

基于区间映射规则的数字直方图处理

费风长, 方志军, 曾卫明, 章 琳

(江西财经大学信息管理学院, 南昌 330013)

摘 要: 在经典的单映射规则 SML 和组映射规则 GML 的基础上, 提出一种新的映射规则——区间映射规则(IML), 并提出基于 IML 的数字图像直方图规定化方法。以 8 阶灰度直方图为例, 通过对常用均衡化方法的实例分析, 引出基于 IML 的直方图均衡化方法——等区间映射均衡化(IIME)。用实验数据分析和证明了 IML 和 IIME 的优越性。

关键词: 直方图规定化; 直方图均衡化; 映射规则; 区间映射规则; 等区间映射均衡化

Histogram Processing of Digital Image Based on Interval Mapping Law

FEI Fengchang, FANG Zhijun, ZENG Weiming, ZHANG Lin

(School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330013)

【Abstract】 This paper presents a new grey-level mapping law for direct histogram specification, interval mapping law(IML). This law is based on the classical single mapping law (SML) and group mapping law (GML). By analyzing an example of the general equalization, IML draws forth the iso-interval mapping law (IIME), a method of direct histogram equalization. In the end, data analysis from the experiment proves the superiority of IML and IIME.

【Key words】 Histogram specification; Histogram equalization; Mapping law; Interval mapping law; Iso-interval mapping equalization

1 概述

灰度直方图是灰度级的函数, 是描述灰度图像中各灰度级的像素分布情况。直方图均衡化处理的基本思想是把原始直方图变换为均匀分布的形式^[1], 直方图的规定化思想则是为了满足实际需要而把直方图变换为我们所需要的形状。这两种直方图的处理方法都是图像空间域增强的方法, 是多种空间域处理技术的基础^[3], 在数字图像处理中处于十分重要的地位。

本文为了讨论方便, 先把会用到的符号与公式统一约定如下^[1,2,4,5]:

n ——原始图像的像素总数;
 N ——原始图像的灰度级个数;
 r_i ——原始图像灰度级, $i=0,1,\dots,N-1$;
 $n(r_i)$ ——原始图像灰度级为 r_i 的像素个数;
 $p_r(r_i)$ ——原始直方图;
 $P_r(r_i)$ ——原始累计直方图;
 M ——规定直方图的灰度级个数;
 u_j ——规定直方图的灰度级, $j=0,1,\dots,M-1$;
 $p_u(u_j)$ ——规定直方图;
 $P_u(u_j)$ ——规定累计直方图;
 s_i ——原始图像中灰度级为 r_i 的像素经映射规则处理后灰度级被映射为 s_i , $i=0,1,\dots,N-1$;
 $p_s(s_i)$ ——映射后的直方图;
 $r_i \rightarrow s_i$ ——灰度级 r_i 按一定规则映射为 s_i ;
' \rightarrow '——指映射为;
 z_j ——经直方图处理后得到的目标灰度级, $j=0,1,\dots,M-1$;
 $p_z(z_j)$ ——目标直方图。

并有相应的公式^[1,2,4,5]:

$$p_r(r_i) = n(r_i) / n \quad (1)$$

$$P_r(r_i) = \sum_{k=0}^i p_r(r_k) \quad (2)$$

$$p_u(u_j) = \sum_{k=0}^j p_u(u_k) \quad (3)$$

$$p_z(z_j) = \sum_{k=a}^b p_r(r_k) \quad (4)$$

$$a, b \in i, a \leq b \text{ 且 } r_a \rightarrow s_a, r_b \rightarrow s_b \text{ 且 } s_a = s_b = z_j$$

常用的均衡化取整扩展公式^[1]:

$$s_i = \text{int}[(N-1)P_r(r_i) + 0.5] \quad (5)$$

常用均衡化过程可以分为 3 个步骤:(1)计算原始累计直方图 $P_r(r_i)$;(2)用式(5)进行取整扩展, 得到相应的灰度级 s_i ;(3)求目标灰度级 z_j 和目标直方图 $p_z(z_j)$ ^[1]。

直方图的规定一般分为 4 个步骤:(1)计算原始累计直方图 $P_r(r_i)$;(2)计算规定累计直方图 $P_u(u_j)$;(3)将原始累积直方图 $P_r(r_i)$ 根据相应的映射法则映射到规定累积直方图 $P_u(u_j)$, 求出 s_i , 一般采用的映射方法有单映射规则 SML(Single mapping law)和组映射规则 GML(Group mapping law)^[4];(4)求目标灰度级 z_j 和目标直方图 $p_z(z_j)$ ^[1,2]。

