

# 基于多通道联合变换相关器的旋转和缩放不变性目标识别方法

林超<sup>1\*</sup>, 韩艳丽<sup>1</sup>, 娄树理<sup>2</sup>, 刘珮<sup>1</sup>, 张文龙<sup>3</sup>, 杨自康<sup>3</sup>

<sup>1</sup>海军航空大学航空作战勤务学院, 山东 烟台 264000

<sup>2</sup>烟台大学光电信息科学与工程学院, 山东 烟台 264000

<sup>3</sup>中国人民解放军 92485 部队, 辽宁 大连 116041

**摘要** 为充分利用光学相关识别系统的空间-频谱带宽,提高光学相关识别系统的速度和精度,提出一种基于多通道联合变换相关器和综合鉴别函数的旋转和缩放不变性目标识别方法,并将峰值位置变化标准差作为识别判据。首先,分析以局域峰值杂波均值为相位优选约束条件时存在的问题,提出新的相位优选约束条件——峰值位置变化值,用于选择一定旋转和缩放范围内目标的公共相位,将公共优选相位用于多通道联合变换相关器并结合综合鉴别函数实现了旋转和缩放不变性识别;然后,分析了当目标缩小或综合鉴别函数集成训练图像数量增加时系统的识别容限;最后,分析了场景图像中背景发生变化时所提方法的适用性。研究表明,在设定的图像尺寸和背景复杂度条件下,所提方法能够在 9 通道并行处理的前提下对目标缩小 60%、综合鉴别函数集成 9 幅图像的旋转和缩放目标进行有效识别,且对像素数变化不超过 50%的背景具有较好的适应性,提高了目标识别的速度和精度,对多通道光学相关器的实用化具有重要意义。

**关键词** 信息处理; 光学模式识别; 多通道联合变换相关器; 综合鉴别函数; 旋转和缩放不变性识别

**中图分类号** O438; TP751

**文献标志码** A

**DOI:** 10.3788/CJL202249.1309001

## 1 引言

大数据和智能化时代的到来对信息处理的实时性要求越来越高,高性能计算(HPC)技术的发展显著提升了信息系统的计算能力,然而也带来了巨大的能耗<sup>[1]</sup>。以光信号为载体的光互联技术和光计算技术是高性能计算的发展方向之一,得益于硅基光电子技术<sup>[2]</sup>和光子神经网络技术<sup>[3]</sup>的发展,光子芯片的功能性和能耗比得到显著提高。

光电信息处理系统具有天然的二维并行处理能力以及大容量信息互连和传输能力,可以实现对图像的超快并行处理,尤其适用于模式识别等需要处理海量图像数据的应用场合。随着光电子技术的发展,学界提出了两类光学模式识别方法,分别是基于光电相关器的模式识别方法<sup>[4]</sup>和基于光学神经网络的模式识别方法<sup>[5-7]</sup>。光学神经网络兼具高速处理、低功耗和识别精度高的优势,但是需要对大量图像数据进行训练。传统数字相关器需要使用移位寄存器对输入图像进行串行加载,而光电混合相关器(HOC)基于空间光调制器和透镜的二维傅里叶变换性质实现了海量图像数据的并行处理,具有

处理速度快等优势,但是存在集成度低和识别精度不高等问题<sup>[8]</sup>。虽然 HOC 在实现复杂数学算法和识别图像精细特征方面没有优势,但是 HOC 可以作为神经网络系统的协处理器,用于对海量图像进行预先筛选<sup>[9]</sup>。因此,研究进一步提高 HOC 系统处理速度和精度的方法具有重要意义。根据存储介质的不同,HOC 可分为光学全息相关器和数字全息相关器两类,其中,基于体全息的相关识别技术已趋于成熟,能够实现典型场景下  $10^4$  量级图像数量的存储和  $10^3$  量级图像数量的并行相关识别<sup>[10]</sup>。基于数字全息的相关识别技术发展缓慢,其识别速度和精度与体全息识别技术有一定差距,主要是因为电荷耦合器件(CCD)相比于全息晶体不具备对海量图像的复用能力,导致基于数字全息的相关识别技术较难实现多图像的并行记录和相关输出。

基于联合变换相关器(JTC)的识别方法是一种典型的 HOC 识别方法,光路结构简单且无需匹配滤波器的精确对准,适用于对实时性要求较高的场合<sup>[11]</sup>。然而,基于数字全息的 JTC 系统并行处理能力不强且识别精度不高,存在零级项较强、相关峰能量分散、空间频谱带宽利用率较低、对旋转和缩放目标与相似目

收稿日期: 2021-09-26; 修回日期: 2021-11-08; 录用日期: 2021-12-14

基金项目: 国家自然科学基金(62005318)

通信作者: \*vestigelinchao@163.com

标的鉴别率较低等问题<sup>[12-15]</sup>。虽然很多改进的 JTC 识别方法被提出<sup>[16-18]</sup>,但是大多数方法是基于预处理或后处理的在线的非线性图像处理算法,在一定程度上削弱了 JTC 实时性强的优势。

提高 JTC 在空域和频域的带宽利用率是提升其并行处理能力的关键。研究人员将多组场景-参考图像对加载到空间光调制器(SLM)上,利用达曼光栅将均匀相干平面波分束为具有相同强度且互不重叠的二维光束阵列并分别照明多组场景-参考图像对,其功率谱被写入光寻址 SLM,相关输出时各个通道互不干扰,实现了多通道 JTC(MC-JTC)<sup>[19]</sup>。但是,理论上任何图像的傅里叶谱带宽是无限的,因此,多条联合变换功率谱(JPS)在频谱面仍存在相互叠加和串扰。另一种方法是基于电寻址 SLM 的像素结构对光调制时产生的高阶衍射级,将 SLM 像素结构的填充因子作为调控参量,具有不同填充因子的参考图像和场景图像经过傅里叶变换,其零级衍射叠加生成 JPS;再用棱镜或光楔将另一幅参考图像的傅里叶谱平移并叠加到场景图像的一级衍射项;通过调节场景图像和参考图像的谱宽和 SLM 填充因子,实现零级 JPS 和一级 JPS 在空间上的分离,实现了 MC-JTC<sup>[20]</sup>。该方法需要精确调节输入面图像灰度分布,且需要调节场景图像和参考图像所处位置处像素的填充因子比例,实现难度较大。此外,研究人员利用菲涅耳波带片、SLM 和傅里叶透镜构成光学倍增器,将场景图像复制出多幅图像并写入光寻址 SLM,与其上加载的多幅参考图像输入 JTC,并通过调控菲涅耳波带片的参数来避免不同通道高低阶衍射项的叠加,实现了 MC-JTC<sup>[21]</sup>。然而,上述方法需要设计特殊的波带片,光路结构较为复杂,调校难度大。

MC-JTC 充分发挥了 HOC 并行处理的优势,但上述 MC-JTC 方法主要是对照明光束或输入图像进行调控,没有对其傅里叶谱进行调控。由于任意图像的傅里叶谱带宽是无限的,谱线之间必然会产生串扰和叠加,无法真正实现不同通道的完全独立处理。本课题组提出一种基于功率谱压缩-平移的 MC-JTC 方法,将基于迭代算法优化后的相位函数叠加到场景和参考图像上,实现了在频谱面几乎无串扰的 JPS 空分复用,提高了 JTC 的并行处理能力<sup>[22]</sup>。

前期研究对提出的 MC-JTC 识别方法进行了验证,初步证明了其可行性<sup>[22]</sup>,即能够从错误参考图像中识别出正确参考图像,然而,这一结论不具备实用价值。实际应用中,场景图像的目标姿态和参考图像是不完全一样的,即目标在场景中会发生“旋转和缩放”。当目标在场景图像中发生旋转和缩放等姿态变化时,为了使相关识别系统能够识别该旋转和缩放目标,即识别结果不随目标旋转和缩放状态的变化而变化,常用的方法分为两类:第一类是解析法,基于纯数学的坐标变换或分解来实现。例如,采用极坐标梅林变换(PMT)可实现缩放不变识别<sup>[23]</sup>,基于圆谐振分解的匹

配滤波器(CHF)可实现面内旋转不变识别<sup>[24]</sup>。解析法计算简单、无需迭代优化,且具有较好的旋转和缩放不变识别效果,但是坐标变换需要在线计算,一定程度上降低了 HOC 系统的实时性。第二类是合成法,将目标的具有不同旋转和缩放形态的参考图像集或对应的滤波器库作为待训练数据库,基于某一优化指标,通过解析或迭代运算,合成一幅参考图像或匹配滤波器,并将其称为综合鉴别函数(SDF)<sup>[25]</sup>,SDF 可以预先离线计算,适合对实时性要求较高的应用场合。

本文在现有的 MC-JTC 研究中引入 SDF,将不同旋转和缩放范围的目标训练图像集综合为若干幅 SDF 图像,并将 SDF 图像作为参考图像同时加载到多个通道中,当场景中出现参考 SDF 覆盖旋转和缩放范围的目标时,系统都可以正确识别出旋转和缩放目标,既可以充分利用多通道并行处理的优势,又可以对一定旋转和缩放范围内的目标进行正确识别,在不增加系统硬件复杂度的前提下,有效提高了 MC-JTC 的识别速度和识别精度。

## 2 原 理

### 2.1 联合功率谱的并行记录

图 1 为 9 通道 JTC 的实验装置示意图。将场景图像和  $N$  幅 SDF 图像同时加载到纯振幅空间光调制器(AO-SLM)上,用于实现多通道并行处理的优选相位模板加载到纯相位空间光调制器(PO-SLM)上,采用  $4f$  中继投影系统将 AO-SLM1 的后表面投影到 PO-SLM 的前表面,通过合理调节 2 个 SLM 前、后 4 个偏振片的透光方位,实现复振幅调制。将 PO-SLM 的后表面置于傅里叶透镜(FL)的前焦平面上,准直相干光经由傅里叶透镜进行二维傅里叶变换,其联合变换功率谱由置于傅里叶透镜后焦平面的 CCD1 记录,并输出到计算机进行显示。分别记录场景图像和每一幅参考图像单独的傅里叶谱强度分布,在计算机中对 JPS 进行消零级和条纹调整等图像处理,将处理后的 JPS 加载到空间光调制器 AO-SLM2 上,经准直相干光照明,通过傅里叶透镜进行傅里叶逆变换,由 CCD2 采集其相关输出并在计算机上显示。

理论上,场景图像和参考图像都具有无限大的频谱带宽,即其在频谱面上占据的空间范围是无限的,因此,要实现无串扰的多通道并行处理,必须限制图像傅里叶频谱的空间范围和空间位置,从而可以在输入 SLM 面积和 CCD 感光面积一定的条件下实现多通道并行处理<sup>[22]</sup>。采用 GS(Gerchberg-Saxton)迭代算法<sup>[26]</sup>来约束图像傅里叶谱的空间范围。对图像傅里叶谱的空间位置进行调控,既可以采用线性相位函数对傅里叶谱进行平移,也可以在相位优化过程中直接对频谱面的不同区域进行约束,研究表明,两者效果基本相同。相位迭代优化算法的流程如图 2 所示,其中 FT 表示傅里叶变换,IFT 表示傅里叶逆变换。

