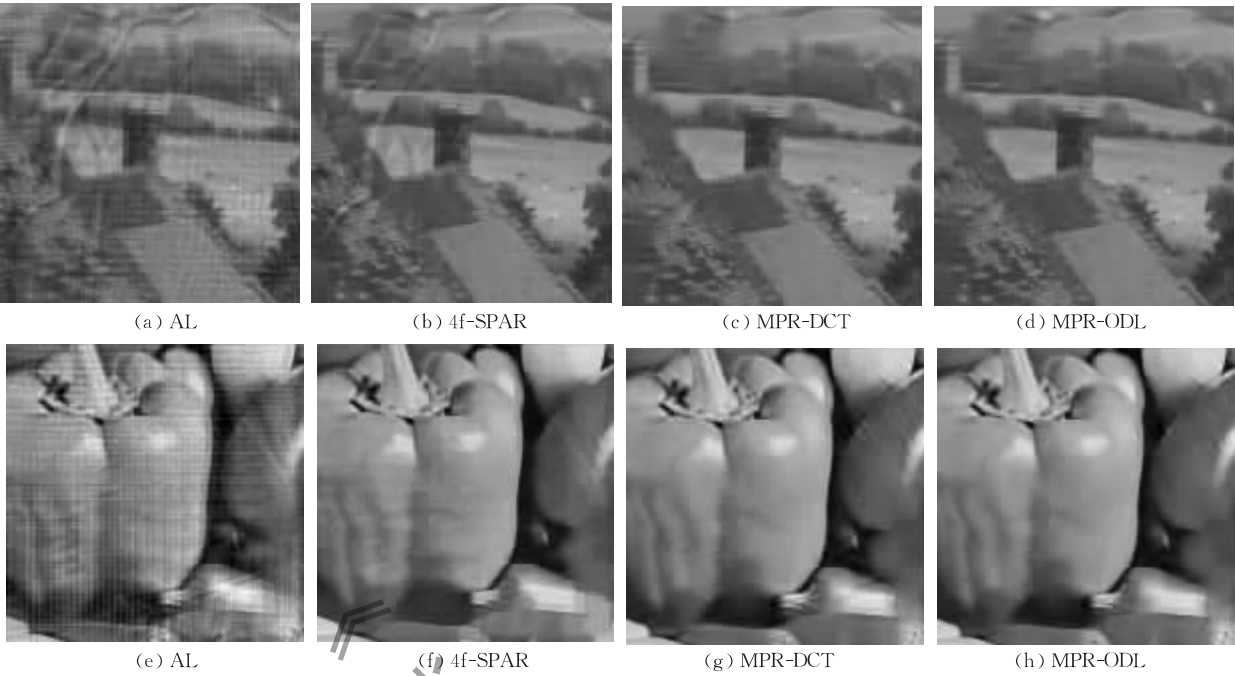


图 7 4 种算法对幅值和相位的重建细节($K=2$)



图 8 4 种算法对幅值和相位的重建结果($RMSE, SSIM$) ($K=2$)

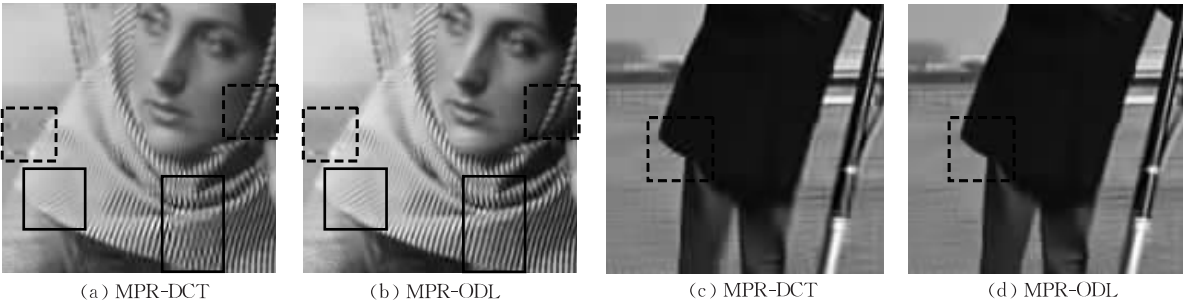


图 9 2 种算法对幅值和相位的重建细节($K=2$)

比较图 8 中的重建图像可知,AL 算法重建效果最差;相比于 4f-SPAR 算法,MPR-DCT 算法和 MPR-ODL 算法的重建图像中含有更少的伪迹.

