

文章编号:1671-6833(2022)06-0042-07

用于分割无损检测图像的快速模糊 C 均值算法

王军芬^{1,2}, 刘培跃³, 董建彬^{1,2}, 朱占龙¹

(1. 河北地质大学 信息工程学院,河北 石家庄 050031; 2. 河北地质大学 智能传感物联网技术河北省工程研究中心,河北 石家庄 050031; 3. 石家庄职业技术学院 机电工程系,河北 石家庄 050081)

摘 要:无损检测图像中目标类和背景类差异较大,模糊 C 均值算法无法有效地将目标分割出来,因此提出一种用于分割无损检测图像的快速模糊 C 均值算法。在聚类过程中,引入局部空间信息和灰度信息,以提高算法的鲁棒性;用条件值表征样本容量来平衡不同大小的类群,以解决类大小敏感问题;基于新的约束条件得到新的隶属度和聚类中心表达式,并给出算法具体步骤;对预处理后图像的灰度级进行分类,图像分割所需要的时间不再取决于图像的尺寸,而是图像的灰度级数,大幅度降低了算法的时间复杂度。采用类大小差异较大的合成图像和无损检测图像进行仿真实验,以分割精度(SA)、F-value、G-mean 以及图像分割所需要的时间为评价指标来评价算法的性能。实验结果表明:在原始测试图像被高斯噪声、椒盐噪声、瑞利噪声和乘性噪声污染时,与其他模糊聚类算法相比,本文算法具有更好的鲁棒性,分割精度更高,为 97.93%,F-value 为 88.50%,G-mean 为 93.83%,图像分割时间也更少,为 14.06 ms。实验证明了本文算法的有效性。

关键词:模糊 C 均值算法;图像分割;无损检测;鲁棒性

中图分类号: TP391.4 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2022.04.016

0 引言

图像分割在许多领域都有着重要的应用,如机器视觉、医学影像、目标识别等^[1-3]。模糊分割方法,尤其是模糊 C 均值(FCM)算法^[4]及其变化在许多图像分割和目标分类应用中取得了良好的效果^[5-11]。但是,对于不平衡数据,FCM 的分割效果并不理想。到目前为止,仅有少数研究试图改善不平衡数据的聚类性能,例如有半监督的模糊 C 均值(ssFCM)算法^[5];对聚类尺寸不敏感的模糊 C 均值(cluster size insensitive FCM,CsiFCM)算法^[12],其实验结果表明 CsiFCM 的分割效果(或聚类性能)优于 ssFCM;基于完整性的、对聚类尺寸不敏感的模糊 C 均值算法(size-insensitive integrity-based FCM,siibFCM)^[13],其聚类思想与 CsiFCM 大致相同,都是对较大的类群赋较小的条件值,较小的类群赋较大的条件值,不过 siibFCM 设计了更为复杂的条件值计算方法;多聚类中心

的模糊 C 均值算法^[14]在聚类过程中,设置类的个数大于实际值,这样多数类就会被拆分到少数类中,聚类之后再合并,从而改善聚类尺寸敏感问题。

当图像被噪声污染后,这些算法针对不平衡数据分割的鲁棒性较差,分割效果不理想。为了改善上述情况,本文提出一种快速鲁棒的,且对聚类尺寸不敏感的模糊 C 均值聚类图像分割算法。首先,利用局部空间算子和局部灰度算子对图像 I_0 进行预处理得到非线性加权求和图像 I ,以提高算法的鲁棒性;其次,对非线性加权求和图像 I 的灰度级进行分类,在分类的过程中引入条件变量 f_g ,将低条件值分配给较大类群的对象,并将高条件值分配给较小类群的对象来防止小类群向大类群漂移^[12],图像的分割时间不再取决于图像的大小,而是图像的灰度级数,从而大大降低算法的时间复杂度。

收稿日期:2022-05-14;修订日期:2022-06-20
基金项目:国家自然科学基金资助项目(61806069);河北省高等学校科学技术研究项目(QN2020263)
作者简介:王军芬(1979—),女,河北石家庄人,河北地质大学讲师,硕士,主要从事数字图像处理和信
息处理研究, E-mail:wangjunfen1979@163.com。
通信作者:朱占龙(1984—),男,河北石家庄人,河北地质大学讲师,博士,主要从事数字图像处理和信
息处理研究, E-mail:zzl_seu@163.com。