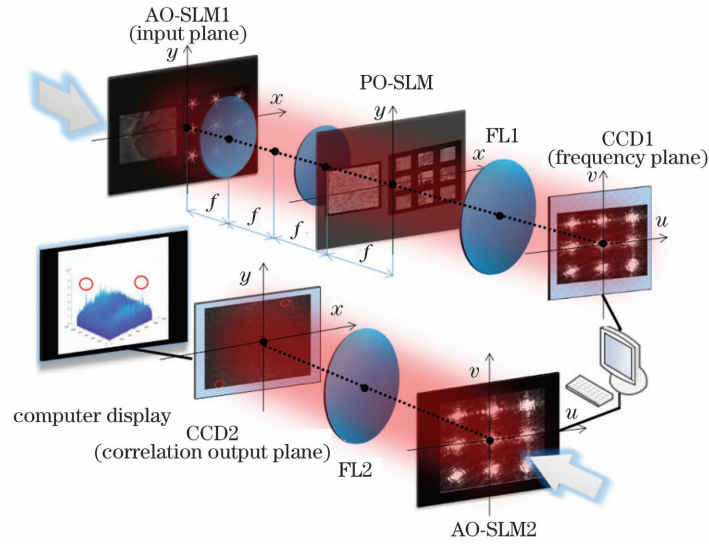


图 1 9 通道联合变换相关器三维光路图

Fig. 1 3D optical setup of nine-channel JTC

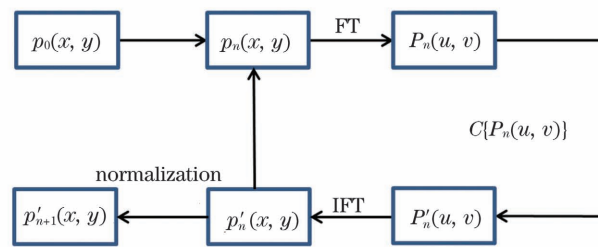


图 2 相位优化算法流程图

Fig. 2 Flow chart for phase optimization algorithm

首先, 建立一个初始随机相位函数  $p_0(x, y)$ , 其中  $(x, y)$  表示空域坐标, 通过傅里叶变换得到其傅里

叶谱, 其中  $(u, v)$  为 CCD 面频域坐标,  $C$  表示频谱面上的压缩限制操作, 如式(1)所示:

$$P'_n(u, v) = C\{P_n(u, v)\} = \begin{cases} P_n(u, v) & (u, v) \in \left(u_j - \frac{1}{2}u' : u_j + \frac{1}{2}u', v_j - \frac{1}{2}v' : v_j + \frac{1}{2}v'\right) \\ 0 & (u, v) \notin \left(u_j - \frac{1}{2}u' : u_j + \frac{1}{2}u', v_j - \frac{1}{2}v' : v_j + \frac{1}{2}v'\right) \end{cases}, \quad (1)$$

式中:  $P'_n$  为第  $n$  次迭代时压缩后的傅里叶谱;  $P_n$  为第  $n$  次迭代时相位函数的傅里叶谱;  $(u_j, v_j)$  表示第  $j$  条通道的傅里叶谱中心点的坐标;  $u', v'$  表示限制区域的坐标范围, 如图3所示。  $C$  表示将压缩限制区域以外

的傅里叶谱赋 0 值, 位于压缩限制区域以内的傅里叶谱保留原值。对压缩后的傅里叶谱进行傅里叶逆变换回到空域, 对傅里叶逆变换后的结果进行振幅归一化处理, 完成一次迭代过程。经过  $n$  次迭代后, 得到优化后的相位函数  $p'_{n+1}(x, y)$ 。

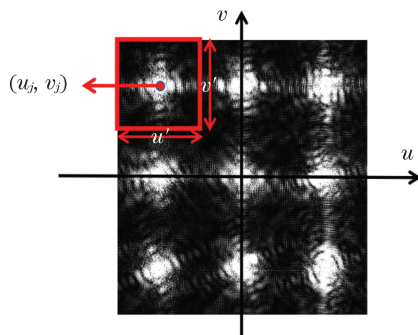


图 3 频谱面功率谱压缩示意图

Fig. 3 Captured image for compression of JPS in Fourier plane

选取 AO-SLM1 中心为原点, 设场景图像的振幅透过率分布函数为  $s(x-x_1, y-y_1)$ , 其中  $(x_1, y_1)$  为场景图像中心的坐标。  $N$  幅 SDF 图像的振幅透过率分布函数为  $\sum_{j=1}^N r_{\text{SDF}_j}(x+x'_j, y+y'_j)$ , 其中  $(x'_j, y'_j)$  为第  $j$  幅 SDF 图像中心的坐标。第  $j$  幅 SDF 图像的振幅透过率分布函数为  $r_{\text{SDF}_j} = \sum_{k_j=1}^K a_{k_j} r_{k_j}(x+x'_j, y+y'_j)$ , 其中  $r_{k_j}$  表示第  $k$  幅旋转和缩放参考图像,  $a_{k_j}$  表示第  $k$  幅旋转和缩放参考图像的权重,  $K$  为单幅 SDF 图像中集成的旋转和缩放参考图像数量。需要说明的是, SDF 图像的



数学表达式由所选用的 SDF 计算和优化方法决定,本研究选用经典的等相关峰值 SDF<sup>[25]</sup>来计算,具体计算方法不再赘述。 $\sum_{j=1}^N p_{\text{SDF}_j}^s(x-x_1, y-y_1)$ 为采用 GS 迭代算法得到的叠加到场景图像上的优选相位函数,  $\sum_{j=1}^N$

$p_{\text{SDF}_j}^r(x+x'_j, y+y'_j)$ 表示采用 GS 迭代算法得到的分别叠加到  $N$  幅 SDF 图像上的优选相位函数,其中场景图像和  $N$  幅 SDF 图像加载到 AO-SLM1 上,优选相位加载到 PO-SLM 上。经傅里叶变换运算,在 CCD 光敏面得到的联合功率谱分布为

$$I_{\text{JPS}} = \left| \mathcal{F} \left\{ \sum_{j=1}^N [s(x-x_1, y-y_1) \cdot p_{\text{SDF}_j}^s(x-x_1, y-y_1)] + \sum_{j=1}^N [r_{\text{SDF}_j}(x+x'_j, y+y'_j) \cdot p_{\text{SDF}_j}^r(x+x'_j, y+y'_j)] \right\} \right|^2 = \left| \sum_{j=1}^N \mathcal{F} \{s(x, y) \cdot p_{\text{SDF}_j}^s(x, y)\} \cdot \exp[i2\pi(ux_1 + vy_1)] + \sum_{j=1}^N \mathcal{F} \{r_{\text{SDF}_j}(x, y) p_{\text{SDF}_j}^r(x, y)\} \cdot \exp[-i2\pi(ux'_j + vy'_j)] \right|^2, \quad (2)$$

式中: $\mathcal{F}$ 表示傅里叶变换运算。

为了进一步提高相关输出性能,引用消零级和条纹调整 JTC<sup>[27]</sup>来实现多通道并行处理,具体处理过程

不再赘述。令  $A = \mathcal{F} \{s(x, y) p_{\text{SDF}_j}^s(x, y)\}$ ,  $B = \mathcal{F} \{r_{\text{SDF}_j}(x, y) p_{\text{SDF}_j}^r(x, y)\}$ ,则消零级和条纹调整后的第  $j$  条通道的联合变换功率谱(FA-JPS)可表示为

$$I_{\text{FA-JPS}}^j = \frac{2|A| \cos\{\varphi_A - \varphi_B + 2\pi[u(x_1 + x'_j) + v(y_1 + y'_j)]\}}{|B|}, \quad (3)$$

式中: $\varphi_A, \varphi_B$ 分别表示  $A$  和  $B$  的相位。

### 2.2 相关识别输出

经过迭代优化后,图像傅里叶谱的能量主要集中在约束区域内(本研究的约束区域面积占比为 67%),不同通道 JPS 相互叠加的能量很小,相当于在相关输出上叠

加一个噪声项,研究表明该噪声项对相关输出的影响较小。因此,主要分析第  $j$  条通道的相关输出特性。

对条纹调整的联合变换功率谱进行傅里叶逆变换,令  $c = 2\pi[u(x_1 + x'_j) + v(y_1 + y'_j)]$ ,可得到最终的相关输出可表示为

$$G(x, y) = \mathcal{F}^{-1} \left\{ \frac{2|A|}{|B|} \cos\{\varphi_A - \varphi_B + 2\pi[u(x_1 + x'_j) + v(y_1 + y'_j)]\} \right\} = \mathcal{F}^{-1} \left( \frac{|A|}{|B|} \right) * \mathcal{F}^{-1} \{ \exp[i(\varphi_A - \varphi_B)] \cdot \exp(ic) + \exp[-i(\varphi_A - \varphi_B)] \cdot \exp(-ic) \} = \mathcal{F}^{-1} \left( \frac{|A|}{|B|} \right) * \{ \mathcal{F}^{-1} \{ \exp[i(\varphi_A - \varphi_B)] \} * \mathcal{F}^{-1} [\exp(ic)] + \mathcal{F}^{-1} \{ \exp[-i(\varphi_A - \varphi_B)] \} * \mathcal{F}^{-1} [\exp(-ic)] \} = \mathcal{F}^{-1} \left\{ \frac{|A|}{|B|} \cdot \exp[i(\varphi_A - \varphi_B)] \right\} * \delta(x - c_x, y - c_y) + \mathcal{F}^{-1} \left\{ \frac{|A|}{|B|} \cdot \exp[-i(\varphi_A - \varphi_B)] \right\} * \delta(x + c_x, y + c_y) = \mathcal{F}^{-1} \left\{ \frac{|\mathcal{F}[s(x, y) p_{\text{SDF}_j}^s(x, y)]|}{|\mathcal{F}[r_{\text{SDF}_j}(x, y) p_{\text{SDF}_j}^r(x, y)]|} \cdot \exp[i(\varphi_A - \varphi_B)] \right\} * \delta(x - c_x, y - c_y) + \mathcal{F}^{-1} \left\{ \frac{|\mathcal{F}[s(x, y) p_{\text{SDF}_j}^s(x, y)]|}{|\mathcal{F}[r_{\text{SDF}_j}(x, y) p_{\text{SDF}_j}^r(x, y)]|} \cdot \exp[-i(\varphi_A - \varphi_B)] \right\} * \delta(x + c_x, y + c_y), \quad (4)$$

式中: $\mathcal{F}^{-1}$ 表示傅里叶逆变换;\*表示卷积运算; $c_x = x_1 + x'_j, c_y = y_1 + y'_j$ 。相关输出峰值强度(CPI)表示为  $\kappa_{\text{CPI}} = |G(x, y)|_{\text{max}}$ 。

在经典 JTC 系统中,当场景图像和参考图像相同时,CPI 为脉冲函数,表现为很强的相关峰。引入 SDF 后,当场景图像中出现某个 SDF 图像旋转和缩放范围内的目标时,基于等相关峰值 SDF 的相关输出为  $\mathcal{F}^{-1} \{ \mathcal{F}[r_{k_j}(x, y)] / \mathcal{F}[r_{\text{SDF}_j}(x, y)] \} = c_k, c_k$ 对任意  $k \in [1, K]$ 均为相等的常数,研究表明 SDF 的引入会

