

基于多层次灰关联分析的图像分割性能评估

The Performance Evaluation Based on the Grey Multi-hierchal Appraise Model of Image Segmentation

狄宇春¹ 邓雁萍²

(宁波大学科技学院计算机系 宁波315211)¹

(上海交通大学图像处理与模式识别研究所 上海200030)²

Abstract Image segmentation is the key component in automatic target recognition system (ATRS). The evaluation of a segmentation algorithm plays an important role in the estimation of performance for an off-line ATRS. A model of grey multi-hierchal appraise combining grey appraise method with the Analytic Hierarchy Process (AHP) is proposed in this paper. It is used to evaluate the performance of image segmentation. The performance evaluation can be performed from both the global or local aspects. Its application shows that the new method is more reasonable than other methods.

Keywords Grey relational grade, Analytic hierarchy process, Automatic target recognition, Segmentation, Performance evaluation

1 引言

目标图像的自动分割是自动目标识别系统(ATRS)的关键部分,而 ATRS 技术是成像制导的关键环节,其结果直接影响到其后的信息处理过程,评估分割算法的性能是 ATRS 离线评估的重要组成部分之一。分割算法的性能评估可以在实验和理论两个方面为改进算法的设计提供定性和定量的依据。

在图像分割及自动目标识别任务中,用于识别的特征的选择是随具体的识别任务而确定的,各特征在识别任务中所起的作用也不是完全等同的。那么我们在对图像分割及自动目标识别系统(ATRS)的性能进行评估时必须根据其识别任务的特点来确定评估的性能指标(准则),实际工作对评估的方法的实用性(如通用性和特殊性)提出了要求(譬如从整体性能上的综合评估,或从某个侧面,或从某个特性如实时性上来进行评估,指标的多层次性要求)。这就要求建立相应的评估模型来适应工作的要求。目前对该问题的研究已有一些成果^[1~5],如 ERIM (Environmental Research Institute of Michigan)、TI(Texas Instruments)以及 Honeywell 采用集合运算中的“交”、“并”算子分别提出了一种性能评估指标^[1],文[2]在文[1]的基础上进一步提出了在进行评估分割算法性能时,不仅要考虑分割结果(ST)在图像中的面积因素,同时需要考虑目标的形状因素。现已提出的许多评估方法^[1~5]大多是从某一个或少数几个方面进行的,存在指标值奇异性和对分割性能不灵敏的现象。本文根据运筹学和系统工程思想,采用多层次灰色关联度分析和层次分析法,提出了一个基于多层次灰色关联度分析的图像分割性能评估模型 MGAHP,并通过一些试验对其进行分析。应用实例显示,该模型评判的结果与算法的性能

一致,评估结果数据离散度高,克服了以往方法不灵敏的不足,同时也避免了奇异性,比现有结果更为有效;该方法计算量小,对样本的数量没有要求,也不需要典型的分布规律。需要指出的是由于分割评估缺少绝对正确的分割参照结果(RT),因而这是一个病态的问题。不过,本文仅局限于分割算法的性能评估,而不涉及如何取得分割的 RT 过程,所有图像的 RT 将作为已知项来处理。

2 多层次灰关联综合评估模型

基于单层次灰色关联度分析综合评判建模过程可用简单的框图1来描述。图中的 X 表示经归一化处理后的属性矩阵,通过关联度分析,得其关于优向量和次向量的关联度,经相关计算、排序得评估结果^[7]。

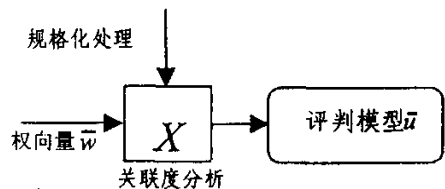


图1

在实际问题中,譬如 ATRS 的性能评估,由于需要考虑的因素往往很多,而且不同因素间还存在不同的层次,因此需要考虑多层次的综合评判问题。多层次综合评判是在单层次综合评判的基础上进行的。多层次综合评判模型(以二级为例)框图如图2。图中的 $X_1, \bar{w}_1, X_2, \bar{w}_2, \dots, X_n, \bar{w}_n$ 分别表示在一级评估的经归一化处理后的各子系统的属性矩阵和权向量。 $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n$ 分别表示在一级评估中各子系统的评估结果向量(它们以及 \bar{u} 的维数等于参与评估的分割算法的个数)。X 表示在二级评估中的属性