1 基于灰度级的模糊C均值算法

灰度图像的尺寸 $N = m \times n$, 灰度级数为 G , 图像灰度级的集合 $I = \{I_1, I_2, \dots, I_G\}$, I_g 表示第 g 个灰度级的值, 把图像分割为 C 个区域, 得到 C 个模糊子集, 每一个模糊子集对应一个聚类中心 v_i , 聚类中心集合 $V = \{v_1, v_2, \dots, v_C\}$ 。目标函数为

$$J_m = \sum_{i=1}^C \sum_{g=1}^G \mu_{ig}^m \|I_g - v_i\|^2. \quad (1)$$

在满足 $\sum_{i=1}^C \mu_{ig} = 1, g = 1, 2, \dots, G$ 约束条件下利用拉格朗日乘子法优化目标函数, 可得隶属度为

$$\mu_{ig} = \frac{1}{\left(\frac{\|I_g - v_i\|^2}{\sum_{j=1}^C \|I_g - v_j\|^2} \right)^{\frac{1}{m-1}}}; \quad (2)$$

$$v_i = \frac{\sum_{g=1}^G \mu_{ig}^m I_g}{\sum_{g=1}^G \mu_{ig}^m}. \quad (3)$$

式中: $i = 1, 2, \dots, C; g = 1, 2, \dots, G$ 。

利用式(2)和式(3)不断迭代更新直到满足设定的终止条件, 得到各聚类中心和每个灰度级的隶属度值, 根据式(4)将灰度级 I_g 分配给具有最高隶属度的 k 类:

$$k = \operatorname{argmax} \mu_{ig}, i = 1, 2, \dots, C. \quad (4)$$

对于不平衡数据, 由于FCM使用了一个平方和形式的代价函数, 使得小类群的中心倾向于向邻近的大类群漂移^[12], 分割结果趋向于均衡聚类类群, 分割效果不理想。

2 本文算法

原始灰度图像 I_0 大小 $N = m \times n$, 总像素集合 $X = \{x_1, x_2, \dots, x_N\}$, 灰度级数为 G 。本文中图像的预处理是指对原始图像 I_0 按照式(5)进行预处理, 形成一个非线性加权求和图像 I :

$$I_k = \frac{\sum_{j \in N_k} S_{kj} x_j}{\sum_{j \in N_k} S_{kj}}. \quad (5)$$

式中: $I_k \in [0, G-1]$ 为非线性加权求和图像 I 的第 k 个像素的灰度值; N_k 为以第 k 个像素为中心, 大小为 $M \times M$ 的窗口邻域像素集合; x_j 为以第 k 个像素为中心的窗口中第 j 个像素的灰度值; S_{kj} 为第 k 个像素和第 j 个像素之间的局部相似性度量, 表示第 j 个像素的权重, 具体表述见文献[6]。

2.1 分割处理

非线性加权求和图像 I 中 $G \ll N$, 将图像分割为 C 个区域, 可得到 C 个模糊子集, 每一个模糊子集对应一个聚类中心 v_i , 聚类中心集合为 $V = \{v_1, v_2, \dots, v_C\}$ 。目标函数定义为

$$J_s = \sum_{i=1}^C \sum_{g=1}^G \gamma_g \mu_{ig}^m (I_g - v_i)^2. \quad (6)$$

式中: μ_{ig} 为第 g 个灰度级相对于聚类中心 v_i 的模糊隶属度; γ_g 为第 g 个灰度级像素的数目, 且有 $\sum_{g=1}^G \gamma_g = N$ 。