降低相关峰的强度,增大旁瓣强度,导致对于相关峰的鉴别度降低<sup>[28]</sup>。而对于所提出的 MC-JTC 系统,由式(4)可知,当场景图像中存在与某幅 SDF 图像  $r_{\text{SDF}_j}$  匹配的目标  $r_{k_j}$ ,即  $s(x, y) = r_{k_j}$  时,一方面,由于优选相位函数  $p_{\text{SDF}_j}^s(x, y) \neq p_{\text{SDF}_j}^r(x, y)$ ,因此  $A \neq B$ ,导致 CPI 数值下降且杂波进一步增强;另一方面,SDF 的引入会导致 CPI 下降和杂波增强,二者共同作用下相关峰的鉴别度会显著降低。前期研究表明,对未引入 SDF 的 MC-JTC 系统必须进行相位优选并用峰值位

置变化标准差 (SPCP;  $D_{SPCP}$ ) 作为相关识别的判据<sup>[22]</sup>。由于原相位优选条件是基于未旋转和缩放的参考图像和场景图像设计的, 引入 SDF 后, 由式 (4) 可知叠加到场景图像和 SDF 图像的优选相位是唯一的, 而场景图像中出现 SDF 旋转和缩放范围内任意参考图像均产生相关输出, 因此需要对包含多幅旋转和缩放参考图像的场景图像和 SDF 图像进行共同相位优选; 另外, 由于峰值的进一步下降和杂波的增强, 需要修改新的相位优选约束条件, 下面结合数值模拟和实验研究进行深入分析。

### 3 研究内容及方法

将一定缩放尺度下不同旋转角度的图像作为计算 SDF 的训练图像集, 基于经典的等相关峰值约束条件 (ECP) 和计算方法<sup>[25]</sup> 将间隔固定角度的旋转图像集成为一幅综合图像即 ECP-SDF, 如图 4 所示。在不同图示缩放尺度 (0.6、0.8 和 1.0) 条件下, 将不同初始角度、相同角度间隔训练图像集计算得到的多个 SDF 分别加载到 MC-JTC 的各条通道。当目标状态和任意通道中 SDF 所覆盖的旋转和缩放范围相匹配时, 相应通道就会产生相关输出, 从而实现对目标状态的准确识别。需要说明的是, 本研究不考虑通道数和功率谱压缩率对旋转和缩放不变识别性能的影响, 而是将通道数固定为 9, 更多通道下的研究过程与 9 通道相同, 只是具体参数结果会有不同。

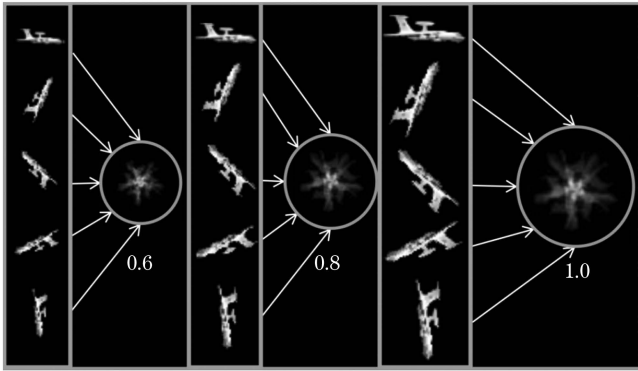


图 4 初始角度为  $0^\circ$ 、角度间隔为  $70^\circ$  的训练图像集在不同缩放尺度下集成 5 幅旋转图像生成的 SDF 图像  
Fig. 4 SDF images synthesized with 5 training images under different scaling sizes when initial angle is  $0^\circ$  and angle interval is  $70^\circ$

为了验证所提方法的有效性, 基于上述前提条件, 本文的研究内容为:

1) 原相位优选过程以场景图像和参考图像进行相关输出时的峰值杂波均值为约束条件<sup>[22]</sup>, 没有考虑到目标状态的旋转和缩放。而本文场景图像中可能包含任意旋转和缩放状态的目标, 且参考图像为 SDF 图像, 因此需要首先分析基于局域峰值杂波均值的相位优选条件是否适用并对其进行改进, 提出可行的相位优选方法。

2) 采用 SDF 实现旋转和缩放不变识别时, 对目标旋转和缩放存在一定的容限, 因此必须分析多通道 JTC 对于目标缩放比例和 SDF 集成图像数量的有效识别阈值。给出在本文设定条件下、以峰值位置变化标准差为识别判据时系统可识别的目标最小缩放比例和 SDF 可集成的最大图像数量。

3) 相位优选过程以场景图像和参考图像的相关输出特性为约束参量, 而在实时目标识别过程中, 不仅目标的姿态会发生旋转和缩放, 场景图像中的背景也会连续变化, 因此必须分析当背景变化时本文方法的有效性, 给出在本文背景条件下可正确识别的背景变化量阈值。

研究过程中采用理论分析、数值模拟和实验验证相结合的方法, 相关软硬件的参数如下: 输入 SLM 的分辨率为  $1024 \text{ pixel} \times 1024 \text{ pixel}$ , 像素尺寸为  $8 \mu\text{m}$ ; CCD 的分辨率为  $1024 \text{ pixel} \times 1024 \text{ pixel}$ , 像素尺寸为  $10 \mu\text{m}$ ; 相位优化算法的初始值由 MATLAB 软件中的均匀随机数函数 rand 生成; 为了使不同通道在频谱面叠加的能量尽可能小, 对于任意通道数, 傅里叶面的限制区域统一为  $(u, v) = (32, 32)$ , 考虑到 CCD 可以记录的动态范围和饱和效应, 该限制区域内基本能够保证不同通道串扰叠加的能量不被 CCD 记录, 即 CCD 只记录单个通道的功率谱, 基本没有通道之间的串扰; 为了保证有足够的能量用于相关输出, 对不同通道总数  $N$ , 选取傅里叶逆变换的空间范围为  $(1024/N, 1024/N)$ 。

### 4 结果及讨论

#### 4.1 L-PCM 作为相位优选约束条件的可行性分析

文献[22]基于局域峰值杂波均值 (L-PCM;  $\xi$ ) 优选输入相位函数, 并计算每条通道的 SPCP 数值, 将 SPCP (下文图表中均为 SPCP 的对数值) 最小值作为 MC-JTC 系统正确识别的判据。引入 SDF 后, 参考图像不再是某个角度的单一目标, 而是多角度的目标合成图像, 这对优选相位的计算提出了新的要求。前者仅需对单一目标满足 L-PCM 大于阈值, 后者要求当 SDF 训练图像集中任一状态的目标出现在场景图像中时, 使用一组固定的优选相位函数均能满足相应的 SPCP 数值最小, 即一组优选相位必须对所有训练图像同时满足 L-PCM 大于阈值, 将这一过程称为对一定旋转和缩放范围内目标的公共相位优选。由文献[22]可知, 随着 L-PCM 的增大, SPCP 逐渐减小, 当 L-PCM 增大到一定程度时, SPCP 趋于稳定值。因此, 选定一个阈值  $t$ , 采用优选相位进行相关输出时的 L-PCM 要满足

$$\xi_{s_r, \alpha + \langle n-1 \rangle, \beta}^{p_{SDF_j}} \geq t. \quad (5)$$

以图 4 所示的 SDF 集成 5 幅训练图像为例,  $p_{SDF_j}$  表示针对第  $j$  条通道中 SDF 所包含的 5 幅训练图像计算得到的公共优选相位,  $s_r$  表示包含训练图像  $r$  的

场景图像,其中  $\alpha$  表示训练图像的初始角度,  $\beta$  表示旋转角度间隔,  $n$  表示第  $n$  幅训练图像。设定  $\alpha = 0^\circ + 10^\circ \times (j - 1)$ ,  $\beta = 70^\circ$ , 在给定缩放尺度下, 式(5)对通道数  $j = 1, 2, 3, 4$  时  $n = 1, 2, \dots, 5$  的训练图像同时成立, 即使用 4 条通道可以覆盖给定缩放尺度条件下  $0^\circ \sim 30^\circ, 70^\circ \sim 100^\circ, 140^\circ \sim 170^\circ, 210^\circ \sim 240^\circ$  和  $280^\circ \sim 310^\circ$  范围内旋转角度的旋转和缩放目标。当然, 上述条件是本研究预先选定的, 实际应用中, 该旋转和缩放范围可以通过调节 SDF 训练图像集缩放尺度和旋转角度的范围灵活选取。

前期研究发现, 如果只对某一旋转角度下的训练图像进行相位优选, 会导致优选相位只在该角度下的 L-PCM 值很高, 而在其他角度下 L-PCM 值很低的现象, 因此基于式(5)进行公共相位优选是必须的。基于 GS 迭代算法进行相位优化, 由于每次迭代时相位的初始值是随机的, 需要通过多次运行, 选出同时对多角度训练图像优化效果都很好的公共优选相位。

采用式(5)作为约束条件, 设  $t = 10$ , 经过多次循环实验, 成功选出满足每幅训练图像的 L-PCM 数值都在 10 以上的优选相位。需要注意的是, 如果阈值  $t$  设置过高, 会导致无法选出满足条件的相位函数,  $t$  的合理数值与场景图像和 SDF 均有关。

得到一组公共优选相位后, 利用 9 通道 JTC 对其有效性进行验证。基于 ECP-SDF 生成 9 幅 SDF 图像作为 9 条通道的参考图像, 如图 5 所示。每幅 SDF 图像分别集成一定缩放尺度下  $n = 1, 2, \dots, 5$  的 5 幅训练图

像。其中第 1~4 条通道集成 0.8 缩放尺度下, 对应  $\alpha = 0^\circ, 10^\circ, 20^\circ, 30^\circ$  计算得到的 4 幅 SDF 图像; 第 5~8 条通道集成 1.0 缩放尺度下, 对应  $\alpha = 0^\circ, 10^\circ, 20^\circ, 30^\circ$  计算得到的 4 幅 SDF 图像; 第 9 条通道只集成 0.6 缩放尺度下,  $\alpha = 0^\circ$  计算得到的 1 幅 SDF 图像。

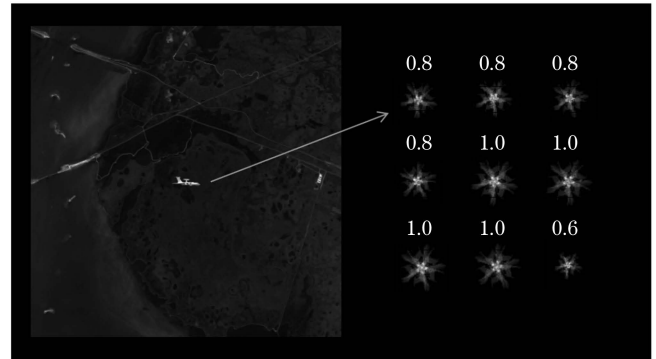


图 5 输入面联合图像示意图

Fig. 5 Schematic of joint images on the input plane

假设目标以 0.8 缩放尺度和  $\alpha = 0^\circ, \beta = 70^\circ$  的 5 个旋转角度中任一角度出现在场景图像中, 基于公共优选相位进行相关识别, 得到 SPCP 随目标在场景图像中旋转角度和通道数的变化关系, 如图 6 所示。图 6(a) 的横坐标为通道数, 纵坐标为旋转角度值, 三维柱状图的高度表示 SPCP 的对数值; 图 6(b) 为三维柱状图的二维展开图。在本文所有的 SPCP 与旋转角度的变化关系图中, 每个角度下从左到右的柱条分别对应第 1~9 条通道。