观察图 9 中的重建图像可以看出,与图(a)相比,图(b)中的实线矩形区域内包含更多的纹理并且纹理更加清晰;与 MPR-DCT 算法相比,图 MPR-ODL 算法中的虚线矩形区域内含有更少的伪迹.结合上述两点可知,MPR-ODL 算法能够更好地恢复原始图像的细节,重建效果优于 MPR-DCT 算法,从而验证了自适应正交字典的优越性.

4.2 含噪情况

在进行实际的相位恢复时,观测值往往是含有噪

声的,为测试本文算法对噪声的鲁棒性,本文首先在强度观测值中加入高斯白噪声,此时强度观测值表示如下:

$$\boldsymbol{o}_r = |\boldsymbol{u}_r|^2 + \boldsymbol{\varepsilon}_r, \quad r=1, \cdots, K \quad (23)$$

其中, $\boldsymbol{\varepsilon}_r$ 表示第 r 个观测面上标准差为 σ_r 的高斯白噪声,假设每个像面的噪声强度相同,即 $\sigma_r = \sigma$. 本文在噪声强度 $\sigma=0.05、0.1$ 以及 0.15 时分别对 4 种算法进行多像面相位恢复仿真实验,对每种算法独立运行 3 次,选取 3 次实验结果的平均值作为重建结果进行对比和分析.利用图像 Lena 和 Cameraman 构造复图像,表 1 和表 2 分别给出了 4 种算法的 RMSE 和 SSIM 数值比较.

表 1 不同噪声强度下对复图像进行相位恢复实验结果 RMSE 比较

噪声强度 (σ)	像面 个数	幅值图像 Lena				相位图像 Cameraman			
		AL	4f-SPAR	MPR-DCT	MPR-ODL	AL	4f-SPAR	MPR-DCT	MPR-ODL
0.05	10	0.0313	0.0164	0.0142	0.0136	0.0241	0.0124	0.0113	0.0106
	5	0.0693	0.0231	0.0203	0.0191	0.0551	0.0198	0.0162	0.0151
	3	0.0937	0.0274	0.0267	0.0240	0.0806	0.0243	0.0215	0.0200
	2	0.1160	0.0384	0.0319	0.0301	0.1089	0.0307	0.0261	0.0246
0.10	10	0.0550	0.0235	0.0228	0.0217	0.0461	0.0176	0.0181	0.0170
	5	0.1007	0.0314	0.0294	0.0277	0.0878	0.0258	0.0234	0.0218
	3	0.1231	0.0363	0.0348	0.0329	0.1105	0.0332	0.0277	0.0260
	2	0.1461	0.0448	0.0387	0.0371	0.1337	0.0432	0.0313	0.0301
0.15	10	0.0755	0.0307	0.0297	0.0283	0.0653	0.0228	0.0238	0.0223
	5	0.1166	0.0388	0.0354	0.0337	0.1113	0.0316	0.0283	0.0266
	3	0.1436	0.0443	0.0403	0.0388	0.1293	0.0422	0.0325	0.0306
	2	0.1670	0.0512	0.0492	0.0479	0.1533	0.0499	0.0385	0.0372