矩阵(它以 $\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n$ 为行向量):

$$X = (\bar{u}_1, \bar{u}_2, \dots, \bar{u}_n)^T$$

\bar{W} 表示在二级评估中用于关联度分析的权重向量, \bar{u} 表示二级评估的最终评估结果向量。

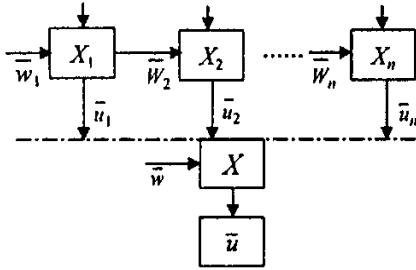


图2

3 指标设计及计算

出于对多层次灰色关联度分析的图像分割性能评估模型效果的检验,由试验分析拟选用面积因子、形状因子、重心距离、汉明距离、相似度如下:

(1) 面积因子:

$$m_A = T_c / T \quad (1)$$

T = 分割参照结果 RT 的面积; T_c = 对目标分割正确的区域 $ST \cap RT$ 的面积

(2) 形状因子:

$$m_F = (T_c / T)^{\frac{1}{3}} (1 - 2\sqrt{\pi} |\sqrt{T} / l_T - \sqrt{T_{ST}} / l_{ST}|) \quad (2)$$

即 m_F 应与正确分割的面积有关。

两幅图像重心间的距离,重心坐标按下式计算:

$$\bar{x} = m_{10} / m_{11}, \quad \bar{y} = m_{01} / m_{11}$$

对于数字图像 $f(x, y)$, 矩可以表示为:

$$m_{pq} = \sum_i \sum_j i^p j^q f(x, y) \quad p, q = 0, 1, 2, \dots$$

(3) 重心距离:

$$d(a, b) = |a_x - b_x| + |a_y - b_y| \quad (3)$$

(4) 两幅图像 f_1, f_2 的汉明距离 HIMDIST 由下式给出:

$$HIMDIST = \sum_y \sum_x |f_1(x, y) - f_2(x, y)| \quad (4)$$

HIMDIST 越小,表示两幅图像越相似, HIMDIST 越大越不相似。HIMDIST 等于0时,则两幅图像完全一致。

(5) 两幅图像 f_1, f_2 的相似度 S 由下式求出:

$$S = \frac{\sum_y \sum_x f_1(x, y) f_2(x, y)}{\sqrt{\sum_y \sum_x f_1^2(x, y) \sum_y \sum_x f_2^2(x, y)}} \quad (5)$$

相似度 S 越大,表示图像越相似, $S=1$ 时,两幅图像完全一致。

4 (单层次)灰色关联度分析评估

灰色关联度分析是一种以灰色关联度为测度衡量因素间关联程度大小的量化方法,是对系统统计数据列几何关系的一种比较。它对试验样本数量没有要求,不需样本具备典型的分布规律,计算量小。

设系统的 n 个待评估对象在 m 个评价因素下的评价指标属性值构成属性矩阵: $Y = (y_{ij})_{m \times n}$ 式中的 y_{ij} 代表第 j 个评估对象在第 i 个评价因素下的指标属性值。由于评估指标相互之间通常具有不同的量纲和数量级,不能直接进行比较,因此需要对原始指标值进行规格化处理(或称归一化处理)。同时,当数据的极差小于1时,归一化处理还可以增大数据的离散度。一般根据对象的属性指标有三种归一化方法:“越大越优型”(如本文中定义的面积因子、形状因子相似度等),“越小越优型”(如两幅图像的 HIMMING 距离、重心距离、虚警率等),“适当型”即“越接近于某一标准值 γ_i 越优”如亮度等^[7]。对其进行规格化处理的公式如下:

$$x_{ij} = \begin{cases} (y_{ij} - \min_i y_{ij}) / (\max_i y_{ij} - \min_i y_{ij}) & \text{越大越优型} \\ (\max_i y_{ij} - y_{ij}) / (\max_i y_{ij} - \min_i y_{ij}) & \text{越小越优型} \\ 1 - |y_{ij} - \gamma_i| / \max_i |y_{ij} - \gamma_i| & \text{适当型} \end{cases} \quad (6)$$

于是,得归一化矩阵 $X = (x_{ij})_{m \times n}$ 。显然, $0 \leq x_{ij} \leq 1$, x_{ij} 越大,表明第 j 个分割算法的第 i 个因素评价越优; x_{ij} 越小,表明第 j 个分割算法的第 i 个因素评价越次。