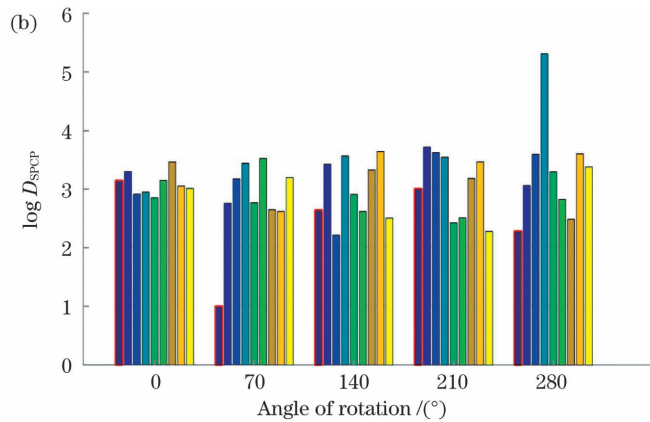
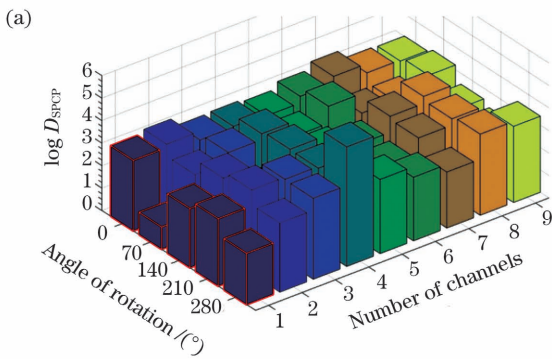


图 6 SPCP 与旋转角度和通道数的变化关系(基于 L-PCM 优选相位, 以红色边框标示第 1 通道)。(a) 三维柱状图; (b) 二维展开图

Fig. 6 Relationship among SPCP with rotation angles and number of channels (preferred phases are chosen based on L-PCM, the first channels are marked with red margins). (a) 3D histogram; (b) 2D unfold chart

当目标以 SDF 旋转和缩放范围内任一角度出现在场景图像中时, 在本文条件下, 第 1 条通道的 SPCP 数值应为所有通道中的最小值, 表明目标以该尺度和角度出现在场景中。然而, 观察图 6(a)、(b) 可以发现, 与第 3 节中直接用单一角度图像作为参考图像的情况相比, 采用 SDF 作为参考图像无法保证旋转和缩放范围内的 5 个角度下正确通道的 SPCP 数值均为最小值, 出现两条通道 SPCP 数值相近或某条通道 SPCP

数值低于正确通道 SPCP 数值等情况, 导致系统不能准确判断目标存在与否及其状态, 从而导致识别准确度显著下降。

基于上述分析和验证, 当参考图像为 SDF 图像时, 采用 L-PCM 作为约束条件进行相位优选, 以 SPCP 作为识别判据的准确性不高。在缩放尺度为 0.8,  $\alpha = 0^\circ, \beta = 70^\circ$  条件下, 对不加优选相位和加入公共优选相位后的相关输出强度分布进行分析, 结果如



图 7 所示,对比经典 JTC 系统的相关输出,将 SDF 作为参考图像时,叠加优选相位后的 CPI 显著降低,约为 0.04,文献[22]中叠加优选相位后的 CPI 能够保持在 0.2 以上。分析原因:首先,SDF 的引入导致 CPI 降低,对相关峰的鉴别率减小;其次,SDF 图像由多个旋转和缩放参考图像合成,其图像形态较为复杂,容易使场景图像中的某些背景和 SDF 图像产生相似性,导致输出的“伪相关峰”较强,尤其是在背景图像较为复杂的情况下;最后,使用优选相位进行功率谱压缩-平移,考虑到相位函数对光的干涉和衍射影响很大,基于文献[22]的结论,未优选的相位函数会产生很强的背景噪声,相比于对未旋转和缩放目标的识别,这一背景噪声在识别旋转和缩放目标时产生的影响更大,更容易使场景中的某些背景与旋转和缩放参考图像产生较强的伪相关峰,从而导致相关输出面出现“多峰值”现

象且错误目标产生的峰值可能超过正确目标产生的峰值,如图 7(b)所示。该现象与背景图像本身的复杂度有较大关系,将原背景中和参考目标相似的白色条状物体手动去除后生成一幅简单背景图像,用同样的优选相位进行相关输出,可显著降低出现“多峰值”现象的概率,如图 7(c)所示。计算 L-PCM 时,其 CPI 是基于相关输出面的强度最大值计算得到的,在“多峰值”情况下,由于该峰值的位置不是正确相关峰所处的像素位置,此时仍使用 L-PCM 作为约束条件来优选相位,会导致正确相关峰周围仍存在较强杂波,使得 SPCP 数值变大。综合上述分析,考虑到实际应用中背景的复杂性和不可控性,当 SDF 作为参考图像时,基于 L-PCM 优选相位会导致 SPCP 作为识别判据的有效性降低,需要对相位优选的约束条件进行进一步研究。

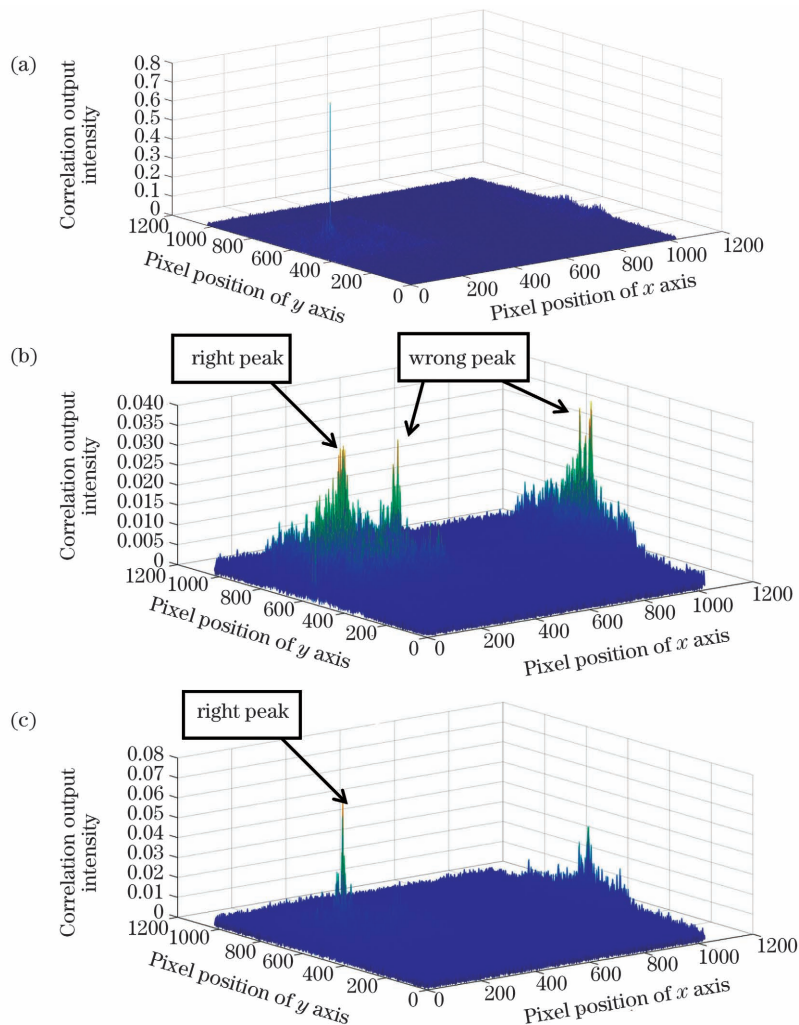


图 7 未加优选相位和加入公共优选相位后的相关输出。(a)未加优选相位的相关输出;(b)加入公共优选相位后的相关输出(原背景);(c)加入公共优选相位后的相关输出(简单背景)

Fig. 7 Correlation outputs without and with preferred phases. (a) Correlation output without preferred phase; (b) correlation output with public preferred phase (original background); (c) correlation output with public preferred phase (simple background)

#### 4.2 峰值位置变化值作为相位优选约束条件的可行性分析

要解决上述问题,理论上应该将图 7(a)中未加相位的 CPI 所在像素位置作为加入优选相位时计算

L-PCM 的约束条件之一,或者直接以相关输出强度最大值的像素位置变化值作为约束条件,实际研究发现,前者是相位优选的“强条件”,后者则为“弱条件”,使用“强条件”需要运行更多次才能选出满足条件的相位,

而后者所需运行次数显著减少。考虑到实际应用中场景图像的背景发生较大变化时需要重新在线计算优选相位,运行次数过多会影响目标识别的实时性,将优选的约束条件直接设置为相关输出强度最大值的位置变化值小于阈值  $t'$ 。由于正确图像所在通道的 CPI 在相关输出面所处像素位置具有稳定性,在相位优选时将不加优选相位时峰值的位置与加入优选相位后峰值位置的差值作为优选条件,即同时控制两像素点横坐标之差  $\Delta x$  的绝对值与纵坐标之差  $\Delta y$  的绝对值,且优选相位需同时满足

$$|\Delta x|_{s_{r_{a+c_{n-1}} > \beta}}^{p_{SDF_j}} \leq t' (n = 1, 2, \dots, 5), \quad (6)$$

$$|\Delta y|_{s_{r_{a+c_{n-1}} > \beta}}^{p_{SDF_j}} \leq t' (n = 1, 2, \dots, 5). \quad (7)$$

由于错误峰值的位置不具有稳定性,上述约束条

件可有效避免图 7(b)所示错误峰值对相位优选结果的影响。

设阈值  $t' = 1$ , 计算得到一组新的公共优选相位并用于相关输出,得到的结果如图 8(a)~(c)所示。图 8(a)、(b)为离散角度下拟合得到的曲面图的二维顶视图,3 个变量分别是通道数、旋转角度和 SPCP 大小(对数值),虚线标示出 SPCP 最小值所处的角度值和通道数。在曲面拟合时取若干个角度中间值进行验证,结果表明优选相位能够实现  $\pm 5^\circ$  范围内旋转和缩放目标图像的正确识别,且正确图像所在通道的 SPCP 数值相比其他通道来说很小,当场景图像中不存在目标时,SPCP 数值很大,说明本文方法具有很高的鉴别度。纵坐标的通道数为离散值,其拟合不具有实际意义。如图 8(a)所示,在 0.8 缩放尺度下,  $0^\circ \sim 30^\circ$