直方图的均衡化其实是一种特殊的直方图规定化, 特殊

基金项目: 国家自然科学基金资助项目(60462003); 江西省教育厅科技研究基金资助项目(赣教技字[2005]115); 江西财经大学创新基金资助项目

作者简介: 费风长(1979-), 男, 硕士生, 主研方向: 图像通信与数字电视及多媒体技术; 方志军、曾卫明, 副教授; 章 琳, 硕士生

收稿日期: 2006-01-25 **E-mail:** ffcbox@yahoo.com.cn

在于其规定直方图为一均匀分布^[2]。一个好的规定化方法可以推出一个好的均衡化方法，又由于规定化的好坏是由映射规则来决定的，由此知映射规则对直方图的规定化和均衡化的影响是十分大的。本文将提出一个不同于以往的映射规则——区间映射规则，并基于该规则推导出相应的直方图规定化方法和均衡化方法。

2 区间映射规则

正是由于映射规则的重要性，直方图在规范化过程中第(3)步选取什么样的映射规则求 s_i 才显得十分关键，在这步中我们常用SML或GML来得到 s_i 。这两种映射规则都是通过原始累积直方图 $P_r(r_i)$ 和规定累积直方图 $P_u(u_j)$ 的点对点运算，来寻找差值最小的点，最后得到 s_i 。它们不同之处在于：单映射规则是从 $P_r(r_i)$ 向 $P_u(u_j)$ 搜索对应点，而组映射规则是从 $P_u(u_j)$ 向 $P_r(r_i)$ 搜索对应点^[2]。笔者基于实践提出一种点到区间的映射规则，它是把原始累积直方图 $P_r(r_i)$ 映射到规定累积直方图区间来得到 s_i 。

区间映射规则如下：

寻找满足下面公式的 k 和 l ：

$$\begin{cases} \sum_{j=0}^{l-1} p_u(u_j) < \sum_{i=0}^k p_r(r_i) \leq \sum_{j=0}^l p_u(u_j) & k=0, 1, \dots, N-1 \\ 0 \leq \sum_{i=0}^k p_r(r_i) \leq p_u(u_l) & l=1, 2, \dots, M-1 \end{cases} \quad (6)$$

然后把 $p_r(r_i)$ 映射到对应的 $p_u(u_j)$ 去。从公式可以看出 $p_r(r_i)$ 是投影到一个区间范围内，所以称为区间映射规则IML(Interval mapping law)。

式(6)的算法如下：

```
for i=0: N - 1
if 0 <= Pr(ri) <= Pu(u0)
    si = u0;
else
    for j=1: M - 1
    if Pu(uj-1) < Pr(ri) <= Pu(uj)
        si = uj;
    end
end
end
```

把 $P_r(r_i)$ 看作一个区域，映射规则的实质就是如何划分 $P_r(r_i)$ ，SML和GML是通过 $P_r(r_i)$ 与 $P_u(u_j)$ 点对点的运算来划分 $P_r(r_i)$ ，而IML是以 z_j 出现的概率来划分 $P_r(r_i)$ 。基于IML是可以得出相应的直方图规定化方法。

3 基于区间映射规则的直方图均衡化方法

3.1 常用均衡化方法

在常用均衡化步骤中，式(5)的取整扩展是均衡化的关键，下面以一个8阶灰度直方图的例子来分析常用均衡化方法。为了讨论的方便，本文会把常用均衡化简称为式(5)均衡化。

给出一个均匀分布的直方图 $p_r(r_i)=1/8$ ， $i=0,1,\dots,7$ ，如图1。把它代入式(5)进行均衡化，得到的目标直方图(图2)，其中 $p(z)=0.125$ 的水平直线为理想均衡化结果^[1]。我们发现本来是均匀分布的直方图经过式(5)均衡化后反而不再是均匀分布了。

为何出现这样的结果呢？下面用IML的思想来分析整个均衡化过程，把 $P_r(r_7)$ 看作一个大小为1的条形区间； $s_i=0,1,\dots,7$ 代入式(5)求得相应的 $P_r(r_i)$ ，可以得到满足 $r_i \leq s_i$ 时 s_i 所对应的 $P_r(r_i)$ 区间。又 s_i 不会出现有相等值的情况，根据式(4)如果 $s_i = z_j$ 那么 $p_r(r_i) = p_s(s_i) = p_z(z_j)$ ， s_i 对应的 $P_r(r_i)$ 区间也