分别定义系统优向量 \bar{G} 和次向量 \bar{B} 如下:

$$\begin{cases} \bar{G} = (g_1, g_2, \dots, g_m) = (x_{11} \vee x_{12} \vee \dots \vee x_{1n}, x_{21} \vee x_{22} \vee \dots \vee x_{2n}, \dots, x_{m1} \vee x_{m2} \vee \dots \vee x_{mn}) \\ \bar{B} = (b_1, b_2, \dots, b_m) = (x_{11} \wedge x_{12} \wedge \dots \wedge x_{1n}, x_{21} \wedge x_{22} \wedge \dots \wedge x_{2n}, \dots, x_{m1} \wedge x_{m2} \wedge \dots \wedge x_{mn}) \end{cases}$$

其中 \vee, \wedge 分别为取大、取小运算符,又对第 j 个分割算法,记

$$\bar{X}_j = (x_{1j}, x_{2j}, \dots, x_{mj}), \quad j = 1, 2, \dots, n$$

灰关联分析:各个评价因素对待评系统的影响程度是不同的;通常要采用权重来表征各因素对待评系统的影响程度。至于权重通过层次分析(AHP)法来确定。记评价因素的权向量为

$$\bar{W} = (w_1, w_2, \dots, w_m), \quad \sum_{i=1}^m w_i = 1$$

对归一化矩阵 $(x_{ij})_{m \times n}$ 、第 j 个算法向量 \bar{X}_j 与优向量 \bar{G} 和次向量 \bar{B} 的关联系数为

$$\begin{cases} \xi_i(\bar{X}_j, \bar{G}) = \frac{\min_i \max_i |x_{ij} - g_i| + \rho \max_i \max_i |x_{ij} - g_i|}{|\min_i |x_{ij} - g_i| + \rho \max_i |x_{ij} - g_i|} \\ \xi_i(\bar{X}_j, \bar{B}) = \frac{\min_i \max_i |x_{ij} - b_i| + \rho \max_i \max_i |x_{ij} - b_i|}{|\min_i |x_{ij} - b_i| + \rho \max_i |x_{ij} - b_i|} \end{cases}$$

其中 ρ 为分辨系数, $\rho \in [0, 1]$, 一般取 $\rho = 0.5$, 第 j 个算法向量 \vec{X}_j 与优向量 \vec{G} 和次向量 \vec{B} 的关联度分别为

$$\gamma(\vec{X}_j, \vec{G}) = \sum_{i=1}^m w_i \xi_i(\vec{X}_j, \vec{G}) \text{ 和 } \gamma(\vec{X}_j, \vec{B}) = \sum_{i=1}^m w_i \xi_i(\vec{X}_j, \vec{B})$$

假设第 j 个算法向量 \vec{X}_j 以 u_j 从属于优向量 \vec{G} , 那么 \vec{X}_j 即以 $(1-u_j)$ 从属于次向量 \vec{B} . 根据经典最小二乘法思想, 提出目标函数如下

$$F(\vec{u}) = \sum_{j=1}^n \{ [(1-u_j) [\gamma(\vec{X}_j, \vec{G})]^2 + [u_j [\gamma(\vec{X}_j, \vec{B})]^2]] \}$$

为了求解系统的最优解向量 $\vec{u} = (u_1, u_2, \dots, u_n)$, (这时 \vec{X}_j 以 u_j 从属于优向量 \vec{G} 的关联程度最大), 由

$$\frac{\partial F(\vec{u})}{\partial u_j} = 0 \text{ 得}$$

$$u_j = \{1 + [\gamma(\vec{X}_j, \vec{B}) / \gamma(\vec{X}_j, \vec{G})]^4\}^{-1} \quad (6)$$

其中 $j=1, 2, \dots, n$

对模型的一些分析:

(1) 利用该模型得到了第 j 个备择算法向量 \vec{X}_j 从属于优向量的程度, 即反映了该算法的优次程度, 从而能根据 u_j 的大小对各个备择算法进行优次排序, 达到了评判的目的。

(2) 该模型考虑了两个具有优次标准的参数, 即 $\gamma(\vec{X}_j, \vec{G}), \gamma(\vec{X}_j, \vec{B})$, 更利于我们的排序决策, 一定程度上克服了仅考虑 $\gamma(\vec{X}_j, \vec{G})$, 当二备择算法在此参数下相差很小时的束手无策、难于决策的局面。该模