为了改善算法对不平衡数据的聚类性能, 在聚类过程中, 引入条件变量 f_g ($0 \leq f_g \leq 1$), 将低条件值分配给较大类群的对象, 并将高条件值分配给较小类群的对象来防止小类群向大类群漂移^[12]。定义条件变量 f_g 为

$$f_g = \frac{1}{1 - P_{i_{\min}}}(1 - P_i). \quad (7)$$

式中: $P_{i_{\min}}$ 为所有类群先验概率的最小值; f_g 为第 g 个灰度级去模糊后被分类为 i 类的条件值; P_i 为第 i 类类群先验概率。

$$P_i = \frac{N_i}{N}. \quad (8)$$

式中: N_i 为属于类群 i 的所有灰度级 g 的像素总数。

$$N_i = \sum_g \gamma_g. \quad (9)$$

每个灰度级一旦被分配到某个特定的类群, 就具有相对于所有类群的固定条件值, 并且同一类群中的所有灰度级都具有相同的条件值。

此时 $\sum_{i=1}^C \mu_{ig} = f_g, g = 1, 2, \dots, G$ 。利用拉格朗日乘子法优化目标函数, 可得隶属度和聚类中心的更新式, 如式(10)、(11)所示:

$$\mu_{ig} = \frac{f_g}{\left(\frac{\|I_g - v_i\|^2}{\sum_{j=1}^C \|I_g - v_j\|^2} \right)^{\frac{1}{m-1}}}; \quad (10)$$

$$v_i = \frac{\sum_{g=1}^G \gamma_g \mu_{ig}^m I_g}{\sum_{g=1}^G \gamma_g \mu_{ig}^m}. \quad (11)$$

式中: $i = 1, 2, \dots, C; g = 1, 2, \dots, G$ 。

2.2 时间复杂度分析

本文算法的时间复杂度包括图像的预处理和分割处理。图像预处理主要是为了获得非线性加权求和图像 I , 其时间复杂度为 $O(nM^2)$; 分割处理主要是算法的迭代过程, 其时间复杂度为 $O(GC^2 \text{Iter})$, Iter 为迭代次数。所以本文算法的时

间复杂度为 $O(nM^2 + GC^2Iter)$ 。

2.3 具体步骤

步骤 1 设置图像分割的区域数目 $C \geq 2$ 、迭代终止条件 $\varepsilon > 0$ 、模糊指数 m 、最大的迭代次数 $Iter_{max}$;

步骤 2 图像预处理,利用式(5)计算非线性加权求和图像 I ;

步骤 3 用 0 到 $G - 1$ 之间的随机值初始化类中心 $V^{(0)} = \{v_1^{(0)}, v_2^{(0)}, \dots, v_c^{(0)}\}$;

步骤 4 初始化 $f_g = 1/C, g = 1, 2, \dots, G$;

步骤 5 用式(10)更新隶属度;

步骤 6 根据隶属度按式(4)将每个灰度级 g 分配到适当的类群;

步骤 7 用式(7)更新 f_g ;

步骤 8 用式(11)更新各个类的聚类中心 V ;

步骤 9 重复步骤 5~8 直到满足终止条件:迭代次数 $Iter > Iter_{max}$ 或 $|V^{(Iter)} - V^{(Iter-1)}| < \varepsilon$;

步骤 10 根据最终的隶属度按式(4)对图像 I 中的每个灰度级进行去模糊。

3 实验与结果分析

3.1 实验参数设置及说明

实验图像由合成图像和无损检测图像构成,合成图像由 4 类图像数据组成,其中大类群与小类群的比值为 29.16。选择无损检测图像 Im1~Im6,其图像灰度分布具有不均衡性,目标类和背景类差异较大,所选图像满足对实验图像的要求。

首先,本文研究了条件变量 f_g 随 N_i 的变化情况,其次,为了更好地验证算法优越性,用高斯噪声(均值为 0,归一化方差为 2%~8%)、椒盐噪声(噪声百分比为 10%~15%)、瑞利噪声(均值为 0,标准差为 15~30)和乘性噪声(均值为 0,方差为 4%~8%)去污染实验图像,使用 FCM、CsiFCM、siibFCM 作为对比算法进行实验。设置模糊指数 $m=2$,迭代精度 $\varepsilon=0.000\ 01$,最大迭代次数为 100。为了使分割结果不受初始类中心的影响,

在相同噪声水平下进行对比实验时,选用相同的初始类中心。实验所用计算机的 CPU 为 Intel CORE i7-8565U@ 1.80 GHz, RAM 为 8 GB,运行环境为 Windows 7,编程环境为 MATLAB R2014a。