就是 z_j 对应的 $P_r(r_i)$ 区间，把 z_j 对应的 $P_r(r_i)$ 区间大小记为 Q_j ，也可以说是 z_j 出现的概率，于是构造出图3。在图3中发现：(1) z_j 对应 $P_r(r_i)$ 的区间大小 Q_j 不全一致， $Q_1=Q_2=Q_3=Q_4=Q_5=Q_6=1/7$ ， $Q_0=Q_7=1/14$ ；(2)原始图像灰度级为 r_i 的像素通过式(5)后灰度级被映射为 s_i ：0 1, 1 2, 2 3, 3 4, 4 4, 5 5, 6 6, 7 7，灰度级个数由8个减少到7个，即 z_j 只有7个。通过分析，原因是： z_j 对应的 Q_j 不一致，即 z_j 出现的概率不一致，目标灰度级 z_j 中有些灰度级容易出现而有些灰度级不容易出现，也就是说直方图经式(5)均衡化后得到的目标直方图会出现：直方图向容易出现的 z_j 偏向的问题，这随之可能导致灰度级个数的减少。在图3中的表现就是灰度级为 r_i 的像素经过式(5)均衡化后，最终得到 $z_j = 1, 2, 3, 4, 5, 6$ 的机会要多些，而得到 $z_j = 0, 7$ 的机会要小些。最终的结果就是使得原本是均匀分布的直方图经过式(5)均衡化后反而不再是均匀分布的了。

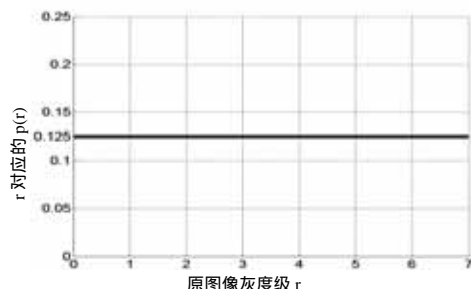


图1 均匀分布的原始直方图

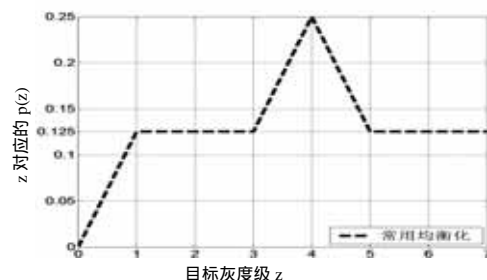


图2 均匀分布的直方图经常用均衡化得到的目标直方图

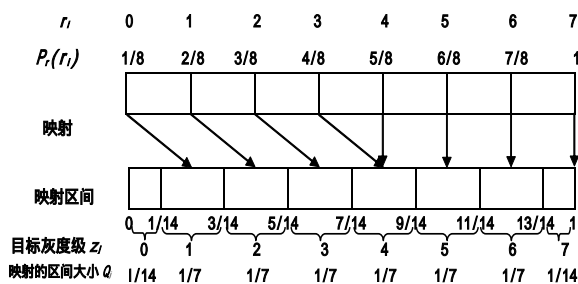


图3 用IML分析均匀分布直方图经常用均衡化过程

3.2 等区间映射均衡化方法

上面用IML的思想解释了为什么均匀分布的直方图经式(5)均衡化后，反而得到了不均匀分布的目标直方图的原因，又由于均衡化是一种特殊的直方图规定化^[2]，于是本文提出了基于IML的直方图均衡化方法——等区间映射均衡化IIME(iso-interval mapping equalization)，简称等区间均衡化。IIME过程如下：

- (1)计算原始累积直方图 $P_r(r_i)$ ；
- (2)计算规定累积直方图 $P_u(u_j)$ ；
- (3)将 $P_r(r_i)$ 根据IML映射到 $P_u(u_j)$ 区间内，得到规定后的灰

度级 s_i ;

(4)求目标灰度级 z_j 和目标直方图 $p_z(z_j)$ 。

这里的规定直方图为 $p_u(u_j) = 1/M$ ，所以第(3)步的映射算法可以表示如下：

```

for i=0: N - 1
if  $P_r(r_i)=0$ 
 $s_i = u_0$ ;
else
for j=0: M - 1
if  $j/M < P_r(r_i) \leq (j+1)/M$ 
 $s_i = u_j$ ;
end
end
end

