

打印参数可变模式下3D打印批调度问题研究

王彬¹, 唐昊^{1,2†}, 戴飞¹, 谭琦¹

(1. 合肥工业大学 电气与自动化工程学院, 安徽 合肥 230009; 2. 工业自动化安徽省工程技术研究中心, 安徽 合肥 230009)

摘要: 研究打印参数可变模式下3D打印批调度问题, 旨在对打印任务随机到达的3D打印服务系统进行优化建模. 考虑到工作台尺寸的限制以及打印参数层高对打印质量的影响, 论文以各个任务队列长度作为系统联合状态, 以任务组合以及打印参数层高作为系统的联合控制变量, 以提高生产率、减少打印质量损失、节约电能为综合目标, 将优化问题描述为半马尔科夫决策过程(SMDP), 以便采用策略迭代算法、Q学习等算法求解系统最优调度策略. 仿真结果表明, 与层高参数固定模式以及先到先服务(FCFS)模式相比, 本文求解的批调度策略能够有效提高生产性能.

关键词: 3D打印; 批调度; 半马尔科夫决策过程; 策略迭代

引用格式: 王彬, 唐昊, 戴飞, 等. 打印参数可变模式下3D打印批调度问题研究. 控制理论与应用, 2021, 38(9): 1351 – 1360

DOI: 10.7641/CTA.2021.00679

Batch scheduling of 3D printing with variable printing parameter

WANG Bin¹, TANG Hao^{1,2†}, DAI Fei¹, TAN Qi¹

(1. School of Electrical Engineering and Automation, Hefei University of Technology, Hefei Anhui 230009, China;
2. Industrial Automation Engineering Technology Research Center of Anhui Province, Hefei Anhui 230009, China)

Abstract: Batch scheduling problem of 3D printing in the mode of variable printing parameter is studied. It aims to optimize and model on 3D printing service system with random arrival of printing tasks. First, the limitation of the size of the worktable and the influence of the layer of the printing parameters on the printing quality is considered. Then the lengths of all task queues are taken as the joint state of the system, and the combination of tasks and the layer of the printing parameters is taken as the joint control variable of the system. The aim is to increase productivity, reduce the loss of printing quality and save energy. Furthermore, the optimization problem is described as a semi-Markov decision process (SMDP), in order to solve the optimal scheduling strategy by using the strategy iteration algorithm and Q learning algorithm. The simulation results show that the proposed batch scheduling strategy can effectively improve the production performance compared with the fixed layer height model and the first come first service (FCFS) model.

Key words: 3D printing; batch scheduling; semi-Markov decision process; strategy iteration algorithm

Citation: WANG Bin, TANG Hao, DAI Fei, et al. Batch scheduling of 3D printing with variable printing parameter. *Control Theory & Applications*, 2021, 38(9): 1351 – 1360

1 引言

3D打印技术被称为“具有工业革命意义的制造技术”, 是制造领域中一项正在飞速发展的新兴制造技术. 采用3D打印技术进行生产的主要流程是: 应用3D模型设计软件设计出打印件的立体模型, 然后通过3D打印设备, 用液化、粉末化、丝化的打印材料逐层堆积“打印”出打印件^[1]. 随着网络技术与先进制造技术结合得越来越紧密, 一种通过物联网将3D打印设备接入制造网络, 接收客户设计的3D打印件模型进

行打印的网络化制造系统被提出^[2], 并在实际生产中开始广泛应用. 但现阶段3D打印产业处于发展阶段, 设备资源比较短缺, 因此, 针对这种打印任务随机的3D打印服务系统, 如何提高系统的处理率是急需解决的问题.

针对该问题, 已有学者为了减少设备预热时间、工作台调整时间和打印完成后设备冷却时间^[3], 研究了3D打印批次规划问题^[4]. 该问题可归类为工件尺寸不同的单机批调度问题, 对该问题的相关研究成果较多.

收稿日期: 2020-10-09; 录用日期: 2021-03-17.

†通信作者. E-mail: htang@hfut.edu.cn; Tel.: +86 13856089699.

本文责任编辑: 赵千川.

政府间国际科技创新合作重点专项(2017YFE0129700), 国家自然科学基金项目(61573126)资助.

Supported by the National Key R&D Program of China-Key Program for International Science and Technology Cooperation Projects of China (2017YFE0129700) and the National Natural Science Foundation of China (61573126).

例如, Ikura等人首先提出单机单工件簇的批处理优化控制问题^[5]. Lee等进一步研究了多工件簇的单机批处理优化问题^[6], 且已有学者提出采用微粒群算法和改进蚁群算法可对问题进行求解^[7-8]. 同时根据各批次加工时间与批中工件的关系, 该类批调度问题可分为3类, 第1类假定各批次加工时间与批中工件无关^[9]; 第2类假定各批次加工时间取决于批中最大工件^[10]; 第3类则假定各批次的加工时间为批中工件的加工时间之和^[11]. 由于3D打印过程中扫描填充时间与所选批中打印件的总体积成正比, 而层间调整所需时间取决于批中最高打印件的高度, 因此本文研究的问题属于第2类和第3类的结合. 由于打印任务随机到达, 因此本文研究的问题也属于在线批调度优化问题^[12-14].

上述研究主要是对任务进行调度以提高系统的处理率. 除了对任务进行调度外, 还可通过改变服务率提高系统性能. 例如在单站点传送带给料生产加工站(conveyor-serviced production station, CSPS)系统的优化控制问题中, 将服务率作为系统的一个控制变量, 求解系统的最优控制策略^[15]. 有学者进一步研究了顾客需求随机到达和服务率可动态实时变化的情况, 将系统的最优控制问题建成半马尔科夫决策过程(semi-Markov decision process, SMDP)模型, 采用策略迭代算法和Q学习算法求解了系统的最优控制策略, 提高了系统的生产率^[16]. 通过以上研究可知, 通过动态设定打印参数, 调整打印速度, 以提高系统的处理率^[14].

因此, 本文研究在打印任务随机到达情况下, 如何根据任务到达情况动态选择任务组合以及设置打印参数, 以提高系统的处理率. 由于对打印速度影响最大的是切片参数层高^[17-18], 因此本文选择打印参数层高作为控制变量. 虽然层高设置的得越大, 打印速度

越快, 但打印出的产品质量会降低^[19], 如何减少打印质量损失, 也是本文需要考虑的问题. 本文根据系统工作模式和决策过程的特点, 以各个任务队列长度作为系统联合状态, 以任务组合以及打印参数层高作为系统的联合控制变量, 以提高生产率、减少打印质量损失、节约电能为综合目标, 将该问题描述为SMDP模型, 采用策略迭代及Q学习算法对系统的最优策略进行求解.