针对不平衡数据聚类的图像分割结果,本文采用了 3 种评价标准:分割精度(SA)、 $F-value^{[15]}$ 、 $G-mean^{[16]}$ 。分割精度定义为分类正确的像素数与图像像素总数的比值; $F-value$ 用来评价聚类算法对少数类的聚类性能; $G-mean$ 用来评价聚类算法的总体性能。 SA 、 $F-value$ 、 $G-mean$ 值越大说明图像分割效果越好。当图像分割区域数为 2 时,选用评价标准 $F-value$ 、 $G-mean$;当图像分割区域数多于 2 时,选用 SA 。

3.2 条件变量 f_g 随 N_i 的变化过程

合成图像像素为 $256 \times 256 = 65\ 536$,共由 4 类图像数据组成,如图 1(a)所示。类 1:灰度值为 0 的小正方形,1 600 像素;类 2:灰度值为 80 的水平线,8 640 像素;类 3:灰度值为 160 的垂直线,8 640 像素;类 4:灰度值为 255 的大正方形,46 656 像素。

以被高斯噪声(均值为 0,归一化方差为 2%)污染后的图像为实验图像,如图 1(b)所示,在图像的分割过程中条件变量 f_g 随 N_i 的变化情况如图 1(c)、1(d)所示。从图 1(c)、1(d)中可以看出,当类群所包含的像素数目越大时,分配给所有属于类群的灰度级的条件值就越小;当类群 i 所包含的像素数目越小时,分配给所有属于类群的灰度级的条件值就越大。 f_g 可以平衡不同大小的类群,解决类大小敏感问题。

3.3 合成图像分割实验

合成图像添加不同强度的噪声进行图像分割测试,结果如表 1 所示,其中 NA 表示失效。表 1 表明:从整体上看本文算法分割精度 SA 为 97.93%,明显优于对比算法,甚至当对比算法失效时,本文算法仍能够得到较好效果。当图像没有噪声或被瑞利噪声污染时,本文算法与对

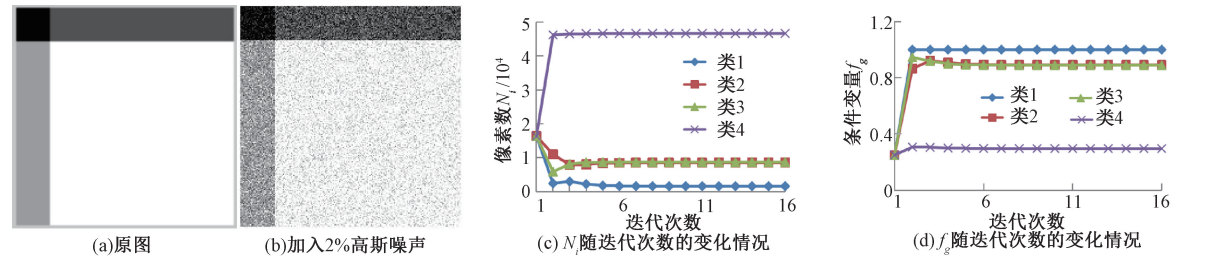


图 1 合成图像及 N_i 、 f_g 测试结果

Figure 1 Synthetic image and test results of N_i and f_g

表 1 不同噪声水平下的分割精度对比

Table 1 Segmentation accuracy with different level noises

噪声水平	初始类中心	SA/%			
		FCM	CsiFCM	siibFCM	本文算法
无噪声	(101,64,221,49)	NA	100.00	NA	100.00
高斯 2%	(124,2,187,253)	71.74	76.20	83.47	99.61
高斯 8%	(59,28,230,122)	60.56	61.62	66.24	90.43
椒盐 10%	(137,246,756,67)	93.85	93.61	NA	99.18
椒盐 15%	(106,35,168,187)	NA	90.45	NA	97.39
瑞利 15	(139,26,148,228)	81.31	81.74	100.00	99.96
瑞利 30	(35,253,34,87)	99.15	98.95	98.89	99.84
乘性 4%	(218,31,44,103)	74.96	75.40	NA	99.52
乘性 8%	(206,5,4,81)	57.03	63.33	NA	95.48
平均值		76.94	82.37	87.15	97.93