表 2 不同噪声强度下对复图像进行相位恢复实验结果 SSIM 比较

噪声强度 (σ)	像面 个数	幅值图像 Lena				相位图像 Cameraman			
		AL	4f-SPAR	MPR-DCT	MPR-ODL	AL	4f-SPAR	MPR-DCT	MPR-ODL
0.05	10	0.7696	0.9320	0.9511	0.9529	0.6658	0.8971	0.9293	0.9336
	5	0.5192	0.8898	0.9186	0.9204	0.4098	0.8370	0.9004	0.9048
	3	0.4095	0.8673	0.8950	0.9054	0.3081	0.8068	0.8577	0.8665
	2	0.3490	0.8191	0.8666	0.8754	0.2427	0.7327	0.8241	0.8342
0.10	10	0.5693	0.8851	0.9039	0.9068	0.4593	0.8377	0.8602	0.8662
	5	0.3868	0.8225	0.8629	0.8643	0.2908	0.7507	0.8064	0.8095
	3	0.3427	0.8021	0.8405	0.8480	0.2479	0.7028	0.7823	0.7893
	2	0.2936	0.6912	0.8236	0.8324	0.2087	0.5412	0.7541	0.7616
0.15	10	0.4463	0.8329	0.8638	0.8681	0.3490	0.7737	0.8087	0.8138
	5	0.3329	0.7697	0.8310	0.8382	0.2386	0.6862	0.7649	0.7719
	3	0.2900	0.7498	0.8004	0.8105	0.2126	0.6161	0.7353	0.7421
	2	0.2571	0.6496	0.7868	0.7929	0.1823	0.4862	0.7128	0.7154

由表 1、表 2 可知,在上述噪声强度下,AL 算法对应的 RMSE 值最高,SSIM 值最低,说明其对高斯噪声的抗噪性最差;与 AL 算法相比,4f-SPAR 算法和 MPR-DCT 算法的抗噪性有较大的提升;本文 MPR-ODL 算法对应的 RMSE 值最低,SSIM 值最高,说明其恢复效果最好.因此,本文提出的 MPR-ODL 算法对高斯白噪声具有更好的鲁棒性.

高斯噪声的概率密度函数服从正态分布,主要

来源在图像采集期间;泊松噪声是由光具有量子特性而造成的测量不确定性引起的,二者对强度测量值影响不同.为验证本文算法对不同类型噪声具有鲁棒性,在强度观测值中加入泊松噪声,衡量观测值中泊松噪声强度的指标 R_{noise} ^[31] 定义如下:

$$R_{\text{noise}} = \frac{\| |\boldsymbol{F}_r \boldsymbol{u}_0|^2 - \boldsymbol{o}_r \|_1}{\| |\boldsymbol{F}_r \boldsymbol{u}_0|^2 \|_1}, \quad r=1, \cdots, K \quad (24)$$

在像面个数 $K=5$, 观测值中泊松噪声强度

$R_{\text{noise}}=10\%$ 时,分别利用 AL 算法、4f-SPAR 算法、MPR-DCT 算法以及 MPR-ODL 算法对复图像进行多像面相位恢复仿真实验,对每种算法独立运行 3 次,选取 3 次实验结果的平均值作为重建结果进行对比和分析.利用图像 Lena 和 Cameraman 构造复图像,表 3 给出了上述 4 种算法的重建结果对应的 $RMSE$ 和 $SSIM$ 数值比较.

由表 3 可知,根据 $RMSE$ 和 $SSIM$ 两个评价指标,AL 算法对泊松噪声的鲁棒性最差,另外 3 种算法对泊松噪声的鲁棒性明显提升,其中本文提出的 MPR-ODL 算法抗噪性最强.为了直观地显示各算法的抗噪性,图 10 给出了 $K=5,R_{\text{noise}}=10\%$ 时 4 种算法的重建图像.由图 10 可以看出,AL 算法对应的重建图像中含有大量的噪声;4f-SPAR 算法能有效去除噪声,但重建图像 Lena 中存在块效应,重建

图像 Cameraman 中存在比较明显的失真,例如相机架变细、变暗等;MPR-DCT 算法去除了大部分噪声,但重建图像中包含较多的伪迹,并出现模糊;MPR-ODL 算法能有效去噪,相比于 4f-SPAR 算法,该算法更加平滑,没有产生明显的失真,重建图像中包含更多细节;相比于 MPR-DCT 算法,该算法的重建图像含有更少的伪迹,视觉效果更好.因此,相比于其它 3 种算法,本文提出的 MPR-ODL 算法对泊松噪声具有更强的鲁棒性.