```

等区间均衡化的特点是：(1)基于区间映射规则；(2)规定函数为一均匀分布。

现在以 8 阶灰度直方图为例，同样给出一个均匀分布的直方图 $p_r(r_i)=1/8, i=0,1,\dots,7$ 来分析其经过IIME的过程。由于是 8 阶灰度直方图，根据式 (5)和IIME过程得到图 4。由给出的条件可以求出 z_j 对应 $P_r(r_i)$ 的区间大小 Q_j 分别为 $Q_0=Q_1=Q_2=Q_3=Q_4=Q_5=Q_6=Q_7=1/8$ ， Q_j 是完全相等的，那么就会使得每个 z_j 出现的机会都是一样的，那么均衡化后的目标直方图就不会出现向某些灰度级偏向的问题。图 4 显示出了原始图像灰度级为 r_i 的像素经过IIME后灰度级被映射为 s_i ：0 0, 1 1, 2 2, 3 3, 4 4, 5 5, 6 6, 7 7。可以看出，一个均匀分布的直方图经过IIME后，目标直方图仍然是一个均匀分布，这就克服了式(5)均衡化中目标灰度级偏向的问题。

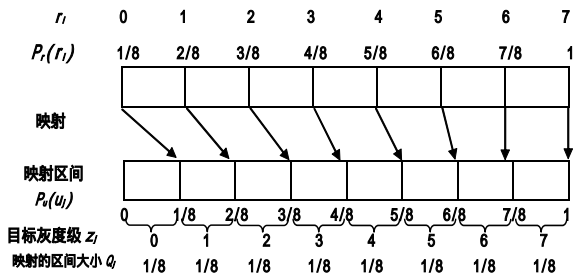


图 4 概率均匀分布直方图经等区间均衡化过程

4 实验与总结

下面给出几个直方图，分别经过常用均衡化和等区间均衡化 IIME 处理，通过分析它们的目标直方图来比较常用均衡化方法和 IIME 方法。给出的直方图如下： $p_0=[1/8 \ 1/8 \ 1/8 \ 1/8 \ 1/8 \ 1/8 \ 1/8 \ 1/8]$ ， $p_1=[1/14 \ 1/7 \ 1/7 \ 1/7 \ 1/7 \ 1/7 \ 1/14]$ ， $p_2=[0.19 \ 0.25 \ 0.21 \ 0.16 \ 0.08 \ 0.06 \ 0.03 \ 0.02]$ ， $p_3=[0.5 \ 0 \ 0.2 \ 0.04 \ 0.07 \ 0.14 \ 0.02 \ 0.03]$ 。以上的直方图经过两种均衡化后得到的目标直方图如图 5~图 8。图中实曲线为经过 IIME 得到的目标直方图曲线，虚曲线为经常用均衡化得到的目标直方图曲线。

由于现在还缺少对直方图变换效果进行评价的通用标准^[2]，因此只有通过经验参数来评价均衡化方法的优劣。本文得到的目标直方图都是基于以下几个方面来比较的：

(1)尽量使灰度级 z_j 对应的 $p_z(z_j)$ 都相等，即均衡化后的直方图曲线应该是一条趋于水平的直线（理想均衡化结果）。本文用绝对误差AE(Absolute Error)^[1]作为一个评定参量。

$$AE = \sum_{j=0}^{M-1} |p_z(z_j) - p_u(u_j)| \quad (7)$$

(2)尽量保留细节，直方图均衡化后的灰度级要尽可能多，覆盖的灰度级也要尽可能多，简并的灰度级要尽量少。

即均衡化后的直方图曲线要尽可能地拉伸^[1]。

(3)一个理想的均匀分布直方图（直方图曲线为一水平直线）通过均衡化后，其直方图应尽可能不变，也应是一条水平直线。

可以把图 5~图 8 中 AE 和简并的灰度级个数构造数据比较，如表 1。

在图 5 中，常用均衡简并了一个灰度级，IIME 没有简并灰度级。均匀分布的直方图经 IIME 后不发生改变，而经过常用均衡化方法后却成为一条曲线。

在图 6 中，通过比较可以看出IIME的目标直方图要好。但是如果分析公式 $s_i=\text{int}[(N-1)P_r(r_i)+0.5]$ 就会发现，原灰度级 $[0,1,2,3,4,5,6,7]$ 经过式(5)均衡化后应被映射为 $[1,2,3,4,5,6,7,7]$ ， $p_z(z_0)=0$ ， $p_z(z_1)=1/14$ ， $p_z(z_2)=p_z(z_3)=p_z(z_4)=p_z(z_5)=p_z(z_6)=1/7$ ， $p_z(z_7)=3/14$ ， $AE1=0.3571$ ，只简并一个灰度级，目标直方图应该是相当不错的。但是图 6 的结果却是实际目标直方图与我们理论分析的结果是截然相反。这主要是因为计算机中 $1/14$ 和 $1/7$ 会被计算为双精度值 0.071429 和 0.142857 ，这时产生了极小的误差，再经过概率的累加又会产生累加误差，这时的误差相对而言仍然是很小的，但是经过取整后，很小的误差却最终导致最后的结果与理论分析结果完全不同。如 $P_r(r_2)=p_r(r_0)+p_r(r_1)+p_r(r_2)=2.99997$ ，经过程序计算后得 $s_2=2$ ，但理论计算上 s_2 应该为 3。这就解释了为什么 $p_z(z_2)$ 、 $p_z(z_5)$ 会是一般 $p_z(z_j)$ 的 2 倍，而 $p_z(z_4)$ 却为 0。