比算法性能基本一致。但是当高斯噪声、椒盐噪声和乘性噪声归一化方差增大时,对比算法的分割精度都明显降低,而本文所提算法可以达到 90% 以上的分割精度,其主要原因是本文算法在聚类过程中引入了条件变量,提高了算法的分割精度。图 2 给出了当高斯噪声的归一化方差为 2% 时,各算法的分割效果。从图 2 中可以看出,本文算法对像素的不均性分布不敏感,分割效果受噪声影响较小,鲁棒性要优于对比算法。

3.4 无损检测图像分割实验

为了验证本文算法在无损检测图像分割方面的优越性,对尺寸不同、背景类与目标类大小比值不同,但差异较大的 6 幅无损检测图像进行图像分割,分为背景和目标物两类。无损检测图像 Im1~Im6 如图 3 所示,其图像像素依次为 131×232、51×98、70×100、74×111、56×271、256×256,评价指标为 *F-value*、*G-mean* 以及分割图像所需要

的时间(*t*)。本文算法与对比算法在不同噪声水平下各指标测试结果如表 2、图 4 所示。其中的 *G-mean* 和 *t* 为图像在不同噪声水平下的平均值。

当图像被高斯噪声、椒盐噪声、瑞利噪声以及乘性噪声污染时,FCM、CsiFCM、siibFCM 以及本文算法对实验中所有无损检测图像的 *F-value* 平均值分别为 64.78%、67.90%、69.41%、88.50%。另外,FCM、CsiFCM、siibFCM 以及本文算法对实验中所有无损检测图像的 *G-mean* 平均值分别为 84.81%、81.47%、83.07%、93.83%。*F-value* 和 *G-mean* 指标优于对比算法的主要原因是本文采用了条件值表征样本容量来平衡不同大小的类群,因此提升了对少数类的聚类性能。

从表 2 和图 4(a)中可以看出,当无损检测图像被污染时,本文算法的聚类性能要优于对比算法。但是当无损检测图像没有被污染或者当图像被非均匀分布的瑞利噪声污染时,本文算法性能