2 系统组成及决策过程

如图1所示, 本文研究的对象是一种3D打印服务系统, 主要由存储打印任务的队列以及单台3D打印设备组成. 打印任务由不同种类的3D打印件模型组成, 各类打印任务按照独立的泊松过程随机到达, 并自动存储在相应的任务队列中, 任务队列容量有限. 3D打印设备以工作台作为其打印区间. 系统运行时, 根据任务队列中的任务数量, 选择进行打印的任务组合, 并设置打印参数层高, 然后对所选任务组合进行切片处理, 再由3D打印设备逐层累积打印, 直至打印完成. 因此, 系统需要根据任务队列中的任务数量对进行打印的任务组合和打印参数层高进行决策, 在此定义系统的决策时刻为打印完成时刻或打印任务到达时刻, 其典型决策过程如下: 在决策时刻, 系统根据当前任务队列中的任务数量, 选择由各类打印任务组成的任务组合并设置打印参数层高. 若所选组合不包含打印任务, 则不进行打印操作, 等待下一任务到达, 以进行下次任务组合选择. 当有任务到达时, 系统进入下一决策时刻; 若所选组合包含打印任务, 则从各个任务队列中选取相应个数的任务, 按照所选打印参数层高进行打印, 打印过程结束时, 系统进入下一决策时刻. 本文主要研究如何根据任务到达情况动态选择任务组合和层高, 以提高系统和系统的生产性能.

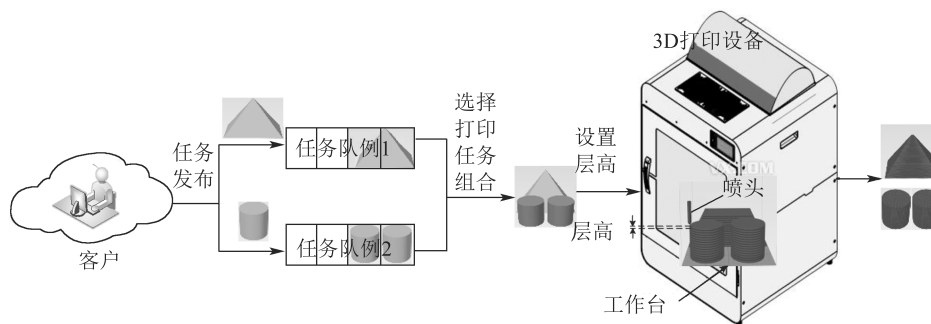


图1 系统的示意图

Fig. 1 Physical model of the system

本文的研究主要基于以下假设:

- 1) 打印任务服从不同参数的泊松过程随机到达并且相互独立, 打印件所需的打印材料相同;
- 2) 忽略打印过程中的热量散失, 喷头在预热和

打印时加热功率相同且恒定;

- 3) 忽略在打印同一层时喷头的空行时间;
- 4) 喷丝截面为矩形, 且忽略由于切片误差产生的打印材料消耗以及打印时间消耗;

5) 工作台每次调整的时间相同.

为了方便地进行数学描述以及进行数学建模, 本文定义了系统的参数和变量, 如表1所示. 本文研究的系统由 M 种打印任务、 M 个任务队列以及一台单喷头的3D打印设备组成. 这 M 种打印任务分别代表 M 种待打印的3D打印件模型, 并按照独立的泊松分布随机到达, 记第 m 种打印任务的到达率为 λ_m , 这 M 种打印任务分别存储在 M 个任务队列中, 第 m 个任务队列用于存储第 m 种打印任务, 其队列容量为 C_m . 3D打印件模型的种类根据底面尺寸 o_m 、体积 q_m 和高度 h_m 来区分. 本文将3D打印件模型的水平投影轮廓最小包络矩形作为底面尺寸, 记

第 m 种3D打印件模型的底面尺寸为 $o_m = n_{m,Length} \times n_{m,Width}$, 记其水平投影轮廓最小包络矩形的长和宽分别为 $n_{m,Length}$ 和 $n_{m,Width}$. 一般3D打印设备的工作台为矩形区域, 本文将3D打印设备的工作台作为打印设备的打印区域, 定义 N_{Length} 和 N_{Width} 为工作台的长和宽. 本文以任务 m 的队列长度 c_m 作为任务队列的状态, 其状态空间为

$$\Phi_m = \{0, 1, 2, \dots, C_m\}, m = 1, 2, \dots, M.$$

现以 M 个任务队列的状态作为系统的联合状态, 记为 $S_n = (c_1, \dots, c_m, \dots, c_M)$, 其状态空间为

$$\Phi = \Phi_1 \times \Phi_2 \times \dots \times \Phi_M.$$

表 1 符号变量表

Table 1 Variable table

符号	物理意义(单位)	符号	物理意义(单位)
M	打印任务种类	q_m	第 m 种打印件的体积(cm^3)
$N_{Length} \times N_{Width}$	工作台的长 \times 宽(cm^2)	h_m	第 m 种打印件的高度(mm)
λ_m	第 m 种打印任务的到达率(h^{-1})	$n_{m,Length} \times n_{m,Width}$	第 m 种打印件的矩形投影的长 \times 宽(cm^2)
t_{heat}	预热时间(h)	C_m	第 m 种任务队列容量
t_{move}	工作台每次调整所需时间(s)	c_m	第 m 种任务队列的长度
d	扫描路径宽(mm)	j_m	所选打印件组合中第 m 种打印件个数
a_{scan}	喷头的扫描速度($\text{mm} \cdot \text{min}^{-1}$)	t_{wait}	等待下一任务到达的时间(h)
t_{cool}	冷却时间(h)	$k_{1,m}$	第 m 种打印任务单位时间等待代价(h^{-1})
θ_{melt}	打印材料熔点($^{\circ}\text{C}$)	k_2	电能单价($1 \cdot (\text{kJ})^{-1}$)
ρ	打印材料的密度($\text{g} \cdot \text{cm}^{-3}$)	k_3	打印材料单价($1 \cdot \text{g}^{-1}$)
δ	打印材料的比热容($\text{J} \cdot (\text{g} \cdot ^{\circ}\text{C})^{-1}$)	ξ_2, ξ_1, ξ_0	磨损强度与层高 h 函数各项系数
θ_{envir}	环境温度($^{\circ}\text{C}$)	$\vartheta_{m,fix}$	打印第 m 种打印件的固定报酬
h	打印参数层高设定值(mm)	ϑ_m	打印第 m 种打印件的浮动报酬系数
H_{min}/H_{max}	层高最小/最大设定值(mm)		

定义在系统状态 S 下采取的行动为 $v_S = (J_S, h_S)$, 其中 J_S 为选择的任务组合,

$$J_S = [j_1 \ \dots \ j_m \ \dots \ j_M], J_S \in D_{J_S},$$

j_m 表示所选任务组合中第 m 种打印任务的个数, 所选任务组合中各类打印任务的数量不超过队列长度, 即 $j_m \leq c_m \leq C_m$. h_S 为设置的打印参数层高, 其取值范围为 $H_{min} \leq h_S \leq H_{max}$. D_{J_S} 为系统在状态 S 下的可选打印件组合的集合. 由于多打印任务同时打印时需要将打印任务布排到工作台上, 在已知工作台尺寸以及打印件水平投影的最小包络矩形条件下, 该问题属于矩形排样问题. 在指定空间布排矩形的方法有多种, 可根据问题的不同采用不同的方法. 对于复杂的排样问题可采用多启发式算法、粒子群算法等算法^[20], 对于简单的排样问题可采用Bottom-Left规则^[21]. 因此本文根据具体排样问题选择合适的算法求解得到系统在状态 S 下的可选任务组合的集合 D_{J_S} . 定义系统平稳策略为

$$v = \{(J_S, h_S) | J_S \in D_S, h_S \in [h_{min}, h_{max}], S \in \Phi\},$$

优化目标是找到一个最优策略使3D打印服务系统在长时间运行情况下代价最小.