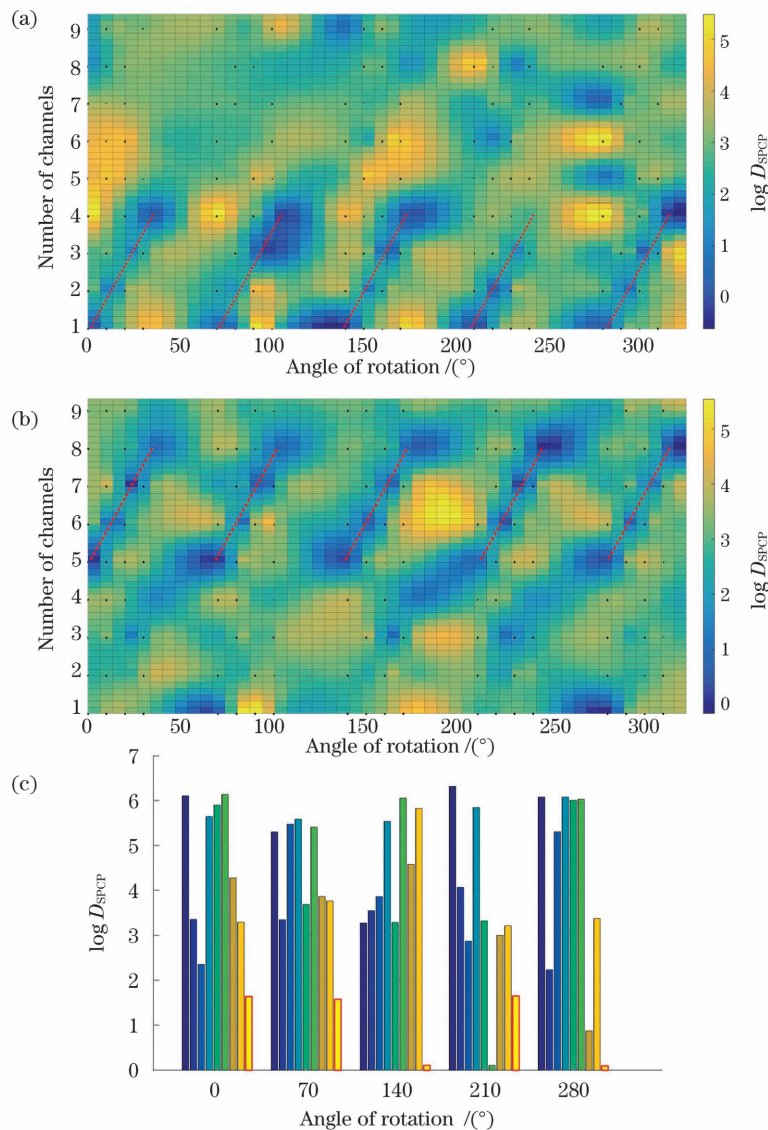


图 8 SPCP 与旋转角度和通道数的关系(基于峰值位置变化值选择相位,虚线和边框标示正确图像所在通道和角度范围)。

(a)第 1~4 条通道(缩放尺度为 0.8);(b)第 5~8 条通道(缩放尺度为 1.0);(c)第 9 条通道(二维展开图,缩放尺度为 0.6)

Fig. 8 Relationship among SPCP with rotation angles and number of channels (preferred phases are chosen based on variation of position of correlation peak, the right ones are marked with dotted lines and solid margins). (a) 1st~4th channels (scaling size is 0.8); (b) 5th~8th channels (scaling size is 1.0); (c) 9th channel (2D unfold chart, and scaling size is 0.6)



范围内通道 1~4 的 SPCP 值分别均为最小值,表明目标分别与通道 1~4 的 SDF 匹配,较为准确地识别出目标的旋转和缩放状态,相比于 L-PCM 作为约束条件,基于峰值位置变化值优选相位方法的识别精度显著提高,验证了新的相位优选条件的有效性。当然,在新的相位优选条件下得到满足条件的相位函数所需要的迭代次数更多、时间更长。需要说明的是,在计算 SPCP 时选了 3 组公共优选相位,整个系统的工作流程如图 9 所示。

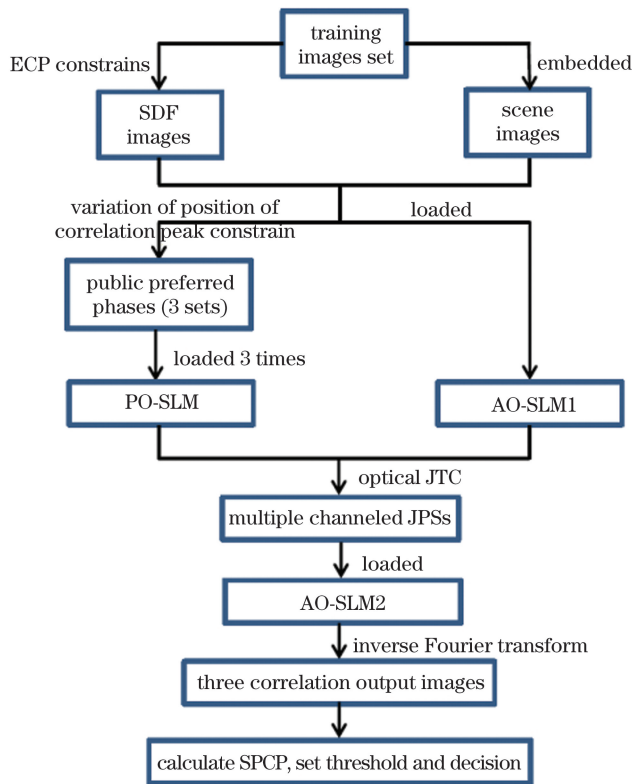


图 9 多通道 JTC 用于旋转和缩放不变识别的工作流程  
Fig. 9 Workflow of proposed distortion-invariant multiple channeled JTC

观察图 8(c)可以发现,目标旋转角度为  $210^\circ$  时第 9 条通道(正确通道)的 SPCP 数值不是最小的,其原因是 0.6 缩放尺度下场景中的目标尺寸和相应的 SDF 尺寸均较小,背景中部分像素区域或假目标与某条通道的 SDF 产生较强的相关输出,导致本文方法在复杂背景条件下对弱小目标的识别能力下降,出现虚警或漏警。后续研究发现当目标的尺寸缩小到 40% 时,目标特征微弱,导致无法计算出公共优选相位。下面分析本文方法对旋转和缩放的识别容限。

### 4.3 缩放比例和 SDF 集成图像数量的阈值分析

#### 4.3.1 目标缩放比例与 SPCP 的关系分析

当目标的尺寸缩小到 60% 时,识别精度下降,因此本节将探究目标的缩放比例变化(主要是目标缩小)对本文方法有效性的影响。研究发现目标越小,需要在相位优选时设定越高的阈值  $t'$  来保证能够计算出满足条件的优选相位。也就是说,目标尺寸变小会使

SPCP 数值变大,当 SPCP 增大到一定值时,其作为相关识别判据的有效性会降低。因此,在本文给定的场景图像像素 ( $450 \text{ pixel} \times 450 \text{ pixel}$ ) 和目标图像像素 ( $50 \text{ pixel} \times 20 \text{ pixel}$ , 缩放尺度为 1.0) 的条件下,研究其对目标尺度的识别极限具有重要意义。

将缩放尺度设置为 0.4,此时目标图像的像素仅为  $20 \text{ pixel} \times 8 \text{ pixel}$ ,其特征很微弱,在一定阈值条件下进行相位优选,得到 3 个旋转角度 ( $0^\circ$ 、 $70^\circ$ 、 $140^\circ$ ) 下目标图像对应的 CPI 位置变化值  $|\Delta x| + |\Delta y|$  随运行次数的变化曲线,并随机选取某个优选相位条件下的相关输出强度值,如图 10 所示。

在图 10(a)中,随着相位优选过程中运行次数的增加,不同旋转角度目标出现在场景图像时,其优选相位对应的峰值位置变化值波动很大,而且 3 条变化曲线没有交点,即不存在同时满足 3 个角度目标图像都能使峰值位置变化值低于某一阈值的优选相位。在本文给定的条件下,对于缩放尺度为 0.4 的目标,无法获得满足式(6)和式(7)的公共优选相位。图 10(b)、(c)对上述结论进行了解释:目标尺寸越小,加入优选相位后背景噪声越强(图 10(c)),多峰值现象越显著,而且错误相关峰峰值超过正确相关峰峰值会导致难以选出满足约束条件的优选相位(图 10(b)、(c)),和基于图 7 得出的结论一致。

将缩放尺度设置为 0.6,此时目标图像的像素为  $30 \text{ pixel} \times 12 \text{ pixel}$ ,如图 4 所示,其特征较为明显。图 11 所示为峰值位置变化值随运行次数的变化关系,可以看出,虽然峰值位置仍有较大波动,但是在位置变化值趋近 0 的位置,3 条曲线存在很多的交点,即能够选出满足一定阈值条件 ( $t' = 4$ ) 的优选相位,但是由于 CPI 峰值位置变化较大,当背景变得复杂时,特别是背景中出现和 SDF 图像相似的像素区域时,正确图像所在通道的 SPCP 并不是最小值(如图 8(c)中  $210^\circ$  条件下),本文方法的识别准确性降低,主要原因是图 7 显示的多峰值现象。需要说明的是,对于很多目标的识别方法,当目标尺寸变小而导致其特征变弱时,识别准确性都会降低。

综上,在 SDF 集成训练图像数量不变的前提下,随着目标尺寸的缩小,相位优选次数增加,约束条件的阈值变大,SPCP 作为识别判据的有效性降低。在本文方法采用的目标图像与场景图像像素一定的条件下,为了保证目标识别的精度,最小可将图像缩小至 60%,即当目标图像压行数为 12 pixel 时,本文方法可以有效识别目标。需要说明的是,当目标尺寸放大时,本文方法对于放大 2 倍的目标仍可以实现有效识别,这说明从识别精度的角度来看,本文方法的性能和已知的基于 JTC 和 SDF 的识别方法<sup>[28]</sup>基本一致。但是,相比已有研究中利用相关峰强度值作为判别参量,以 SPCP 为判别参量所得到的判别结果更稳定,正确通道的 SPCP 数值趋于 0,且无需设定阈值来区分正确

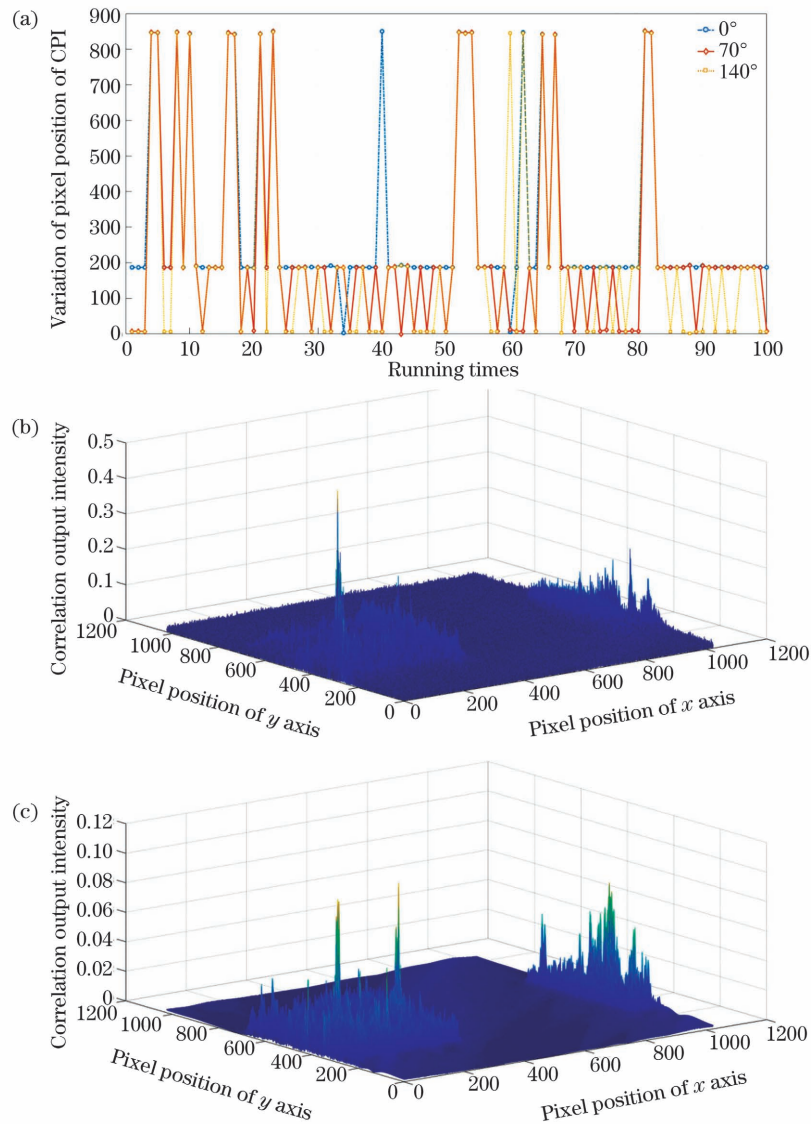


图 10 缩放尺度为 0.4 时的峰值位置变化值和相关输出强度值。(a)峰值位置变化值和运行次数的关系;(b)未加优选相位的相关输出;(c)加入优选相位后的相关输出