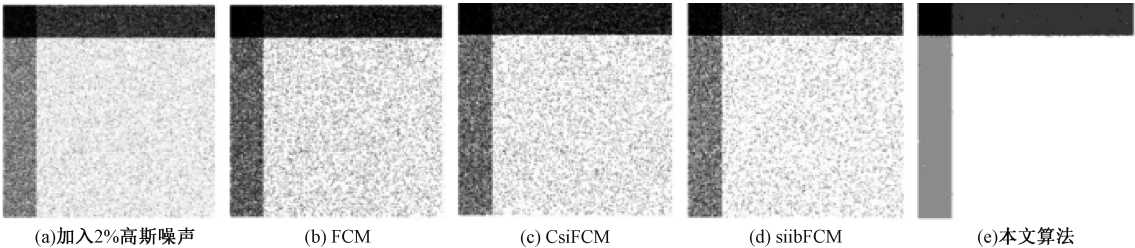


图 2 高斯噪声 2% 时合成图像各算法分割结果

Figure 2 Segmentation results on synthetic image corrupted by 2% Gaussian noise

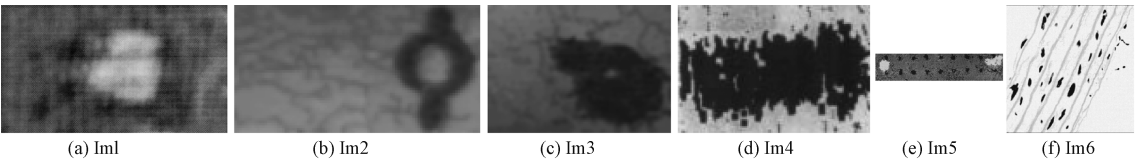


图 3 无损检测图像

Figure 3 Non-destructive testing images

表 2 不同算法得到的 F -value

Table 2 F -value obtained by different algorithms

图像	噪声水平	F -value/%				图像	噪声水平	F -value/%			
		FCM	CsiFCM	siibFCM	本文算法			FCM	CsiFCM	siibCM	本文算法
Im1	无噪声	96.26	98.06	98.06	98.37	Im2	无噪声	89.45	95.4	97.27	94.69
	高斯 2%	76.99	80.83	79.41	97.94		高斯 2%	40.14	42.46	41.71	73.06
	高斯 8%	71.08	72.17	70.78	88.40		高斯 8%	32.18	32.86	32.01	51.39
	椒盐 10%	95.09	96.49	95.80	97.99		椒盐 10%	81.72	9.24	80.61	88.85
	椒盐 15%	94.23	95.04	94.19	97.36		椒盐 15%	71.58	25.48	47.41	85.61
	瑞利 15	93.31	97.58	96.02	98.16		瑞利 15	62.48	88.86	89.88	93.90
	瑞利 30	85.65	94.07	94.42	98.21		瑞利 30	47.96	52.29	56.82	91.02
	乘性 4%	83.53	96.09	93.46	98.02		乘性 4%	41.29	41.50	43.22	81.28
Im3	乘性 8%	79.33	85.70	87.00	97.05	Im4	乘性 8%	39.50	39.79	39.40	65.31
	无噪声	88.20	89.25	84.04	89.17		无噪声	97.17	97.17	96.5	96.62
	高斯 2%	60.66	60.98	58.74	78.36		高斯 2%	92.96	93.11	91.57	96.52
	高斯 8%	50.58	49.58	48.23	69.57		高斯 8%	79.82	80.49	77.93	93.99
	椒盐 10%	80.39	82.68	41.16	86.57		椒盐 10%	92.38	92.23	91.74	96.19
	椒盐 15%	73.33	40.47	40.49	85.43		椒盐 15%	89.54	90.04	89.31	95.39
	瑞利 15	83.54	85.15	78.51	88.39		瑞利 15	96.96	96.97	96.34	96.61
	瑞利 30	74.12	75.91	70.57	87.57		瑞利 30	96.25	96.42	95.61	96.48
Im5	乘性 4%	66.48	65.71	59.14	84.12	Im6	乘性 4%	94.29	94.27	89.90	95.93
	乘性 8%	61.19	60.29	57.20	82.28		乘性 8%	88.05	86.83	83.56	95.51
	无噪声	20.14	97.38	89.70	96.32		无噪声	99.78	92.72	90.85	90.23
	高斯 2%	21.70	22.35	24.72	92.96		高斯 2%	27.19	39.34	85.05	87.04
	高斯 8%	18.73	18.80	20.50	28.19		高斯 8%	21.15	28.51	36.14	84.62
	椒盐 10%	67.01	66.59	66.20	94.03		椒盐 10%	54.06	51.80	51.43	88.04
	椒盐 15%	57.59	11.47	56.99	91.13		椒盐 15%	42.17	41.67	41.82	86.28
	瑞利 15	22.71	92.72	80.36	95.96		瑞利 15	97.23	92.47	90.20	89.62
	瑞利 30	24.77	79.33	44.25	95.19		瑞利 30	94.34	90.54	91.34	90.39
	乘性 4%	24.35	27.78	43.68	93.65		乘性 4%	17.79	90.67	91.07	88.86
	乘性 8%	21.93	22.23	30.60	86.83		乘性 8%	17.92	18.70	25.44	88.39

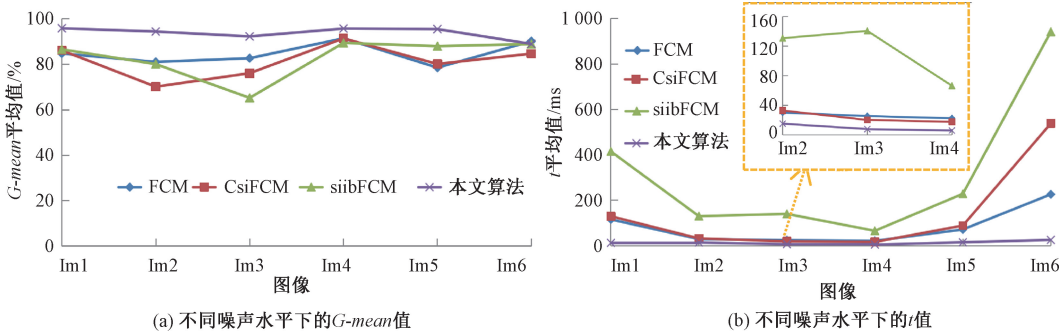


图 4 不同算法得到的 G -mean 和 Time

Figure 4 G -mean and Time obtained by different algorithms

与对比算法基本一致或者稍微弱于对比算法。其主要原因是本文算法采用了对图像中少数类加权的策略,尤其是针对非均匀噪声的攻击,在增强了少数类权重的同时也增加了该区域的噪声强度,导致本文算法对部分图像(如 Im6)的聚类性能略微低于对比算法。