系统状态 S 随时间随机跳变, 记 t 时刻的状态为 S_t , 第 n 次决策时刻记为 T_n , 在决策时刻 T_n 时系统状态为 S_{T_n} , 简记为 $S_n = S_{T_n}$. 在决策时刻 T_n , 系统所选任务组合 J_{S_n} 和层高设置 h_{S_n} 作为决策行动 v_{S_n} . 若 $\sum_{m=1}^M j_m = 0$, 则不进行打印操作, 等待下一任务到达, 记决策过程中等待下一任务到达的时间为 ϖ_n , ϖ_n 服从指数分布. 当有任务到达时, 系统进入下一决策时刻, 记为 T_{n+1} , 满足 $T_{n+1} = T_n + \varpi_n$, 若到达的是第 m 类打印任务, 则系统转移到下一状态 $S_{n+1} = (c_1, \dots, c'_m, \dots, c_M)$, 其中 $c'_m = \min\{c_m + 1, C_m\}$; 若 $\sum_{m=1}^M j_m \neq 0$, 则执行打印操作, 记打印过程所需的时间为 $\tau(v_{S_n})$, $\tau(v_{S_n})$ 由 $v_{S_n} = (J_{S_n}, h_{S_n})$ 确定. 在打印

过程中, 将第 m 种打印任务到达的个数记为 b_m , 则系统转移到下一状态

$$S_{n+1} = (c'_1, \dots, c'_m, \dots, c'_M),$$

其中 $c'_m = \min\{c_m - j_m + b_m, C_m\}$, $m = 1, 2, \dots, M$, 打印完成后进入下一决策时刻 T_{n+1} , $T_{n+1} = T_n + \tau(v_{S_n})$. 若决策时刻任务队列都为满时, 系统必须进行打印, 即 J_{S_n} 满足 $\sum_{m=1}^M j_m \neq 0$, 系统从任务队列中取出打印任务进行打印, 打印完成后进入下一决策时刻.

3 系统的半马尔科夫决策过程模型及优化算法

3.1 系统的半马尔科夫决策过程模型

由系统决策过程可知, 系统状态跳变只与当前状态有关, 与历史状态无关, 因此在一次样本轨道仿真决策过程中产生的状态序列 $\{S_1, S_2, \dots, S_n, \dots\}$ 是一个嵌入式马尔科夫链. 此外由于系统打印过程所需时间为确定值, 不服从指数分布, 因此系统决策过程是连续时间半马尔科夫过程^[22]. 除状态空间和行动集外, 系统的SMDP模型还包括逗留时间分布矩阵 $F^v(t)$ 、状态转移矩阵 P^v 以及性能函数矩阵 R^v , 下面将分别进行分析.

3.1.1 系统逗留时间分布矩阵

首先, 定义在策略 v 下系统逗留时间分布矩阵为

$$F^v(t) = [F_{S_n, S_{n+1}}(t, v_{S_n})]_{S_n, S_{n+1} \in \Phi}, \quad (1)$$

其中 $F_{S_n, S_{n+1}}(t, v_{S_n})$ 表示系统在状态 S_n 选择行动 v_{S_n} 转移到下一状态 S_{n+1} 的时间分布.

在决策时刻 T_n , 系统在状态 S_n 下选择行动 $v_{S_n} = (J_{S_n}, h_{S_n})$, 若 $\sum_{m=1}^M j_m = 0$, 则进行等待操作, 等待下一任务到达, 决策周期为等待下一任务到达的等待时间 $\varpi_n = t_{wait}$, 系统转移到下一状态 S_{n+1} , 此时状态逗留时间分布如下:

$$F_{S_n, S_{n+1}}(t, v_{S_n}) = 1 - e^{-\lambda t}, \quad \lambda = \sum_{m=1}^M \lambda_m, \quad (2)$$

λ 为 M 种任务到达率之和.

若 $\sum_{m=1}^M j_m \neq 0$, 则进行打印操作, 决策周期时间轴

如图2所示, 由于3D打印过程由预热、逐层扫描打印、设备冷却3个步骤组成, 因此打印过程所用的时间由3个部分组成: 预热时间 t_{heat} 、扫描填充时间 $t_{print}(v_{S_n})$ 、冷却时间 t_{cool} , 打印过程所需时间 $\tau(v_{S_n})$ 计算式如下:

$$\tau(v_{S_n}) = t_{heat} + t_{print}(v_{S_n}) + t_{cool}. \quad (3)$$

接下来文中给出预热时间 t_{heat} 、扫描填充时间 $t_{print}(v_{S_n})$ 、冷却时间 t_{cool} 和打印过程所需时间 $\tau(v_{S_n})$ 的具体计算过程.

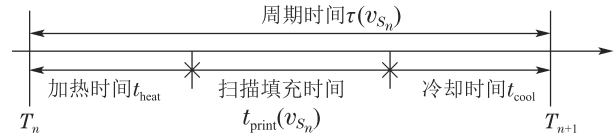


图2 打印操作时间轴

Fig. 2 Timeline of printing operation

3D打印机在打印前需要对工作台以及喷头等部件进行加热, 已有技术实现喷头在较短的时间内加热到设置温度, 且基本实现无超调^[23], 因此本文认为预热时间 t_{heat} 为固定值.

由于打印时喷头需要对每层截面扫描填充, 扫描填充完一层后, 工作台需要向下调整, 然后对下一层进行扫描填充, 因此扫描打印时间 $t_{print}(v_{S_n})$ 由喷头扫描填充时间 $t_{scan}(v_{S_n})$ 与工作台调整时间组成, 喷头扫描填充时间 $t_{scan}(v_{S_n})$ 与打印件的体积成正比, 工作台调整时间取决于工作台的移动次数. 多打印件打印时喷头扫描填充时间 $t_{scan}(v_{S_n})$ 等于各打印件的扫描填充时间之和, 工作台调整时间取决于最高打印件的高度. 喷头扫描填充时间 $t_{scan}(v_{S_n})$ 在忽略切片产生的填充体积误差时, 可根据扫描路径宽度 d 、参数层高 h_{S_n} 、喷头的扫描速度 a_{scan} 以及打印件的总体积求出, 计算式如下:

$$t_{scan}(v_{S_n}) = \frac{J_{S_n} Q^T}{dh_{S_n} a_{scan}}, \quad (4)$$

其中: $Q = [q_1 \ \dots \ q_m \ \dots \ q_M]$ 为 M 种打印件的体积向量.

记所选打印件组合中最高打印件的高度为 h_{max} , $h_{max} = \max\{h_m | j_m \neq 0 \text{ 且 } j_m \in J_{S_n}\}$, 则扫描打印时间 $t_{print}(v_{S_n})$ 计算公式如下:

$$t_{print}(v_{S_n}) = t_{scan}(v_{S_n}) + (\lceil \frac{h_{max}}{h_{S_n}} \rceil - 1)t_{move}, \quad (5)$$

上式中 $\lceil \cdot \rceil$ 为向上取整函数.

3D打印机完成逐层扫描填充后, 需要等待机器冷却后将打印件取出, 记这段时间为冷却时间 t_{cool} , 本文采用固定值.