Fig. 10 Variation of position of correlation peak and correlation output at scaling size of 0.4. (a) Relationship between variation of pixel position of CPI with running times; (b) correlation output without preferred phase; (c) correlation output with public preferred phase

的相关峰值,这对于提升系统的整体判别精度具有重要意义。

#### 4.3.2 SDF 集成图像数量与 SPCP 的关系分析

上文的結果都是基于每幅 SDF 图像集成 5 幅训练图像。理论上,单幅 SDF 图像集成的训练图像数量增加能提高目标状态的识别精度,对于旋转和缩放范围一定的目标,需要使用的通道数量越少,越能充分发挥 MC-JTC 并行处理的优势。为了探究本文方法对 SDF 可以集成训练图像数量的极限,在目标缩放尺度一定的条件下,分别集成 7、9、11 个角度的训练图像生成 SDF,计算优选相位并进行相关输出,结果如图 12~14 所示。

由图 12 可知,当 SDF 集成 7 幅图像时,可以保证正确目标所在通道(第 2 条通道)的 SPCP 最小。当 SDF 集成 9 幅图像时,仍可以保证正确目标所在通道

(第 6 条通道)的 SPCP 最小,但其数值有所增大,如图 13(a)所示,同时由图 13(b)可知,优选相位所需的阈值较小,本文条件下的最小阈值  $t'=5$ 。

由图 14 可知,当 SDF 集成 11 幅图像时,得到优选相位所需的阈值明显增大,本文条件下的最小阈值  $t'=12$ 。该约束条件下得到的优选相位无法保证正确图像所在通道的 SPCP 值为最小值,其原因主要是计算 SDF 的像素数是一定的,当集成训练图像的数量增加时,重叠干扰和灰度饱和等效应使得目标特征减弱,与经典 JTC 中基于 SDF 的旋转和缩放不变识别类似,本文方法对于集成训练图像的数量存在上限。从实验结果总结出的 SDF 集成训练图像数量与相位优选的最小阈值的关系如表 1 所示。

综上,在本文方法采用的目标图像与场景图像像素一定的条件下,为了保证目标识别精度,基于 ECP-SDF

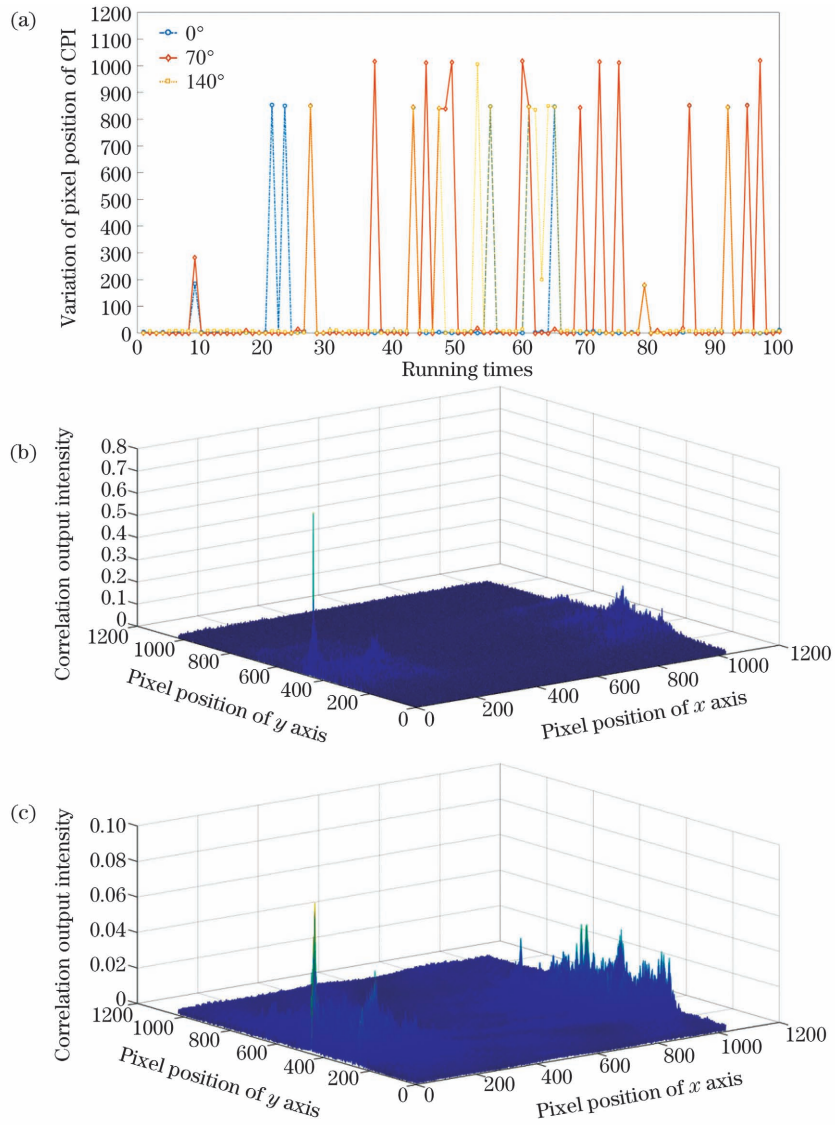


图 11 缩放尺度为 0.6 时的峰值位置变化值和相关输出强度值。(a)峰值位置变化值和运行次数的关系;(b)未加优选相位的相关输出;(c)加入优选相位后的相关输出

Fig. 11 Variation of position of correlation peak and correlation output at scaling size of 0.6. (a) Relationship between variation of pixel position of CPI and running times; (b) correlation output without preferred phase; (c) correlation output with public preferred phase

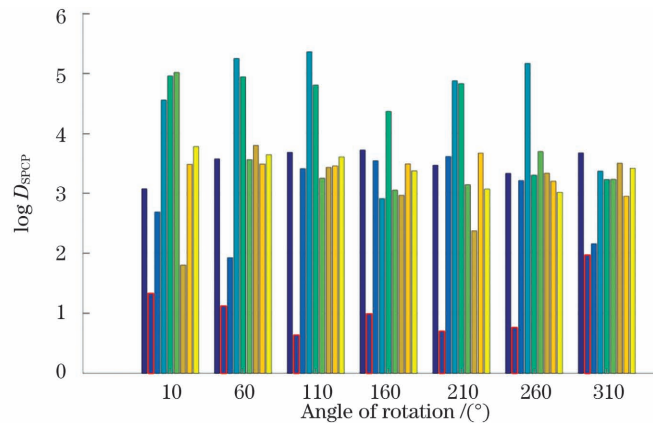


图 12 SPCP 与旋转角度和通道数的关系(SDF 集成 7 幅图像,缩放尺度为 1.0)

Fig. 12 Relationship among SPCP with rotation angles and number of channels (SDF synthesized with 7 training images and scaling size of 1.0)



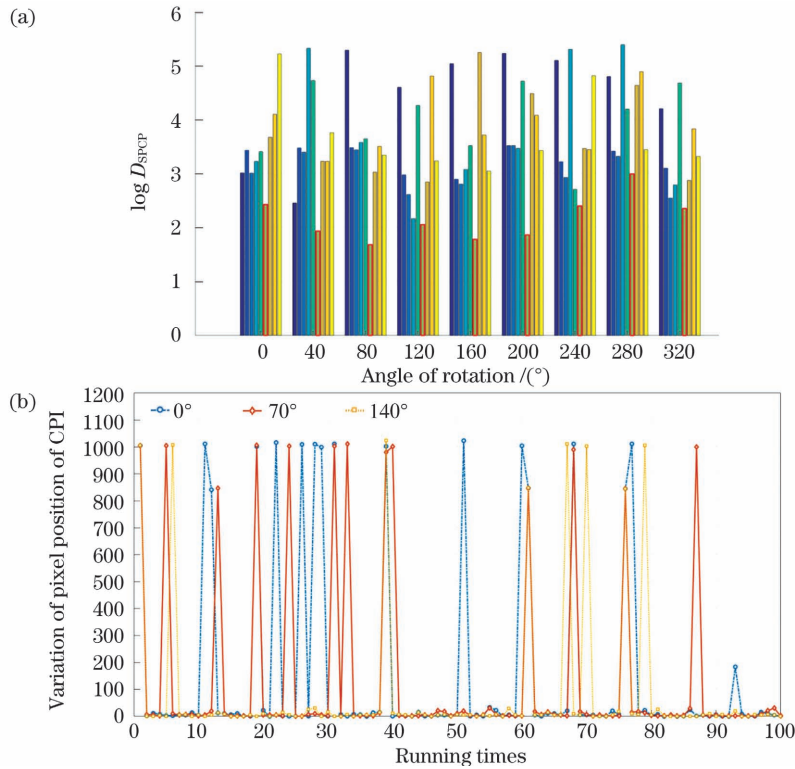


图 13 SPCP 关系图和峰值位置变化曲线(SDF 集成 9 幅图像, 缩放尺度为 1.0)。(a)SPCP 与旋转角度和通道数的关系;  
(b)峰值位置变化值和运行次数的关系

Fig. 13 SPCP diagram and variation of position of correlation peak (SDF synthesized with 9 training images and scaling size of 1.0). (a) Relationship among SPCP with rotation angles and number of channels; (b) relationship between variation of pixel position of CPI with running times