表 3 4 种算法对复图像进行相位恢复实验结果比较				
算法	幅值图像 Lena		相位图像 Cameraman	
	$RMSE$	$SSIM$	$RMSE$	$SSIM$
AL	0.1300	0.2668	0.1270	0.1937
4f-SPAR	0.0445	0.7480	0.0360	0.6747
MPR-DCT	0.0395	0.8059	0.0312	0.7642
MPR-ODL	0.0376	0.8133	0.0292	0.7811

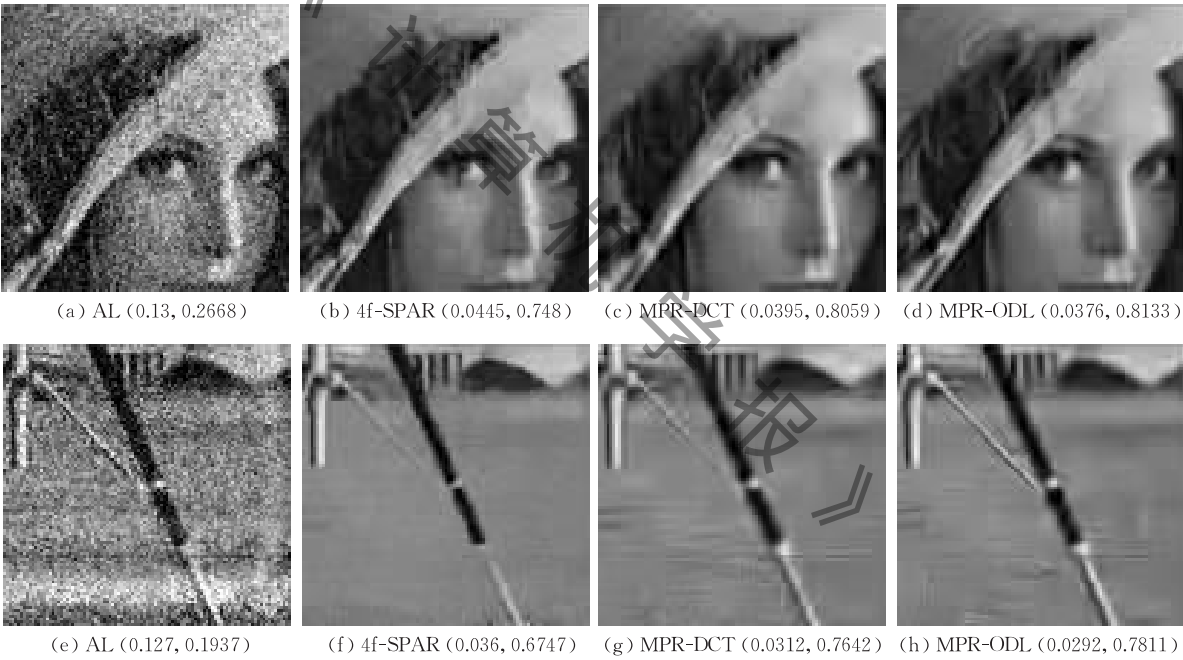


图 10 4 种算法对幅值和相位的重建结果($RMSE, SSIM$)($K=5$, 泊松噪声 $R_{\text{noise}}=10\%$)

4.3 算法时间复杂度

为测试算法的时间复杂度,本文将图像 Hill 作为幅值,图像 Peppers 作为相位构造复图像,在像面个数 $K=10$ 、不含噪声以及 $K=10$ 、含有强度 $\sigma=0.05$ 的高斯白噪声时,分别用上述 4 种算法对该复图像进行一次仿真实验,实验平台为 Intel Core i5-4590 四核 CPU,主频 3.3 GHz,内存 4 GB,Matlab 2013a.表 4 给出了 4 种算法的运行时间.由表 4 可知,AL 算法耗时最短,这是因为该算法对复图像整体施加先验信息并进行重建,每次迭代过程中只需进行简单的矩阵操作即可得到最优解,算法复杂度

较低,耗时较少;基于正交字典的 MPR-DCT 算法和 MPR-ODL 算法需要对图像进行稀疏表示,求解稀疏系数,算法复杂度较高,用时较长;为提高正交字典的性能,MPR-ODL 算法在迭代过程中需要更新字典,因此算法复杂度最高,耗时最长.