综上所述, 在决策时刻 T_n , 系统在状态 S_n 下选择的行动 $v_{S_n} = (J_{S_n}, h_{S_n})$, 若 $\sum_{m=1}^M j_m \neq 0$, 则进行打印操作, 系统转移到下一状态. 在此过程中系统状态逗留时间分布为

$$F_{S_n, S_{n+1}}(t, v_{S_n}) = \begin{cases} 1, & t \geq \tau(v_{S_n}), \\ 0, & t < \tau(v_{S_n}). \end{cases} \quad (6)$$

3.1.2 状态转移矩阵

定义在策略 v 下系统的嵌入链转移矩阵为 $P^v = [P_{S_n, S_{n+1}}(v_{S_n})]_{S_n, S_{n+1} \in \Phi}$, 其中 $P_{S_n, S_{n+1}}(v_{S_n})$ 表示在状态 S_n 采取行动 v_{S_n} , 转移到下一状态 S_{n+1} 的概率. 执行打印操作时, 在 t 时间内, 任务队列 m 由状态 c_m 转

移到 c'_m 的概率为

$$p_{c_m, c'_m}(v_{S_n}, t) = \begin{cases} \frac{(\lambda_m t)^{c'_m - c_m + j_m}}{(c'_m - c_m + j_m)!} e^{-\lambda_m t}, & \text{条件1,} \\ 1, & \text{条件2,} \\ 1 - \sum_{k=0}^{C_m - c_m + j_m - 1} \frac{(\lambda_m t)^k}{k!} e^{-\lambda_m t}, & \text{条件3,} \end{cases} \quad (7)$$

其中: 条件1为 $0 \leq c'_m - c_m + j_m < C_m - c_m + j_m$, 条件2为 $c'_m = c_m = C_m, j_m = 0$, 条件3为 $c'_m = C_m, c_m \geq j_m > 0$.

对于其他状态转移, $p_{c_m, c'_m}(v_{S_n}, t) = 0$. 由于任务队列状态相互独立, 且不同任务到达过程也是相互独立的, 因此在状态 S_n 下采取行动 v_{S_n} 转移到下一状态 S_{n+1} 的概率为

$$P_{S_n, S_{n+1}}(v_{S_n}) = \int_0^\infty \prod_{m=1}^M p_{c_m, c'_m}(v_{S_n}, t) dF_{S_n, S_{n+1}}(t, v_{S_n}), \quad (8)$$

其中: $F_{S_n, S_{n+1}}(t, v_{S_n})$ 为系统状态逗留时间分布函数. 打印操作时, 将式(7)代入得到 $P_{S_n, S_{n+1}}(v_{S_n})$.

当执行等待操作时, 到达的为第 m 种任务的概率为 $\frac{\lambda_m}{\lambda}$, 系统下一状态为

$$S_{n+1} = (c'_1, \dots, c'_m, \dots, c'_M) = (c_1, \dots, c'_m, \dots, c_M),$$

其中 $c'_m = \min\{c_m + 1, C_m\}$. 若 $c_m \neq C_m$, 则 $S_{n+1} = (c_1, \dots, c_m + 1, \dots, c_M)$, 此时在状态 S_n 下采取行动 v_{S_n} 转移到下一状态的概率为

$$P_{S_n, S_{n+1}}(v_{S_n}) = \frac{\lambda_m}{\lambda}. \quad (9)$$

若 $c_m = C_m$, 则 $S_{n+1} = S_n$, 当存在任务队列为满时, 系统转移到相同状态的概率为

$$P_{S_n, S_n}(v_{S_n}) = \frac{\lambda_{\text{full}}}{\lambda}, \quad (10)$$

其中: λ_{full} 为状态 S_n 中所有为满队列对应任务的到达率之和, $\lambda_{\text{full}} = \sum_{m \in B_{\text{full}}} \lambda_m$, B_{full} 为状态 S_n 中所有为满队列对应的任务种类的集合.

3.1.3 性能函数矩阵

本文通过SMDP模型描述3D打印批调度问题, 其优化目标是找到一个最优策略使系统在无穷时段下的单位时间代价最小.

首先, 由于系统采取不同行动时会产生不同代价, 因此定义 $R(S_n, v_{S_n}, S_{n+1})$ 为系统在决策时刻 T_n 采取行动 v_{S_n} 并转移到下一决策时刻 T_{n+1} 过程中的单位时期望代价函数, 则性能函数矩阵为

$$R^v = [R(S_n, v_{S_n}, S_{n+1})]_{S_n, S_{n+1} \in \Phi}. \quad (11)$$

系统的优化目标函数为^[22]

$$\eta_\alpha^v(S_0) = E[\int_0^\infty \alpha e^{-\alpha t} R(S_n, v_{S_n}, S_{n+1}) dt | S_0], \quad (12)$$

其中: $\eta_\alpha^v(S_0)$ 表示系统从状态 S_0 开始在策略 v 下的系统无穷时段期望折扣代价, α 为折扣因子, 若 $\alpha = 0$, 则上式表示平均代价优化准则. 本文的优化目标就是在策略集中寻找一个最优策略 v^* , 使 $\eta_\alpha^v(S_0)$ 最小.

系统在执行打印操作或等待操作会产生多类代价, 例如任务队列中打印任务的等待代价, 打印材料损耗代价, 电能损耗代价等, 下面将具体分析 $R(S_n, v_{S_n}, S_{n+1})$ 的计算.

首先, 打印任务在任务队列等待时需要占用一定的存储空间, 同时任务等待时间会影响任务的交货时间, 因此定义第 m 种任务等待时单位时间等待代价为 $k_{1,m}$.

其次, 3D打印过程需消耗电能, 因此定义电能单价为 k_2 . 打印设备在打印过程中能量损耗主要用于喷头加热用于融化打印材料, 则在忽略热量散失及其他加热装置条件下, 喷头的加热功率 g 近似等于单位时间融化材料所需的热能, 与参数层高 h 成正比. 在参数层高为 h 时, 单位时间打印材料挤出量为 $dha_{\text{scan}}\rho$, 其中 d 为扫描路径宽度, a_{scan} 为喷头的扫描速度, ρ 为打印材料的密度. 则打印材料由环境温度 θ_{envir} 加热到打印材料的融化温度 θ_{melt} 所需的热量为 $dha_{\text{scan}}\rho\delta \times (\theta_{\text{melt}} - \theta_{\text{envir}})$, δ 为打印材料的比热容, 则选择行动 $v_{S_n} = (J_{S_n}, h_{S_n})$ 时, 打印设备的加热功率为

$$g(v_{S_n}) = dh_{S_n} a_{\text{scan}} \rho \delta (\theta_{\text{melt}} - \theta_{\text{envir}}). \quad (13)$$

另外, 打印过程中需要消耗打印材料, 因此定义材料单价为 k_3 . 进行打印行动时, 所需的打印材料 $l(v_{S_n})$ 与选择打印的打印件组合 J_{S_n} 的总积成正比, 则选择打印的打印件组合 J_{S_n} 时消耗的打印材料质量计算式如下:

$$l(v_{S_n}) = \rho J_{S_n} Q^T. \quad (14)$$

此外, 系统打印完产品会获得报酬, 由于层高参数 h 影响产品质量, 因此本文根据产品质量决定产品的打印报酬, 根据文献[19]的研究结果可知, 打印产品的磨损强度与层高参数 h 函数关系近似拟合为

$$S_{\text{wear}}(h) = \xi_2 h^2 + \xi_1 h + \xi_0, \quad (15)$$

式中: ξ_2, ξ_1, ξ_0 为各项系数, 在此定义打印报酬由固定报酬以及浮动报酬这两部分组成, 其中浮动报酬与打印质量成正比, 定义打印第 m 类任务的报酬 $r_m(h_{S_n})$ 与层高参数 h_{S_n} 关系为

$$r_m(h_{S_n}) = \vartheta_m S_{\text{wear}}(h_{S_n}) + \vartheta_{m, \text{fix}}, \quad (16)$$