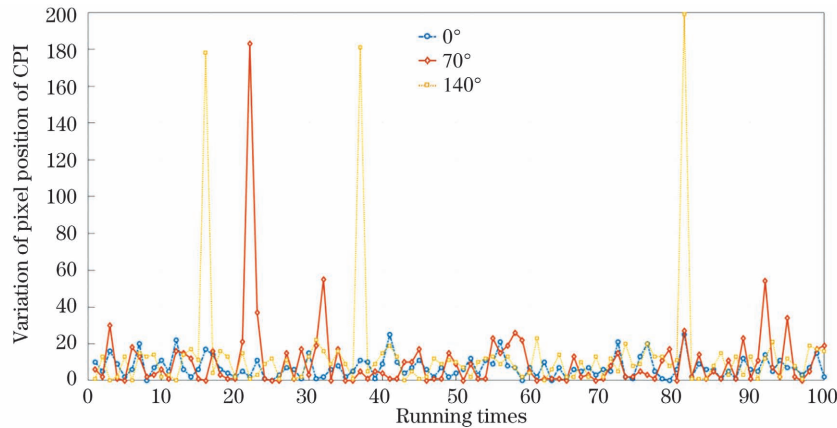


图 14 峰值位置随运行次数的变化关系(SDF 集成 11 幅图像, 缩放尺度为 1.0)

Fig. 14 Variation of position of correlation peak with running times (SDF synthesized with 11 training images and scaling size of 1.0)

算法计算 SDF 最多可以集成 9 幅训练图像。而已知基于 SDF 的单通道目标识别方法可以集成 25 幅或更多训练图像<sup>[29]</sup>, 本文方法在 SDF 集成训练图像数量方面的性能相比已有方法有所退化, 其主要原因是所

选取的公共相位优选算法。研究结果表明, 当训练图像数量增加时, 难以生成满足条件的公共优选相位。下一步将通过分析相位优选条件或者选择其他类型的 SDF 函数来提升可集成的图像数量阈值。

表 1 相位优选约束条件的最小阈值和 SDF 集成训练图像数量的关系

Table 1 Relationship between minimum threshold of phase optimization constraint and number of SDF synthesized images

Number of synthesized images	5	6	7	9	11
Minimum threshold	1	3	4	5	12

#### 4.4 背景变化对 SPCP 的影响分析

在目标识别系统的实际应用中, 由于系统所在平台以及目标本身会发生相对运动, 因此场景图像是实

时变化的。由于本文方法的公共相位优选约束条件针对的是特定场景图像下的相关输出峰值位置变化值, 下面分析当场景图像的背景发生变化时, 原优选相位

能否对目标进行有效识别,以及将背景替换为复杂背景时,基于复杂背景计算得到的优选相位是否有效。原背景是在一幅高分辨率遥感图像中截取的,通过变换截取像素区域,采用原背景下得到的优选相位进行相关输出,并分析其有效性。其中两帧变化后的背景如图 15(b)、(c)所示,图 15(b)相比于原背景变化较小(变化的像素数低于 50%),图 15(c)则变化较大(变化的像素数超过 50%)。图 15(d)所示为较为复杂的地面背景。

当其他条件保持不变,仅更换场景图像中的背景时,使用原背景下的优选相位对新背景中目标进行识

别的结果如图 16 和图 17 所示。可以发现,当背景变化较小时,正确图像所在通道(第 5 条通道)的 SPCP 仍为最小值,如图 16 所示。随着背景图像中发生变化的像素数超过 50%且背景复杂度变高,正确图像所在通道(第 5 条通道)的 SPCP 在某些角度下不是最小值,如图 17 中 0°和 210°,原因是新背景中出现了与目标相似的干扰,出现多峰值现象,导致 SPCP 作为相关识别判据的有效性降低,识别精度减小,产生虚警或漏警。图 18 所示为使用复杂背景图像作为场景图像重新计算优选相位并进行目标识别的结果,正确图像所在通道(第 5 条通道)的 SPCP 仍为最小值。

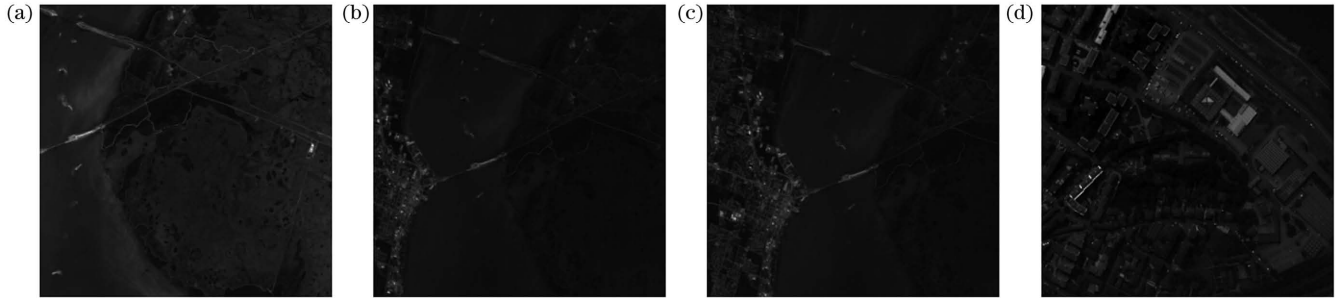


图 15 场景图像中背景发生变化示意图。(a)原背景;(b)变化小的背景;(c)变化大的背景;(d)复杂背景  
Fig. 15 Scene images with varied background. (a) Original background; (b) background with slight variation; (c) background with large variation; (d) complex background

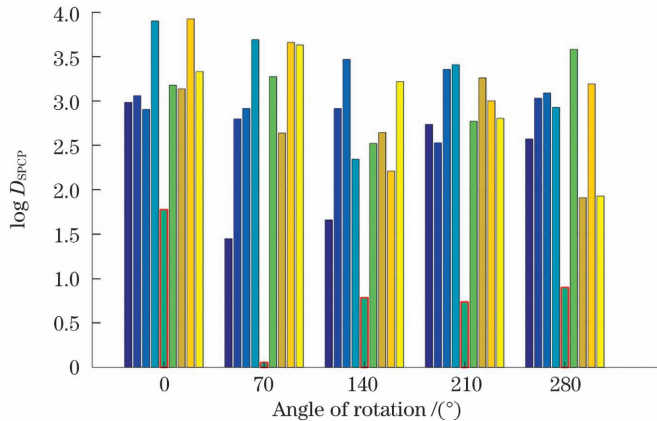


图 16 SPCP 与旋转角度和通道数的变化关系(背景变化较小,SDF 集成 5 幅图像,缩放尺度为 1.0)  
Fig. 16 Relationship among SPCP with rotation angles and number of channels (background with slight variation, SDF synthesized with 5 training images, and scaling size of 1.0)

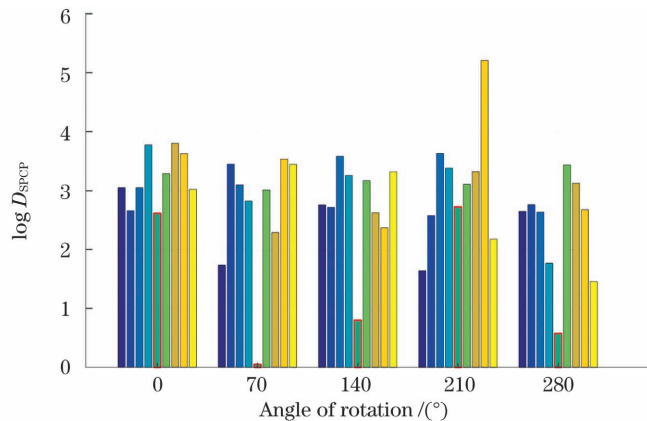


图 17 SPCP 与旋转角度和通道数的变化关系(背景变化较大,SDF 集成 5 幅图像,缩放尺度为 1.0)  
Fig. 17 Relationship among SPCP with rotation angles and number of channels (background with large variation, SDF synthesized with 5 training images, and scaling size of 1.0)

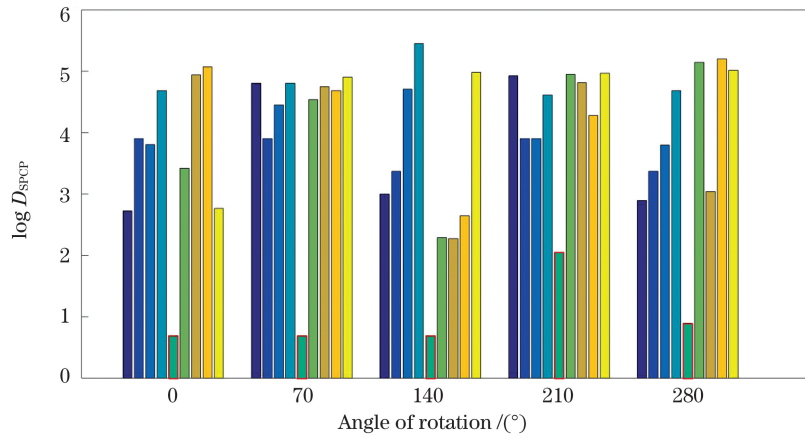


图 18 SPCP 与旋转角度和通道数的变化关系(复杂背景并重新计算优选相位, SDF 集成 5 幅图像, 缩放尺度为 1.0)

Fig. 18 Relationship among SPCP with rotation angles and number of channels (complex background with recalculated phases, SDF synthesized with 5 training images, and scaling size of 1.0)

综上,本文方法具有一定的背景变化适应性,当背景变为复杂地面背景时,仅需重新计算优选相位就能实现正确识别。但是,当背景连续变化且复杂度提高时,原优选相位会失效,识别精度显著下降,此时需要针对变化后的背景重新计算公共优选相位,在线计算优选相位会影响目标识别的实时性,这也是 3.2 节中选用“弱条件”进行相位优选的原因,保证能尽快得到特定背景下满足阈值条件的优选相位(20 次运行即可选出 3~5 个公共优选相位)。

## 5 结论

将基于功率谱压缩-平移的多通道联合变换相关识别方法应用于旋转和缩放不变性目标识别,在提升目标识别速度的前提下,进一步提升了所提方法对旋转和缩放目标的识别精度。分析了以 L-PCM 为约束条件进行相位优选时存在的问题,并提出基于峰值位置变化值进行公共相位优选的新约束条件,实现了对一定旋转和缩放范围内的目标识别。进一步分析了在给定图像尺寸等条件下,当目标缩小或 SDF 集成训练图像数量增加时系统的识别容限。分析了背景发生变化时所提方法的适用性,结果表明,所提方法在将目标识别速度提升 9 倍的条件下,实现了 0.8 和 1.0 缩放尺度下旋转角度范围包括  $0^\circ \sim 30^\circ$ 、 $70^\circ \sim 100^\circ$ 、 $140^\circ \sim 170^\circ$ 、 $210^\circ \sim 240^\circ$  和  $280^\circ \sim 310^\circ$  的目标状态的准确识别,缩放识别极限能够达到 0.6, SDF 集成图像数量的极限能够达到 9 幅。相比于单通道 JTC 需要多次刷新加载不同旋转和缩放范围的 SDF 图像来覆盖目标较大的旋转和缩放可能性,所提方法在维持单通道 JTC 旋转和缩放不变识别性能的前提下,实现了多通道并行处理,提高了识别的速度;而相比于已有多通道 JTC 系统,所设计的系统在简化光学系统结构的基础上,实现了旋转和缩放不变性目标识别,具有潜在的应用价值。