型的结果由于是非线性处理, 因而更具有可信度及精度。

(3) 该模型具有明显直观的数学物理意义:

a) 当 $\gamma(\vec{X}_j, \vec{G}) < \gamma(\vec{X}_j, \vec{B})$ 时, \vec{X}_j 从属于 \vec{G} 的程度 $u_j > 0.5$, 即知 \vec{X}_j 从属于优向量 \vec{G} 的程度较 \vec{X}_j 从属于次向量 \vec{B} 的程度大;

b) 当 $\gamma(\vec{X}_j, \vec{G}) > \gamma(\vec{X}_j, \vec{B})$ 时, 其数学物理意义与 a) 相反;

c) 当 $\gamma(\vec{X}_j, \vec{G}) = \gamma(\vec{X}_j, \vec{B})$ 时, \vec{X}_j 从属于优向量 \vec{G} 的程度等于 \vec{X}_j 从属于次向量 \vec{B} 的程度; 即该备择算法折中, $u_j = 0.5$;

d) 当 $\gamma(\vec{X}_j, \vec{B}) = 0$ 时, \vec{X}_j 与次向量 \vec{B} 的关联度最小, 处于极端, 该备择算法 \vec{X}_j 只能为优向量 \vec{G} , \vec{X}_j 从属于 \vec{G} 的程度当然为 1, $u_j = 1$;

e) 当 $\gamma(\vec{X}_j, \vec{G}) = 0$ 时, \vec{X}_j 与优向量 \vec{G} 的关联度最小, 也处于极端, 该备择算法 \vec{X}_j 最不令人满意, 它只能为次向量, \vec{X}_j 从属于 \vec{G} 的程度当然为 0, $u_j = 0$ 。

5 应用实例与分析

为了与其他评价方法做比较, 对分割算法匹配精度的性能评估按文[1]的测试图像集进行, 其中空格区域指 RT 域, 而阴影区域则表示某分割算法输出区域。图(a)是正确分割的目标(为参照图), 图(b)是分割后不连续的目标区域, 图(c)是分割后仅包含 RT 下半部且连续的目标区域, 图(d)是分割后仅包含 RT 上半部且连续的目标区域, 图(e)是分割连续的目标区域, 图(f)是分割完全不正确的目标区域。

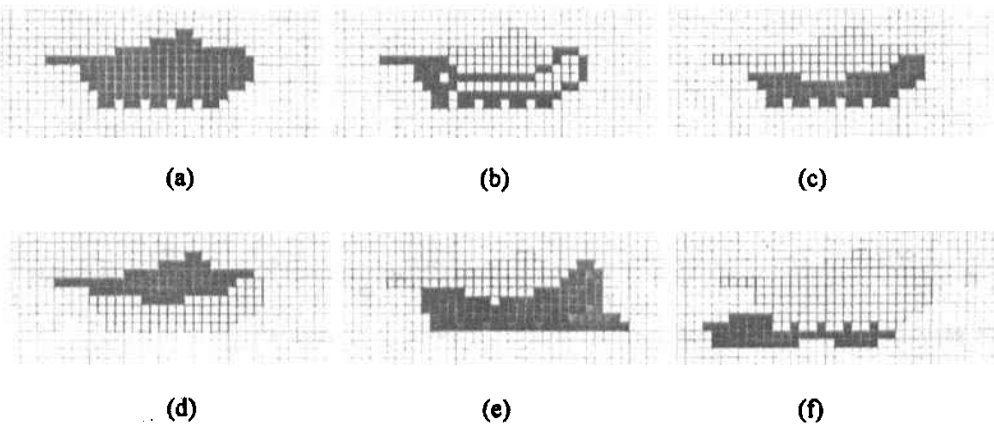


图3 图像分割测试集

(1) 多层次灰关联评估的第一步: 确定系统层次结构、指标体系; 可用图4表示。

面积因子、形状因子、相似度属于“越大越优型”的指标; HIMMING 距离、重心距离属于“越小越优

型”的指标。按归一化公式(6)进行规格化处理。

(2) 指标权重的确定用层次分析法和专家咨询结合来进行。1) 建立递阶层次结构的图像分割性能评估指标体系图4; 2) 确定各层次中因素的优劣次

序;3)计算权向量。在此为检验本方法在等权值时的 (0.5, 0.5)