图 4(b) 为本文算法和对比算法对无损检测

图像在不同噪声水平下分割时间平均值对比图。可以看出,本文算法的分割速度更快,本文算法对实验中所有无损检测图像分割时间的平均值为 14.06 ms。其主要原因是本文算法采用对图像的灰度级进行分类,而不是图像的像素,造成了当图像的尺寸较小时,如 Im2、Im3 和 Im4,本文算法只是稍优于对比算法,也说明本文算法适合图像尺

寸较大的场合。

图 5 为本文算法与对比算法对 Im5 的分割效果图。其中图 5(a)为原图;图 5(b)为噪声污染后的图像;图 5(c)为标准分割结果;图 5(d)~5(g)为本文算法和对比算法的分割结果。从图 5

中可看出,对比算法对噪声比较敏感,而本文算法鲁棒性较高,受噪声的影响较小,分割效果优于对比算法。其主要原因是本文算法引入局部空间信息和灰度信息对图像进行预处理,提高了算法的鲁棒性。

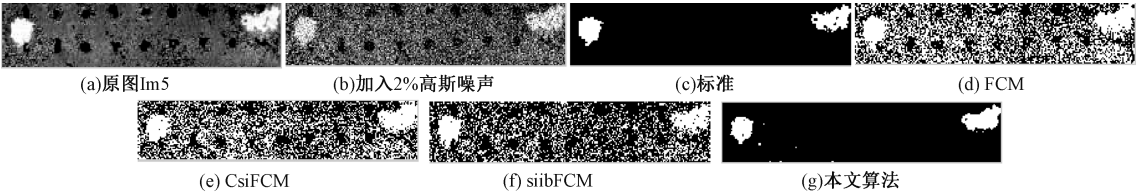


图 5 Im5 图像及其各算法的分割结果

Figure 5 Comparison of segmentation results on non-destructive testing Im5

4 结论

为了改善模糊 C 均值算法不能有效分割类间差异较大的图像这一问题,提出了一种新的用于分割无损检测图像的快速模糊 C 均值算法。该算法主要是通过引入局部空间信息和灰度信息,降低噪声的影响,用条件值表征样本容量来平衡不同大小的类群,避免较小类群的聚类中心向较大类群偏移。基于新的约束条件得到新的隶属度和聚类中心表达式,对预处理后图像的灰度级进行分类,大幅度地降低了算法的时间复杂度。实验数据证明了本文算法的有效性和鲁棒性。下一步研究可改进算法对非均匀噪声的处理及初始化方式,以便更加有效地对不平衡图像数据进行精细的分割。

参考文献:

[1] 刘今越,李顺达,陈梦倩,等. 面向移乘搬运护理机器人的人体姿态视觉识别[J]. 机器人, 2019, 41(5): 601-608.
LIU J Y, LI S D, CHEN M Q, et al. Visual recognition of human pose for the transfer-care assistant robot [J]. Robot, 2019, 41(5): 601-608.

[2] 潘沛克,王艳,罗勇,等. 基于 U-net 模型的全自动鼻咽肿瘤 MR 图像分割[J]. 计算机应用, 2019, 39(4): 1183-1188.
PAN P K, WANG Y, LUO Y, et al. Automatic segmentation of nasopharyngeal neoplasm in MR image based on U-net model[J]. Journal of computer applications, 2019, 39(4): 1183-1188.

[3] 彭金柱,董梦超,杨扬. 基于视觉和肌电信息融合的手势识别方法[J]. 郑州大学学报(工学版), 2021, 42(2): 67-73.
PENG J Z, DONG M C, YANG Y. Human gesture

recognition method based on vision and EMG signal information[J]. Journal of Zhengzhou university (engineering science), 2021, 42(2): 67-73.

[4] BEZDEK J C. A convergence theorem for the fuzzy ISODATA clustering algorithms[J]. IEEE transactions on pattern analysis and machine intelligence, 1980, 2(1): 1-8.

[5] BENSAID A M, HALL L O, BEZDEK J C, et al. Partially supervised clustering for image segmentation [J]. Pattern recognition, 1996, 29(5): 859-871.