## 参 考 文 献

- [1] 金钟, 陆忠华, 李会元, 等. 高性能计算之源起: 科学计算的应用现状及发展思考[J]. 中国科学院院刊, 2019, 34(6): 625-639.  
Jin Z, Lu Z H, Li H Y, et al. Origin of high performance computing: current status and developments of scientific computing applications [J]. Bulletin of Chinese Academy of Sciences, 2019, 34(6): 625-639.
- [2] 周治平, 许鹏飞, 董晓文. 硅基光电计算[J]. 中国激光, 2020, 47(6): 0600001.  
Zhou Z P, Xu P F, Dong X W. Computing on silicon photonic platform[J]. Chinese Journal of Lasers, 2020, 47(6): 0600001.
- [3] 陈宏伟, 于振明, 张天, 等. 光子神经网络发展与挑战[J]. 中国激光, 2020, 47(5): 0500004.  
Chen H W, Yu Z M, Zhang T, et al. Advances and challenges of optical neural networks[J]. Chinese Journal of Lasers, 2020, 47(5): 0500004.
- [4] Manzur T, Zeller J, Serati S. Optical correlator based target detection, recognition, classification, and tracking[J]. Applied Optics, 2012, 51(21): 4976-4983.
- [5] Gao L, Chai Y, Zibar D, et al. Deep learning in photonics: introduction [J]. Photonics Research, 2021, 9(8): DLP1-DLP3.
- [6] Pierangeli D, Marcucci G, Conti C. Photonic extreme learning machine by free-space optical propagation [J]. Photonics Research, 2021, 9(8): 1446-1454.
- [7] Ryou A, Whitehead J, Zhelyeznyakov M, et al. Free-space optical neural network based on thermal atomic nonlinearity[J]. Photonics Research, 2021, 9(4): B128-B134.
- [8] 杨拓, 徐平, 周亮, 等. 折反式  $2f$  系统光学相关器的研究[J]. 光学学报, 2018, 38(1): 0107001.  
Yang T, Xu P, Zhou L, et al. Folding reflective  $2f$  system optical correlator [J]. Acta Optica Sinica, 2018, 38(1): 0107001.
- [9] Monjur M S, Tseng S, Tripathi R, et al. Hybrid optoelectronic correlator architecture for shift-invariant target recognition[J]. Journal of the Optical Society of America A, 2013, 31(1): 41-47.
- [10] 孟凡伟. 基于近化学计量比  $Mg:Fe:LiNbO_3$  晶体的体全息相关识别技术研究[D]. 哈尔滨: 哈尔滨工业大学, 2010.  
Meng F W. Study of volume holographic correlation recognition technology based on near stoichiometric  $Mg:Fe:LiNbO_3$  crystal [D]. Harbin: Harbin Institute of Technology, 2010.
- [11] Weaver C S, Goodman J W. A technique for optically convolving two functions[J]. Applied Optics, 1966, 5(7): 1248-1249.
- [12] Javidi B, Kuo C J. Joint transform image correlation using a binary spatial light modulator at the Fourier plane[J]. Applied



- Optics, 1988, 27(4): 663-665.
- [13] Tang Q, Zhang G S. Chirp encoded joint transform correlators with input scale search[J]. Optics Communications, 1994, 107(1/2): 23-27.
- [14] Lu G W, Zhang Z, Wu S D, et al. Implementation of a non-zero-order joint-transform correlator by use of phase-shifting techniques[J]. Applied Optics, 1997, 36(2): 470-483.
- [15] Zhong S, Jiang J X, Liu S T, et al. Binary joint transform correlator based on differential processing of the joint transform power spectrum[J]. Applied Optics, 1997, 36(8): 1776-1780.
- [16] Cheng C J, Tu H Y. Implementation of a nonzero-order joint transform correlator using interferometric technique[J]. Optical Review, 2002, 9(5): 193-196.
- [17] Zalevsky Z, Rubner A, García J, et al. Joint transform correlator with spatial code division multiplexing [J]. Applied Optics, 2006, 45(28): 7325-7333.
- [18] Zalevsky Z, García J, García-Martínez P. Interlacing multiplexing techniques for optical morphological correlation[J]. Optics Communications, 2006, 264(1): 45-50.
- [19] Feng J H, Chin G F, Wu M X, et al. Multiobject recognition in a multichannel joint-transform correlator [J]. Optics Letters, 1995, 20(1): 82-84.
- [20] Keryer G, de Bougrenet de la Tocnaye J L. A multichannel joint transform correlator[J]. Optics Communications, 1995, 118(1/2): 102-113.
- [21] Thapliya R, Koizumi H, Kodate K, et al. Parallel joint transform correlator applied to Devanagari script recognition[J]. Applied Optics, 1998, 37(23): 5408-5415.
- [22] 林超, 娄树理, 化斌斌, 等. 基于功率谱压缩-平移的多通道联合变换相关识别方法研究 [J]. 光学学报, 2021, 41(22): 2207001.
- Lin C, Lou S L, Hua B B, et al. Recognition method for optical multiple channeled joint transform correlation based on compression and translation of power spectrum[J]. Acta Optica Sinica, 2021, 41(22): 2207001.
- [23] Monjur M S, Tseng S, Tripathi R, et al. Incorporation of polar Mellin transform in a hybrid optoelectronic correlator for scale and rotation invariant target recognition [J]. Journal of the Optical Society of America A, 2014, 31(6): 1259-1272.
- [24] Hsu Y N, Arsenault H H. Optical pattern recognition using circular harmonic expansion[J]. Applied Optics, 1982, 21(22): 4016-4019.
- [25] Casasent D. Unified synthetic discriminant function computational formulation[J]. Applied Optics, 1984, 23(10): 1620-1627.
- [26] Gerchberg R, Saxton W. A practical algorithm for the determination of phase from image and diffraction plane pictures [J]. Optik, 1972, 35: 237-246.
- [27] Alam M S, Karim M A. Fringe-adjusted joint transform correlation[J]. Applied Optics, 1993, 32(23): 4344-4350.
- [28] Alam M S, Chen X W, Karim M A. Distortion-invariant fringe-adjusted joint transform correlation[J]. Applied Optics, 1997, 36(29): 7422-7427.
- [29] Napoléon T, Alfalou A. Pose invariant face recognition: 3D model from single photo[J]. Optics and Lasers in Engineering, 2017, 89: 150-161.

## Distortion-Invariant Target Recognition Based on Multichannel Joint Transform Correlator

Lin Chao<sup>1\*</sup>, Han Yanli<sup>1</sup>, Lou Shuli<sup>2</sup>, Liu Pei<sup>1</sup>, Zhang Wenlong<sup>3</sup>, Yang Zikang<sup>3</sup>

<sup>1</sup> School of Aviation of Operations and Support, Naval Aviation University, Yantai 264000, Shandong, China;

<sup>2</sup> School of Opto-Electronic Information Science and Technology, Yantai University, Yantai 264000, Shandong, China;

<sup>3</sup> Unit 92485 of PLA, Dalian 116041, Liaoning, China

### Abstract

**Objective** With the advent of the big data and intelligence eras, information systems require considerably enhanced performance and low energy costs. Optical computing may become the next-generation computing platform owing to its parallel processing capability and high bandwidth with low energy consumption. In pattern recognition applications, large amounts of image data must be rapidly processed. Two types of optical approaches have been investigated for pattern recognition: optical neural network, which comprises two subclasses including silicon photonic-based neural networks, and free-space-based optical network. The former has undergone considerable advancements recently owing to improved fabrication capability and novel network components based on optics such as microring resonators and Mach-Zehnder interferometers. The latter (e.g., diffractive neural networks) is also important, particularly for computational imaging-based applications. However, optical neural network-based pattern recognition approaches are immature owing to the implementation of nonlinear functions. Pattern recognition approaches founded on free-space-based optical networks are hybrid optoelectronic correlators, far more mature than optical neural network-based ones. The correlator can be codesigned with a neural network to serve as a coprocessor to prefilter some image features for ultrafast processing. However, in conventional optical correlators, both the spatial and spectral bandwidths of systems have not been efficiently used when performing the correlation operation. Hence, the inherent parallel processing capability of optics cannot be fully exploited.

**Methods** In our previous work, a multichannel joint transform correlation method is proposed based on the compression and translation of joint transform power spectrum to fully utilize spatial and spectral bandwidths and enhance the parallel processing efficiency and recognition accuracy of optical correlation systems. In the input plane of this scheme, the scene

image and  $N$  numbers of reference images are uploaded on different zones of the input spatial light modulator; then, the phase maps optimized using the iterative algorithm are superimposed onto the images. In the Fourier plane, interference between the Fourier spectra of scene images and those of every single reference image occurs in different zones of the Fourier plane. When the restriction parameter in the phase optimization algorithm is appropriated adjusted, no interference of the Fourier spectra of the reference images is observed. Consequently, the parallel processing of  $N$  channels is achieved without crosstalk. The relation between the localized peak clutter mean of the Fourier spectra of the preferred phase and the standard deviation of the correlation peak position is analyzed and used as a criterion for preferential preferred phase mask selection. Furthermore, the standard deviation of the correlation peak position is obtained for recognition tasks. In this study, we focus on distortion-invariant pattern recognition by integrating the multichannel joint transform correlator and the synthetic discriminant function. First, the feasibility of the local peak to clutter mean as a constraint for preferential phase selection is analyzed; results indicated that this factor is not appropriate when the synthetic discriminant function is used. Hence, a new phase selection criterion—known as the variation in the correlation peak position—is proposed to obtain the public preferred phase for targets with a specific distortion range. Furthermore, the selected phase is used in the multichannel joint transform correlator with the synthetic discriminant function to achieve distortion-invariant pattern recognition. Then, to determine the system performance in terms of the distortion level, the tolerance of our system on the scaling-down of the size of target and the increase in the number of training images for the synthetic discriminant function are analyzed. Finally, considering that the background may vary in real applications, we take successive video frames as varied input backgrounds and analyze the feasibility of our proposal.

**Results and Discussions** Results indicate that under the considered image file size and background complexity, the proposed method can achieve nine-channel parallel recognition (Fig. 8). For correct recognition, the minimum scaling down factor is 0.6 (Fig. 11). When the number of rotated training images is increased to 9 in the synthetic discriminant function, a correct recognition can be guaranteed (Fig. 13). The relation between the minimum threshold of phase the optimization constraint and the synthesized image numbers of SDF is obtained for calculating the preferred phases (Table 1). Furthermore, a correct recognition can be guaranteed when the values of half the pixels in the background have changed (Fig. 16).

**Conclusions** Herein, a novel distortion-invariant pattern recognition method based on multichannel joint transform correlator is proposed. The local peak to clutter mean is shown to be unsuitable, and we propose a new optimization criterion known as the variation in the correlation peak position, which is feasible in this proposal. We achieve nine-channel pattern recognition within 0.6–1.0 times of the scaling of the image size and rotation ranges of  $0^\circ$ – $30^\circ$ ,  $70^\circ$ – $100^\circ$ ,  $140^\circ$ – $170^\circ$ ,  $210^\circ$ – $240^\circ$ , and  $280^\circ$ – $310^\circ$ . The upper limit of the number of synthesized training images is analyzed, which is nine in this study. Moreover, the proposed method can maintain its performance when the background is varied within the values of half its pixel, indicating robustness to background changes. The recognition speed and accuracy on distortion of the system are considerably improved with our proposal, which will benefit the development of practical multichannel optical correlators.

**Key words** information processing; optical pattern recognition; multiple channeled joint transform correlator; synthetic discriminant function; distortion-invariant pattern recognition