其中: ϑ_m 为打印第 m 类任务的奖励报酬系数, $\vartheta_{m, \text{fix}}$ 为打印第 m 类任务获得的固定报酬. 则采取行动 v_{S_n}

$= (J_{S_n}, h_{S_n})$ 获得的报酬为

$$r(v_{S_n}) = \sum_{m=1}^M j_m r_m(h_{S_n}).$$

因此,系统执行打印操作时,单位时间期望代价为

$$R(S_n, v_{S_n}, S_{n+1}) = \sum_{m=1}^M \frac{(c_m - j_m + c'_m)}{2} k_{1,m} + k_2 g(v_{S_n}) \varphi_1(S_n, v_{S_n}, S_{n+1}) + k_3 l(v_{S_n}) \varphi_2(S_n, v_{S_n}, S_{n+1}) + r(v_{S_n}) \varphi_3(S_n, v_{S_n}, S_{n+1}), \quad (17)$$

其中 $\varphi_1(S_n, v_{S_n}, S_{n+1})$ 为电能代价在整个时间段上的折扣,计算式为

$$\varphi_1(S_n, v_{S_n}, S_{n+1}) = \frac{T_\alpha(t_{\text{heat}} + t_{\text{print}}(v_{S_n}))}{\int_0^\infty T_\alpha(t) dF_{S_n, S_{n+1}}(t, v_{S_n})}, \quad (18)$$

其中: $T_\alpha(t) = \int_0^t e^{-\alpha x} dx$, $\varphi_2(S_n, v_{S_n}, S_{n+1})$ 为材料消耗代价在整个时间段上的折扣,计算式为

$$\varphi_2(S_n, v_{S_n}, S_{n+1}) = \frac{e^{-\alpha(t_{\text{heat}} + t_{\text{print}}(v_{S_n}))}}{\int_0^\infty T_\alpha(t) dF_{S_n, S_{n+1}}(t, v_{S_n})}. \quad (19)$$

$\varphi_3(S_n, v_{S_n}, S_{n+1})$ 为打印立即报酬在整个时间段上的折扣,计算如下:

$$\varphi_3(S_n, v_{S_n}, S_{n+1}) = \frac{\int_0^\infty e^{-\alpha t} dF_{S_n, S_{n+1}}(t, v_{S_n})}{\int_0^\infty T_\alpha(t) dF_{S_n, S_{n+1}}(t, v_{S_n})}. \quad (20)$$

与打印操作不同,执行等待操作时只存在等待代价,此时系统单位时间期望代价为

$$R(S_n, v_{S_n}, S_{n+1}) = \sum_{m=1}^M c_m k_{1,m}. \quad (21)$$

3.2 优化算法

本文的优化目标就是找到一个最优策略 v^* ,使系统在折扣均准则或平均准则下的无穷时段的单位时间代价最小.在此本文给出策略迭代和Q学习这两种算法求解系统最优策略.

文献[22, 24–25]指出通过等价无穷小因子 A_α^v 和等价性能矩阵 f_α^v ,可将半Markov决策过程转换为连续时间Markov决策的过程模型 $X = \langle S_t, \Phi, D, A_\alpha^v, f_\alpha^v \rangle$ [18],然后采用策略迭代算法进行求解,其步骤主要分为策略评估和策略改进两部分,本文不再详述.

尽管策略迭代算法可以求出系统最优解,但是需要求解系统矩阵,并进行矩阵求逆,当系统参数例如到达率未知,或状态庞大造成维数灾使矩阵运算计算困难时,最优制策略无法求解或求解十分困难.因此,可用Q学习算法进行求解.

Q学习算法的主要步骤分为如下几个过程:首先采用矩形排样方法得到打印件组合行动集 D_J ,然后将

层高 h 取值范围用一个很小的常数 Δh 离散化,得到一个离散的紧致行动集 D_h ,通过组合得到系统行动集.在每次决策过程中,通过观测得到一个样本数据 $\langle S_n, v_{S_n}, S_{n+1}, \omega_n \rangle$, ω_n 为第 n 个决策时刻与第 $n+1$ 个决策时刻的时间间隔,则平均准则和折扣准则下统一差分计算式为

$$\Delta_n = R'(S_n, v_{S_n}, S_{n+1}) - T_\alpha(\omega_n) \cdot \eta_n + e^{-\alpha \omega_n} \min_{v \in D} Q_\alpha(S_{n+1}, v) - Q_\alpha(S_n, v_{S_n}), \quad (22)$$

其中: $Q_\alpha(\cdot, \cdot)$ 是在折扣因子 α 下的状态-行动对的值, $R'(S_n, v_{S_n}, S_{n+1})$ 表示在第 n 个决策时刻系统在状态 S_n 下,采取行动 v_{S_n} 转移到状态 S_{n+1} 过程中产生的累计折扣代价.具体代价计算如下:当进行等待操作时

$$R'(S_n, v_{S_n}, S_{n+1}) = T_\alpha(\omega_n) \sum_{m=1}^M c_m k_{1,m}. \quad (23)$$

当进行打印操作时

$$R'(S_n, v_{S_n}, S_{n+1}) = T_\alpha(\omega_n) \sum_{m=1}^M \frac{c_m - j_m + c'_m}{2} k_{1,m} + T_\alpha(t_{\text{heat}} + t_{\text{print}}(v_{S_n})) k_2 g(v_{S_n}) + e^{-\alpha(t_{\text{heat}} + t_{\text{print}}(v_{S_n}))} k_3 l(v_{S_n}) + r(v_{S_n}) e^{-\alpha \omega_n}. \quad (24)$$

此外, η_n 是平均代价的当前估计值,可根据样本轨道直接计算.则 $Q_\alpha(\cdot, \cdot)$ 更新式如下:

$$Q_\alpha(S_n, v_{S_n}) = Q_\alpha(S_n, v_{S_n}) + \gamma(S_n, v_{S_n}) \Delta_n, \quad (25)$$

其中 $\gamma(S_n, v_{S_n})$ 为学习步长.本文介绍的Q学习算法具体步骤已在文献[26]中进行了详细介绍,在此不再详述.

4 仿真结果

本文根据实验室中的3D打印设备的参数对系统的物理参数进行设置,具体参数设置如表2所示.同时设置了仿真参数如表3所示.根据仿真实验给出的3D打印机和打印件的尺寸可知,本文给出的排样问题属于简单排样问题,因此本文选择Bottom-Left规则对打印任务进行布排[4],按照此规则求解系统可行打印件组合集合为

$$D_J = \{(0, 0, 0), (0, 0, 1), (0, 0, 2), (0, 0, 3), (0, 1, 0), (0, 1, 1), (0, 2, 0), (1, 0, 0), (1, 0, 1)\}.$$

仿真实验首先介绍了策略迭代算法和Q学习算法在折扣和平均准则下的优化过程,图3为在平均和折扣准则下采用策略迭代算法和Q学习算法的系统优化的优化曲线,此时折扣因子 $\alpha = 0.01$.由图3可知,平均和折扣准则下采用策略迭代算法对系统进行优化,在第2步或第3步就已经收敛,得到系统的最优策略.在平均准则下采用Q学习算法的优化曲线,每学3000

步评估一次,由图可知曲线在第100次评估时算法基本稳定.与策略迭代算法相比,Q学习算法的收敛步数更大.