[6] CAI W L, CHEN S C, ZHANG D Q. Fast and robust fuzzy c-means clustering algorithms incorporating local information for image segmentation[J]. Pattern recognition, 2007, 40(3): 825-838.

[7] ZHAO F, JIAO L C, LIU H Q, et al. A novel fuzzy clustering algorithm with non local adaptive spatial constraint for image segmentation[J]. Signal processing, 2011, 91(4): 988-999.

[8] GONG M G, LIANG Y, SHI J, et al. Fuzzy c-means clustering with local information and kernel metric for image segmentation[J]. IEEE transactions on image processing, 2013, 22(2): 573-584.

[9] MEMON K H, LEE D H. Generalised kernel weighted fuzzy C-means clustering algorithm with local information[J]. Fuzzy sets and systems, 2018, 340: 91-108.

[10] LEI T, JIA X H, ZHANG Y N, et al. Significantly fast and robust fuzzy c-means clustering algorithm based on morphological reconstruction and membership filtering[J]. IEEE transactions on fuzzy systems, 2018, 26(5): 3027-3041.

[11] ZHANG Y X, BAI X Z, FAN R R, et al. Deviation-sparse fuzzy c-means with neighbor information constraint[J]. IEEE transactions on fuzzy systems, 2019, 27(1): 185-199.

[12] NOORDAM J C, Van Den BROEK W H A M, BUY-

DENS L M C. Multivariate image segmentation with cluster size insensitive fuzzy *c*-means[J]. Chemometrics and intelligent laboratory systems, 2002, 64(1): 65–78.

[13] LIN P L, HUANG P W, KUO C H, et al. A size-insensitive integrity-based fuzzy *c*-means method for data clustering [J]. Pattern recognition, 2014, 47(5): 2042–2056.

[14] LIANG J Y, BAI L, DANG C Y, et al. The *K*-means-type algorithms versus imbalanced data distributions [J]. IEEE transactions on fuzzy systems, 2012, 20(4): 728–745.

[15] HAN H, WANG W Y, MAO B H. Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning[J]. Lecture notes in computer science, 2005, 3644:878–887.

[16] SU C T, CHEN L S, YIH Y. Knowledge acquisition through information granulation for imbalanced data [J]. Expert systems with applications, 2006, 31(3): 531–541.

Fast Fuzzy C Means Algorithm for Segmentation of Non-destructive Testing Image

WANG Junfen^{1,2}, LIU Peiyue³, DONG Jianbin^{1,2}, ZHU Zhanlong¹

(1. School of Information Engineering, Hebei GEO University, Shijiazhuang 050031, China; 2. Intelligent Sensor Network Engineering Research Center of Hebei Province, Hebei GEO University, Shijiazhuang 050031, China; 3. Department of Electromechanics Engineering, Shijiazhuang Vocational Technology Institute, Shijiazhuang 050081, China)

Abstract: The Fuzzy *C* means algorithm cannot effectively segment object pixels from a non-destructive testing (NDT) image because of the great difference between the background and the object region. Therefore, a fast fuzzy *C* means algorithm for segmentation of NDT image was proposed in this study. In the process of clustering, local spatial information and gray information were introduced to improve the robustness of the algorithm. The condition value was used to represent the sample size to balance clusters of different sizes and solve the cluster-size sensitivity problem. Based on the new constraints, a new form of membership degree and cluster center representation could be obtained. The execution time of image segmentation was no longer determined by the size of the image, but by the gray level of the image. The computational complexity of the algorithm was greatly reduced. NDT images and synthetic images with a large difference in cluster size were used for testing. The segmentation accuracy (*SA*), *F-value*, *G-mean* and the execution time of image segmentation were used to evaluate the performance of algorithms. The experimental results showed that when the original test image was polluted by Gaussian noise, salt and pepper noise, Rayleigh noise and multiplicative noise, the proposed algorithm had better robustness, *SA* is 97.93%, *F-value* is 88.50%, *G-mean* is 93.83%, the execution time of image segmentation was less, about 14.06 ms. The simulation tests could verify the effectiveness of the proposed algorithm.

Keywords: fuzzy *C* means algorithm; image segmentation; non-destructive testing; robustness