表 4 算法运行时间				(单位:s)
算法	AL	4f-SPAR	MPR-DCT	MPR-ODL
不含噪	317	319	735	938
含噪	321	322	751	945

为了更具体地测试本文算法的时间复杂度,表 5 给出了本文算法在第 $(iter \times 4/5) + (iter/50) \times i$

($i \in 1, 2, \dots, 10$) 次迭代时关键步骤平均耗时情况. 由表 5 可以看出, MPR-ODL 算法主要在正交字典学习这一步耗时较长.

表 5 MPR-ODL 算法关键步骤耗时 (单位: s)

步骤	更新幅值	更新相位	正交字典学习	更新中间变量
时间	1.2	0.032	25	1.4

4.4 算法收敛性

为了测试本文算法的收敛性, 图 11 给出了像面个数 $K=10$ 、观测值中含有强度 $\sigma=0.05$ 的高斯噪声时, MPR-ODL 算法的重建图像 Lena 和 Cameraman 对应的 $RMSE$ 值随迭代次数的变化曲线图. 该算法在前 250 次迭代中未利用正则项, 仅通过求解基于多个强度观测值的数据保真项对待重建图像进行粗略估计; 从第 251 次开始, 利用图像在正交字典下的稀疏性进行多像面相位恢复. 由图 11 可以看出, 从第 251 次迭代开始, $RMSE$ 值下降明显, 说明本文使用的稀疏正则项能有效提升重建图像的质量, 从而验证了本文算法的优越性.

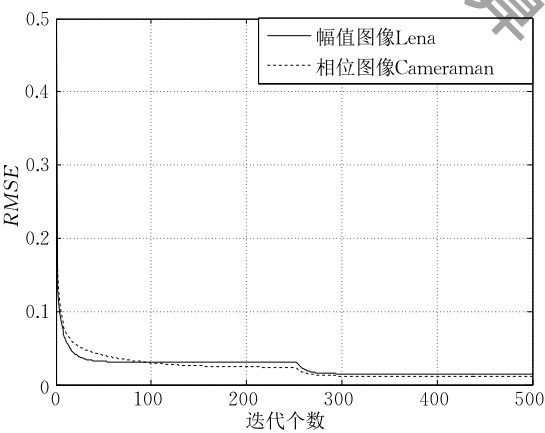


图 11 重建图像的 $RMSE$ 值随迭代次数变化曲线图

5 结 论

本文针对多像面相位恢复问题, 利用图像在字典表示下的稀疏性, 提出了基于正交字典学习的多像面相位恢复算法. 该算法结合图像的幅值和相位在正交字典下的稀疏正则项构造多像面相位恢复最优化问题, 并利用分裂 Bregman 方法和 MM 算法有效地求解了该优化问题. 该算法能够仅利用强度观测值重建图像并学习正交字典. 实验结果证明, 本文提出的 MPR-ODL 算法能够从多个强度观测值中有效地恢复原始图像, 并且在像面个数为 2 时依然能

够获得较高的图像重建质量. 由 $RMSE$ 和 $SSIM$ 两个指标可知, 与 AL 算法、4f-SPAR 等算法相比, 本文算法无论是从客观评价还是从主观视觉评价都存在着明显的优势. 另外, 本文算法在观测强度值含有高斯白噪声和泊松噪声时均能恢复出较为清晰的图像, 并保留部分细节, 对噪声具有鲁棒性.