表2 物理参数表

Table 2 Physical parameters

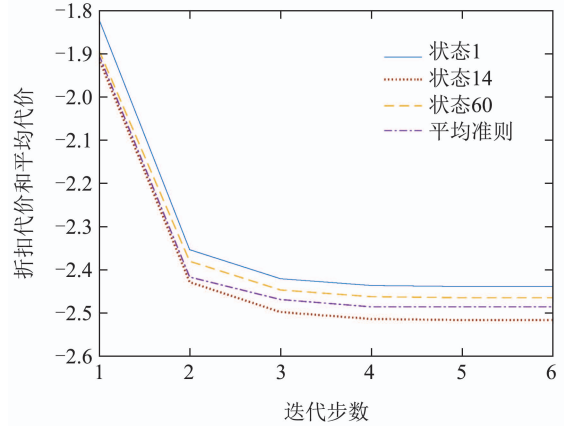
物理参数	取值	物理参数	取值
t_{heat}	0.2 h	t_{cool}	0.17 h
θ_{envir}	30°C	θ_{melt}	220°C
θ_{plat}	80°C	ρ	1.04 g · cm ⁻³
d	1 mm	δ	2.4 J · (g · °C) ⁻¹
a_{scan}	4000 mm · min ⁻¹	$N_{Length} \times N_{Width}$	(15 × 15) cm ²
t_{move}	2 s		

表3 仿真参数表

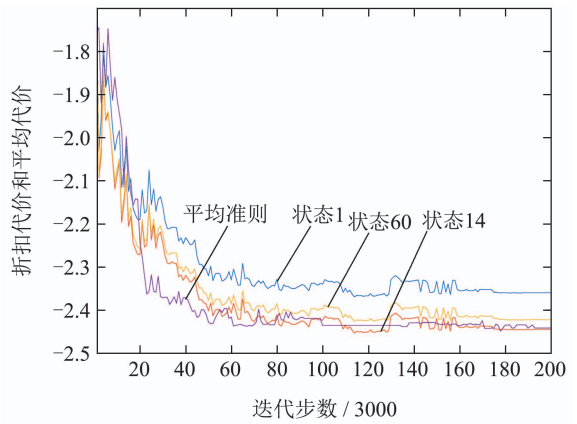
Table 3 Simulation parameters

仿真参数	取值
M	3
$[q_1 \ q_2 \ q_3]$	[32 27 16]
$[o_1 \ o_2 \ o_3]$	[10 × 10 7.5 × 15 5 × 15]
$[h_1 \ h_2 \ h_3]$	[6 10 15]
$[C_1 \ C_2 \ C_3]$	[2 3 4]
$\frac{H_{min}}{H_{max}}$	0.1
Δh	0.3
$[\lambda_1 \ \lambda_2 \ \lambda_3]$	[0.2 0.3 0.5]
$[k_{1,1} \ k_{1,2} \ k_{1,3}]$	[0.4 0.3 0.2]
k_2	0.04
k_3	0.3
$[\xi_2 \ \xi_1 \ \xi_0]$	[-0.25 0.055 0.0315]
$[\vartheta_1 \ \vartheta_2 \ \vartheta_3]$	[-205 -180 -110]
$[\vartheta_{1,fix} \ \vartheta_{2,fix} \ \vartheta_{3,fix}]$	[-6.53 -6.93 -4.35]

通过图3可以看出折扣准则下的优化曲线是相互分离的,在不同状态下系统的代价不同.这里定义3个特殊状态.状态1:任务队列全为空.状态14:队列1长度为0,队列2长度为2,队列3长度为3.状态60:任务队列全为满.在状态1任务队列全为空时,由于没有进行打印,因此没有打印报酬,所以代价曲线在最上方.在状态14时队列1长度为0,队列2长度为2,队列3长度为3,此时可以组合成两个最优的打印件组合进行打印,因此获得较高的打印报酬.在状态60时,任务队列全为满,此时等待代价变大,同时由于系统对任务1的处理效率不高,所以代价曲线在状态14的代价曲线的上方.由策略迭代求解的最优策略可知,在状态14时,系统选择的打印件组合 $J = (0, 2, 0)$,参数层高 $h = 0.1739$,在状态60时,系统选择的打印件组合 $J = (0, 2, 0)$,层高参数 $h = 0.2415$.这表明当任务队列为满时,系统选择最优的打印件组合进行打印,层高参数设置较大,与实际情况相符合.



(a) 策略迭代算法



(b) Q学习算法

图3 平均和折扣准则下两种算法的优化曲线

Fig. 3 Optimization curves of two algorithms under average and discount criteria

为了进一步分析两种算法的特点,仿真实验给出了两种算法的运行时间以及优化后系统的平均代价、打印任务的处理率和产品的平均质量,其中策略迭代算法的结果为3996 s, -2.485, 0.977, 0.0328,而Q学习算法的运行时间以及优化后系统的平均代价、打印任务的处理率和平均质量分别为1222 s, -2.448, 0.972, 0.0327.同时为了对Q学习算法的结果进行方差分析,在实验中对Q学习算法进行5次独立试验取均值并统计各项数据的标准差,对应每项的标准差分别为32 s, 0.003, 0.0019, 0.000103,分析各项标准差可知Q学习算法比较稳定.本文还对两种算法的复杂度进行了对比分析,首先比较空间复杂度,根据实验部分设置的系统参数求得系统状态数为 $3 \times 4 \times 5 = 60$,策略迭代算法需要存储逗留时间分布矩阵、状态转移矩阵以及性能函数矩阵,由于这3个矩阵维度和系统状态数相同,且都是方阵,所需存储空间为 $3 \times 60 \times 60 = 10800$.而Q学习需要存储每个状态-行动对的值,所需存储空间为

$$\frac{60 \times 9 \times (0.3 - 0.1)}{0.01} = 10800.$$

两种存储空间相同.在迭代步数方面,策略迭代算法

在第3步就已经收敛,而Q学习算法在第100次评估时算法基本稳定.在运行时间方面,策略迭代算法运行时间为3996 s,然而Q学习算法运行时间为1222 s.通过上述结果可知,虽然策略迭代算法收敛速度较快,但运行时间更长,这是由于策略迭代算法需要进行矩阵求逆运算,所以算法运行时间长.Q学习算法运行速度快,当模型参数未知时,可采用Q学习算法进行求解.本文接下来对3D打印服务系统的性能分析,将依据策略迭代算法的求解结果.

本文研究的系统作为一种排队服务系统,通过设置合理的任务队列容量,可以提高系统性能.为了探究任务队列容量与系统性能之间的关系,本次实验同时比较了3个队列的容量变化对系统性能的影响,在此次实验中假设 $\lambda_1 = 0.2, \lambda_2 = 0.3, \lambda_3 = 0.5$,实验结果如图4所示.根据实验结果可知,任务队列1容量增大时,系统代价会有所上升,任务队列1容量增大到2以后,系统代价也有所上升,随着任务队列2和任务队列3的容量分别增大到2和3时,系统的平均代价逐渐收敛,这是由于两个第2种打印件和三个第3种打印件分别组成两种等高且完全占满工作台的打印件组合.处理这种能够占满工作台且等高的打印件组合,可以节省打印所花时间,提高系统的处理率.