抗失效性(不灵敏性),取:

$$\bar{w}_1 = (0.5, 0.5), \bar{w}_2 = (1/3 \ 1/3 \ 1/3), \bar{w} =$$

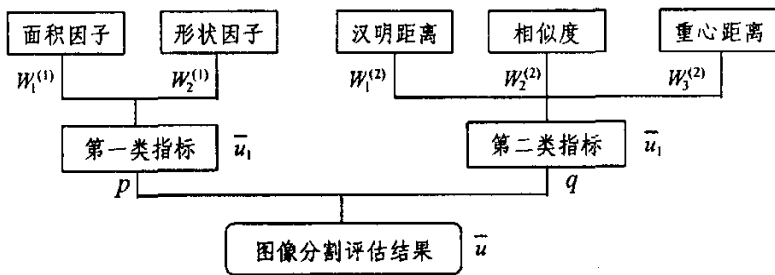


图4 图像分割性能评估指标体系图

(3)灰色关联度评估与分析:分三步:1)单层次灰色关联度分析评估;2)多层次灰色关联度分析评估;3)评估结果分析。

1.关于第一类因素指标:面积因子、形状因子的评判结果。

$$\bar{u}^{(1)} = (1.0, 0.7140, 0.7401, 0.7428, 0.8495, 0)$$

2.关于第二类因素指标:汉明距离、相似度、重心距离的评判结果。

$$\bar{u}^{(2)} = (1.0, 0.7151, 0.7641, 0.8020, 0.6999, 0)$$

3.分割性能的二级综合评估,评估结果。

$$\bar{u} = (u_a \ u_b \ u_c \ u_d \ u_e \ u_f) = (1.0, 0.8510, 0.8871, 0.9038, 0.9076, 0)$$

从第一类指标即面积因子、形状因子上分析, $u_a^{(1)} > u_c^{(1)} > u_d^{(1)} > u_e^{(1)} > u_b^{(1)} > u_f^{(1)}$ 。方法 e 的分割性能相对较好。

从第二类指标即汉明距离、相似度、重心距离上分析,即从两幅图像上对应点的匹配程度的角度来看,方法 d 的分割效果相对较好。

从两类指标的二级综合评估结果看,方法 e 的分割效果相对较好。可以看出本方法 MGAHP 有较大的适应性——它不仅可以从整体上进行综合评估,又可以根据需要从某个侧面进行特殊性需要的评估。这样就部分程度上兼顾了普遍性与特殊性的要求。

表1 分割性能评估的效果对比

	(a)	(b)	(c)	(d)	(e)	(f)
ERIM	1.0	0.5	0.5	0.5	0.5	0
TI	1.0	0.75	0.75	0.75	0.77	0.13335
Honeywell	0	0.0003744	0.0004918	0.0005028	0.0005994	0.0014
MF	1.0	0.13197	0.14391	0.146089	0.24947	0
MGAHP	1.0	0.85101	0.88713	0.90382	0.90763	0

小结 根据实际工作对评估方法的实用性的要求,提出了一个基于多层次灰色关联度分析的图像分割性能评估模型(MGAHP),并通过测试对其进行分析。该模型有如下特点:

(1)该模型 MGAHP 既避免了 TI 与 Honeywell 指标值中可能出现的奇异性,又克服了 ERIM 对分割性能的不灵敏性,大大增强了评判结果数据的离散性,提高了模型的辨析力,与其它方法相比,该方法计算量小,对样本的数量和分布规律没有要求。比其它方法更为合理有效。

(2)该评估模型在部分程度上对实际评估工作中的通用性与特殊性要求作了初步的尝试(譬如从整体性能上的综合评估,或从某个侧面特性上来进行评估,指标的多层次性要求等)。而以往方法尚未做到这一点。

(3)该评估模型的设计思想对整个 ATRS 的性能评估模型的设计有一定程度的启发意义。

致谢:本研究曾得到李介谷教授、杨杰教授的指导。

参考文献

- Bennett L. F. Knowledge-based evaluation of the segmentation component in automatic pattern recognition systems. Opt. Eng., 1991, 30(2): 154~165
- 邓雁萍,李介谷.图像分割的性能评估.模式识别与人工智能,1996,6:144~148
- Schneidewind N. F. Validating Metrics for Ensuring Space Shuttle Flight Software Quality. Computer, 1994 (8): 50~57
- Ng W. S., et al. Comment on Using the Uniformity Measure for Performance Measure in Image Segmentation. IEEE Trans. on PAMI, 1996, 18(9): 933~934
- 侯格贤,毕笃彦,等.图像分割质量评价方法研究.中国图形图像学报, 2000, 5(A): 39~43
- Cho K., et al. Performance Assessment Through Bootstrap. IEEE Trans. on PAMI, 1997, 19(11): 1185~1198
- 狄宇春.基于灰色系统分析的图像分割性能评估.[硕士学位论文文].上海:上海交通大学图书馆,1999