参 考 文 献

[1] Shechtman Y, Eldar Y C, Cohen O, et al. Phase retrieval with application to optical imaging: A contemporary overview. IEEE Signal Processing Magazine, 2015, 32(3): 87-109

[2] Yang Z, Zhang C, Xie L. Robust compressive phase retrieval via L1 minimization with application to image reconstruction //Proceedings of the International Conference on Image Processing (ICIP). Melbourne, Australia, 2013: 1-5

[3] Ivanov V Y, Sivokon V P, Vorontsov M A. Phase retrieval from a set of intensity measurements: Theory and experiment. Journal of the Optical of America A, 1992, 9(9): 1515-1524

[4] Pedrini G, Osten W, Zhang Y. Wave-front reconstruction from a sequence of interferograms recorded at different planes. Optics Letters, 2005, 30(8): 833-835

[5] Gerchberg R W, Saxton W O. A practical algorithm for the determination of phase from image and diffraction plane pictures. Optik, 1972, 35(2): 237-246

[6] Migulin A, Katkovnik V, Astola J. Wave field reconstruction from multiple plane intensity-only data: Augmented Lagrangian algorithm. Journal of the Optical Society of America A, 2011, 28(6): 993-1002

[7] Almoro P, Pedrini G, Osten W. Complete wavefront reconstruction using sequential intensity measurements of a volume speckle field. Applied Optics, 2007, 45(34): 8596-8605

[8] Almoro P, Maallo A M, Hanson S. Fast-convergent algorithm for speckle-based phase retrieval and a design for dynamic wavefront sensing. Applied Optics, 2009, 48(8): 1485-1493

[9] Li Fang, Bi Yong, Kong Xin-Xin, et al. An improved algorithm for multi-plane reconstruction with Gerchberg-Saxton phase retrieval algorithm. Chinese Journal of Lasers, 2013, 40(10): 227-233

[10] Hagemann J, Robisch A L, Luke D R, et al. Reconstruction of wave front and object for inline holography from a set of detection planes. Optics Express, 2014, 22(10): 11 552-11 569

[11] Migulin A, Katkovnik V, Astola J. Optimal phase retrieval from multiple observations with Gaussian noise: Augmented Lagrangian algorithm for phase objects//Proceedings of the SPIE, 2011, 8082: 80820L-80820L-11