本文同时研究了队列容量变化对不同打印任务的

处理率影响,如表4所示.由表4可知当某一种打印任务的队列容量增大时,对应种类打印任务的处理率有较大幅度的提高,当队列容量增大到一定值后打印任务的处理率基本不变.随着某一任务队列容量的增大,其他种类打印任务的处理率会小幅度下降,这是由于该种队列容量增大时,其接收对应种类任务的容量越大,当其处理速度不及打印任务到达速度时,任务产生堆积.同时由于任务等待代价随着队列长度增加而增加,因此会优先处理队列长度长的任务,其他打印任务的处理率会有所降低.

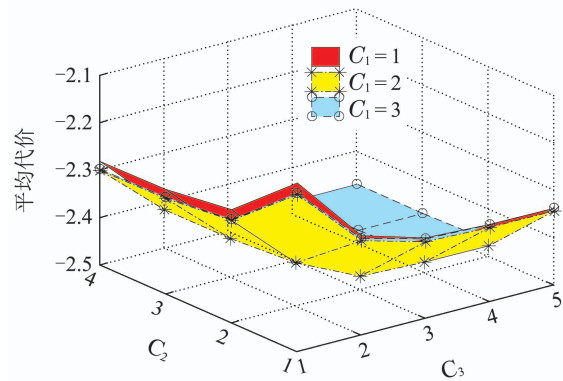


图4 不同队列容量下的平均代价

Fig. 4 Average cost under different queue capacities

表4 不同队列容量下系统处理率

Table 4 System processing rate under different queue capacities

$[C_1 \ C_2 \ C_3]$	[2 2 3]	[3 2 3]	[4 2 3]	[5 2 3]	[1 2 3]	[1 3 3]	[1 4 3]	[1 5 3]	[1 2 2]	[1 2 3]	[1 2 4]	[1 2 5]
任务1	0.938	0.988	1.000	1.000	0.805	0.796	0.790	0.793	0.849	0.805	0.801	0.795
任务2	0.936	0.943	0.931	0.929	0.941	0.991	1.000	1.000	0.957	0.941	0.946	0.943
任务3	0.968	0.965	0.960	0.958	0.967	0.966	0.967	0.964	0.858	0.967	0.989	0.999
总体	0.952	0.963	0.960	0.956	0.927	0.939	0.942	0.943	0.886	0.927	0.939	0.941

上面仿真实验研究了在任务到达率固定时队列容量对系统性能的影响,但是在实际系统中,任务到达率是变化的,因此本次仿真实验探究了任务到达率与系统性能之间的关系.图5为不同到达率下系统的平均报酬,在本次仿真实验中的队列容量设置为 $C_1 = 2, C_2 = 3, C_3 = 4$.由图可知随着任务到达率的增加,系统的平均报酬增大,这是由于在系统打印能力之内,到达率越大,系统单位时间打印的打印件越多,获得的报酬越大.

本次仿真实验还研究了在不同任务到达率下按照先到先服务(first-come-first-served, FCFS)规则、打印参数固定模式和打印参数可变模式这3种模式下系统的处理率和产品平均质量,如表5所示.随着任务到达率增大,系统的处理率和产品质量随之降低.但与FCFS规则和参数固定模式下系统性能相

比,参数可变模式下系统的处理率下降比较缓慢.这是由于任务到达率较小时,参数可变模式下系统为了获得较高的打印报酬,将打印参数层高设置得较小,提高了产品质量.随着打印任务到达率增大,任务队列中会堆积较多的打印任务,参数可变模式下的系统为了提高处理速度,将打印参数层高设置得较大,加快了打印任务的处理速度,因此能够及时处理到达的打印任务,打印任务的流失率小.通过比较可知,在任务到达率为[0.2 0.3 0.1]时,3种情况下系统的处理率十分接近,都在99%左右,但在参数可变模式下系统的产品的平均质量与在其他两种模式下系统产品质量相比提高了4.3%.在任务到达率为[0.5 0.3 0.5]时,参数可变模式下系统处理率比FCFS规则下系统的处理率高出6.6%.根据仿真实验结果可知,参数可变模式下系统对任务到达

率变化的适应性更强,当到达率较小或较大时,通过动态设置打印参数调整打印任务处理速度,提高系统的处理率。

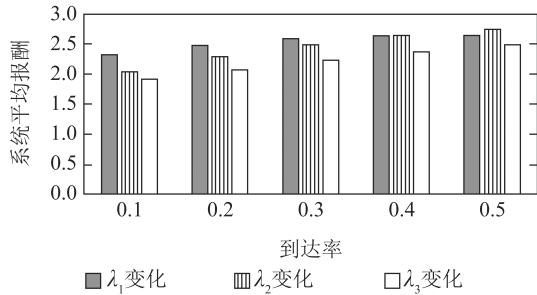


图5 不同到达率下的平均报酬

Fig. 5 Average reward under different arrival rates

表5 不同到达率下系统处理率和产品平均质量

Table 5 System processing rate and average product quality under different arrival rates

[λ_1 λ_2 λ_3]	处理率			平均质量		
	FCFS	参数固定	参数可变	FCFS	参数固定	参数可变
[0.2 0.3 0.1]	0.990	0.994	0.987	0.0325	0.0325	0.0339
[0.2 0.3 0.2]	0.988	0.990	0.986	0.0325	0.0325	0.0336
[0.2 0.3 0.3]	0.984	0.989	0.984	0.0325	0.0325	0.0334
[0.2 0.3 0.4]	0.979	0.984	0.983	0.0325	0.0325	0.0331
[0.2 0.3 0.5]	0.967	0.976	0.977	0.0325	0.0325	0.0328
[0.2 0.1 0.5]	0.994	0.992	0.987	0.0325	0.0325	0.0337
[0.2 0.2 0.5]	0.983	0.984	0.982	0.0325	0.0325	0.0333
[0.2 0.4 0.5]	0.939	0.957	0.966	0.0325	0.0325	0.0323
[0.2 0.5 0.5]	0.897	0.927	0.946	0.0325	0.0325	0.0319
[0.1 0.3 0.5]	0.989	0.993	0.987	0.0325	0.0325	0.0330
[0.3 0.3 0.5]	0.929	0.945	0.955	0.0325	0.0325	0.0323
[0.4 0.3 0.5]	0.878	0.905	0.924	0.0325	0.0325	0.0319
[0.5 0.3 0.5]	0.822	0.851	0.888	0.0325	0.0325	0.0315

5 总结

本文研究了在打印任务随机到达情况下,如何根据任务到达情况动态选择任务组合和层高的问题,并将优化问题描述为SMDP模型,采用策略迭代算法和Q学习算法这两种优化算法求解了系统的最优调度策略。仿真实验结果表明所求调度策略可以提高系统的处理率。本文目前只考虑了简单的排布问题以及单3D打印设备,如何将复杂排样问题纳入到整个优化问题中,以及如何对多台3D打印设备进行资源优化,是未来值得研究的问题。

参考文献:

[1] REN Y. A comparative study of intelligent manufacturing between china and the main developed countries. *Industrial Economy Review*, 2015, 8(2): 68 – 76.