- [12] Katkovnik V, Astola J. High-accuracy wave field reconstruction; Decoupled inverse imaging with sparse modeling of phase and amplitude. *Journal of the Optical Society of America A*, 2012, 29(1): 44-54
- [13] Migukin A, Katkovnik V, Astola J. Advanced multi-plane phase retrieval using graphic processing unit; Augmented Lagrangian technique with sparse regularization. *Spie Photonics Europe*, 2012, 8429(6): 4779-4787
- [14] Dabov K, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 2007, 16(8): 2080-2095
- [15] Katkovnik V, Astola J. Phase retrieval via spatial light modulator phase modulation in 4f optical setup; Numerical inverse imaging with sparse regularization for phase and amplitude. *Journal of the Optical Society of America A*, 2012, 29(1): 105-116
- [16] Lian Qiu-Sheng, Shi Bao-Shun, Chen Shu-Zhen. Research advances on dictionary learning models, algorithms and applications. *Acta Automatica Sinica*, 2015, 41(2): 240-260 (in Chinese)
(练秋生, 石保顺, 陈书贞. 字典学习模型、算法及其应用研究进展. *自动化学报*, 2015, 41(2): 240-260)
- [17] Zhang Q, Li B. Discriminative K-SVD for dictionary learning in face recognition//*Proceedings of the Computer Vision and Pattern Recognition*. Arizona, USA, 2010: 2691-2698
- [18] Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 2006, 15(12): 3736-3745
- [19] Huang J, Guo L, Feng Q, et al. Sparsity-promoting orthogonal dictionary updating for image reconstruction from highly undersampled magnetic resonance data. *Physics in Medicine & Biology*, 2015, 60(14): 5359-5380
- [20] Bao C L, Cai J F, Ji H. Fast sparsity-based orthogonal dictionary learning for image restoration//*Proceedings of the IEEE International Conference on Computer Vision*. Iowa City, USA, 2013: 3384-3391
- [21] Katkovnik V, Migukin A, Astola J. Backward discrete wave field propagation modeling as an inverse problem; Toward perfect reconstruction of wave field distributions. *Applied Optics*, 2009, 48(48): 3407-3423
- [22] Katkovnik V, Astola J, Egiazarian K. Discrete diffraction transform for propagation, reconstruction, and design of wavefield distributions. *Applied Optics*, 2008, 47(19): 3481-3493
- [23] Fessler J A, Noll D C. Iterative image reconstruction in MRI with separate magnitude and phase regularization//*Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. Arlington, USA, 2004: 209-212
- [24] Zhao F, Noll D C, Nielsen J F, et al. Separate magnitude and phase regularization via compressed sensing. *IEEE Transactions on Medical Imaging*, 2012, 31(9): 1713-1723
- [25] Goldstein T, Osher S. The split bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences*, 2009, 2(2): 323-343
- [26] Candes E J, Strohmer T, Vershyninski V. PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming. *Communications on Pure and Applied Mathematics*, 2013, 66(8): 1241-1274
- [27] Ohlsson H, Yang A Y, Dong R, et al. Compressive phase retrieval from squared output measurements via semidefinite programming//*Proceedings of the System Identification*. Brussels, Belgium, 2012: 89-94
- [28] Lange K, Hunter D R, Yang I. Optimization transfer using surrogate objective function. *Journal of Computational and Graphical Statistics*, 2000, 9(1): 1-20
- [29] Qiu T, Babu P, Palomar D P. PRIME: Phase retrieval via majorization-minimization//*Proceedings of the Asilomar Conference on Signals, Systems and Computers*. Pacific Grove, USA, 2015: 1681-1685
- [30] Wang Z, Bovik A C, Sheikh H R, et al. Image assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612
- [31] Rodriguez J A, Xu R, Chen C-C, et al. Oversampling smoothness (OSS): An effective algorithm for phase retrieval of noisy diffraction intensities. *Journal of Applied Crystallography*, 2013, 46(2): 312-318



LIAN Qiu-Sheng, born in 1969, Ph.D., professor, Ph.D. supervisor. His current research interests include nonlinear compressed sensing, deep learning, phase retrieval, image processing.

GAO Li-Ping, born in 1991, M.S. Her research interests focus on phase retrieval.

SHI Bao-Shun, born in 1989, Ph.D., lecturer. His research interests include dictionary learning, phase retrieval, compressed sensing.

CHEN Shu-Zhen, born in 1968, associate professor. Her current research interests include compressed sensing, phase retrieval.

Background

This work is supported by the National Natural Science Foundation of China with the title: “The Study of Phase Retrieval and Image Reconstruction Based on Nonlinear Compressed Sensing and Adaptive Dictionary Learning” (No. 61471313), and the Natural Science Foundation of Hebei Province with the title “Adaptive Dictionary Learning and Its Application on Image Reconstruction” (No. F2014203076). The projects aim to explore the high order statistics and structural model of natural images, and study the high-quality image reconstruction and phase retrieval algorithms based on the synthesis and analysis sparse image representation models.

Phase retrieval refers to the problem of recovering a signal or image from intensity-only measurements of a

transform of that signal. That problem appears in crystallography, optical imaging, astronomical imaging, and other areas. The reconstructed images are often not unique due to the constant global phase change, spatial shift and conjugate inversion. Even though the phase retrieval problem is ill-posed, the ambiguity of the solution can be greatly reduced using multiple measurements which gain an observation redundancy and improve the reconstruction quality. This paper aims to design a multi-plane phase retrieval algorithm which incorporates the sparse regularization under orthogonal dictionary learning. Experiments show that this algorithm can achieve high quality reconstructed images and it is robust to noise.

《计算机学报》