[2] JIANG Yuejuan, LU Bingheng, FANG Xuewei, et al. 3D printing-based internet collect-manufacturing mode. *Computer Integrated Manufacturing Systems*, 2016, 22(6): 1424 – 1433. (姜月娟, 卢秉恒, 方学伟, 等. 基于3D打印的网络化集散制造模式研究. *计算机集成制造系统*, 2016, 22(6): 1424 – 1433.)

[3] REYMUS M, NINA L, STAWARCZYK B. 3D-printed material for temporary restorations: Impact of print layer thickness and post-curing method on degree of conversion. *International Journal of Computerized Dentistry*, 2019, 22(3): 231 – 237.

[4] HAO Nanhai. Batch scheduling of 3D printing based on work-piece layout. *Modern Manufacturing Engineering*, 2017, 1(5): 55 – 58. (郝南海. 考虑打印件排布的3D打印批次规划. *现代制造工程*, 2017, 1(5): 55 – 58.)

[5] IKURA Y, GIMPLE M. Efficient scheduling algorithms for a single batch processing machine. *Operations Research Letters*, 1986, 5(2): 61 – 65.

[6] LEE C Y, UZSOY R, MARTIN-VEGA L A. Efficient algorithms for scheduling semiconductor burn-in operations. *Operations Research*, 1992, 40(4): 764 – 775.

[7] UZSOY R. Scheduling a single batch processing machine with non-identical job sizes. *International Journal of Production Research*, 1994, 32(7): 1615 – 1635.

[8] CHENG B, LI K, CHEN B. Scheduling a single batch-processing machine with non-identical job sizes in fuzzy environment using an improved ant colony optimization. *Journal of Manufacturing Systems*, 2010, 29(1): 29 – 34.

[9] SHEN Fengping, LI Jinjin, YANG Yulong, et al. Batch scheduling model of heat treatment with minimization waiting time and the solution with PSO. *Computer Simulation*, 2020, 37(4): 200 – 204. (申风平, 李京京, 杨玉龙, 等. 极小化等待时间的热处理批调度模型与PSO解. *计算机仿真*, 2020, 37(4): 200 – 204.)

[10] GONG Hua, ZHANG Ermei, LIU Fang. Coordinated scheduling of transportation and production on batching-machine with tracking time constraint. *Control and Decision*, 2017, 32(6): 995 – 1000. (宫华, 张二梅, 刘芳. 传捆时间约束下的运输与批处理机生产协调调度. *控制与决策*, 2017, 32(6): 995 – 1000.)

[11] VERBIEST F, PINTO-VARELA T, CORNELISSENS T, et al. Decomposition approaches for the design and scheduling of multiproduct multistage batch plants with parallel lines. *Computers & Chemical Engineering*, 2019, 127(4): 111 – 126.

[12] GERMS R, FOREEST N D V. Order acceptance and scheduling policies for a make-to-order environment with family-dependent lead and batch setup times. *International Journal of Production Research*, 2013, 51(3): 940 – 951.

[13] WU Song, TANG Hao, TAN Qi. Online optimal control method of conveyor-serviced batch processing system. *Journal of System Simulation*, 2017, 29(4): 730 – 739. (吴松, 唐昊, 谭琦. 传送带给料分批加工系统在线优化控制方法. *系统仿真学报*, 2017, 29(4): 730 – 739.)

[14] WANG Bin. *Research on online optimization scheduling of 3D printing manufacturing system in multi-task printing mode*. Hefei: Hefei University of Technology, 2019. (王彬. 多任务打印模式下3D打印制造系统在线优化调度问题研究. 合肥: 合肥工业大学, 2019.)

[15] QING Q C, TANG H, ZHOU L, et al. The optimization control of s-single conveyor-serviced production station with variable service rate. *Proceedings of the 32nd Chinese Control Conference*. New York: IEEE, 2013: 2180 – 2184.

[16] TANG Hao, XU Lingling, ZHOU Lei, et al. Optimization control of demand-driven conveyor-serviced production station with changeable service rate. *Control Theory & Applications*, 2015, 32(6): 97 – 103. (唐昊, 许玲玲, 周雷, 等. 可变服务率模式下基于需求驱动的传送带给料加工站系统的优化控制. *控制理论与应用*, 2015, 32(6): 97 – 103.)

- [17] VIJAYARAGHAVAN V, GARG A, LAM J S L, et al. Process characterisation of 3D-printed FDM components using improved evolutionary computational approach. *International Journal of Advanced Manufacturing Technology*, 2015, 78(5/8): 781 – 793.
- [18] LIN Guangyi, LIU Feng, ZHAO Jihui, et al. Effects of layer thickness in FDM on mechanical properties and surface quality of printed products. *Plastics*, 2017, 46(3): 83 – 85.
(林广义, 刘峰, 赵辉绩, 等. FDM分层高度对制品力学性能和表面质量的影响. *塑料*, 2017, 46(3): 83 – 85.)
- [19] MAHAPATRA S S. Process characterisation of 3D-printed FDM components using improved evolutionary computational approach. *The International Journal of Advanced Manufacturing Technology*, 2015, 78(5/8): 781 – 793.
- [20] JIANG Xingbo, LV Xiaoqing, LIU Chengcheng. Lowest-level left align best-fit algorithm for the 2D rectangular strip packing problem. *Journal of Software*, 2009, 20(6): 1528 – 1538.
(蒋兴波, 吕肖庆, 刘成城. 二维矩形条带装箱问题的底部左齐择优匹配算法. *软件学报*, 2009, 20(6): 1528 – 1538.)
- [21] DOWSLAND K A, VAID S, DOWSLAND W B. An algorithm for polygon placement using a bottom-left strategy. *European Journal of Operational Research*, 2002, 141(2): 371 – 381.
- [22] CAO X R. Stochastic learning and optimization—A sensitivity-based approach. *Annual Reviews in Control*, 2009, 33(1): 11 – 24.
- [23] MULLER M, ELMAR W. An architecture for hybrid manufacturing combining 3D printing and CNC machining. *International Journal of Manufacturing Engineering*, 2016, 2016(1): 1 – 12.
- [24] YIN B Q, XI H S, ZHOU Y P. *Queuing System Performance Analysis and Process of Markov Control*. Hefei: China Science and Technology University Press, 2004.
(殷保群, 奚宏生, 周亚平. 排队系统性能分析与Markov控制过程. 合肥: 中国科学技术大学出版社, 2004.)
- [25] TANG H, YIN B Q, XI H S. Error bounds of optimization algorithms for semi-Markov decision processes. *International Journal of Systems science*, 2007, 38(9): 725 – 736.
- [26] SUTTON R, BARTO A. *Reinforcement Learning: An Introduction*. London: MIT Press, 2018.

作者简介:

王彬 博士研究生, 目前研究方向为智能生产系统、深度强化学习, E-mail: wang_bin@mail.hfut.edu.cn;

唐昊 博士后, 教授, 目前研究方向为离散事件动态系统、深度强化学习、神经元动态规划等, Email: htang@hfut.edu.cn;

戴飞 博士研究生, 目前研究方向为离散事件动态系统、深度强化学习、神经元动态规划等, Email: df8992@163.com;

谭琦 博士, 讲师, 目前研究方向为多目标优化方法、批调度及相关研究, Email: tanqi@hfut.edu.cn.