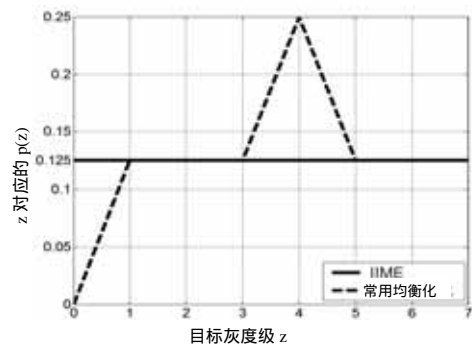


图 5 直方图 p0 经两种均衡化后得到的目标直方图

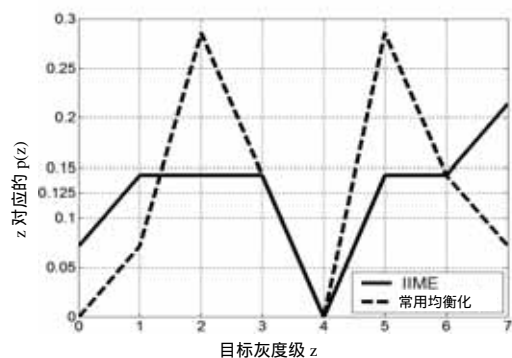


图 6 直方图 p1 经两种均衡化后得到的目标直方图

在图 7 中，两条曲线只有最后两个灰度级的 $p_z(z_j)$ 不同，常用均衡化得到的目标直方图曲线最后两个灰度级不能形成比较水平的直线，而经过IIME得到的目标直方图曲线最后两个灰度级能形成比较好的水平直线。

在图 8 中，绝对误差是一样的，并且常用均衡与 IIME 简并的灰度级一样都是 4 个。虽然各种定量的比较都一样，但可以看到 IIME 得到的目标直方图曲线覆盖的灰度级要多些，这样可以增强对比度，增加图像的视觉效果。

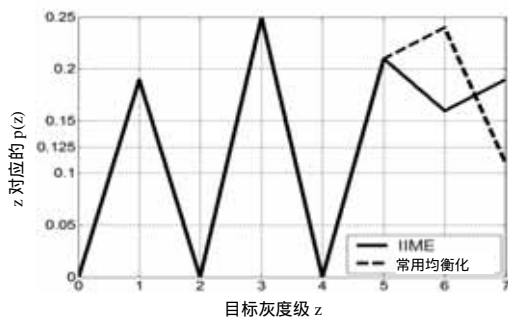


图 7 直方图 p2 经两种均衡化后得到的目标直方图

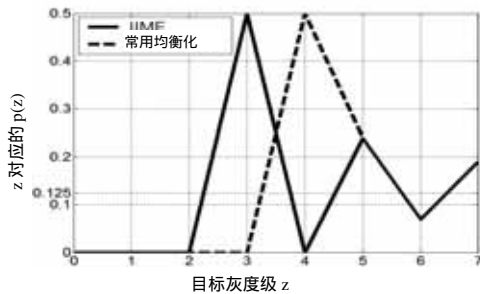


图 8 直方图 p3 经两种均衡化后得到的目标直方图

表 1 常用均衡化与 IIME 的目标直方图数据比较

	P0		P1		P2		P3	
	AE	简并的灰度级个数	AE	简并的灰度级个数	AE	简并的灰度级个数	AE	简并的灰度级个数
常用均衡化方法	0.250 0	1 个	0.714 3	2 个	0.780 0	3 个	1.110 0	4 个
IIME	0	0 个	0.357 1	1 个	0.750 0	3 个	1.110 0	4 个

从以上可以总结 :区间映射规则(IML)是基于经典的SML

和GML的思想发展而来的。IML的思想,即以 z_j 出现的概率来划分 $P_r(r_i)$ 的思想可以十分方便地解释直方图变换中的一些问题,本文后来就是用IML的思想发现了常用直方图均衡化方法中的不足——灰度级偏向的问题。为解决这个偏向问题,于是提出了基于IML的均衡化方法——IIME。通过实验发现在处理 8 阶灰度直方图时,IIME方法产生的结果图比常用均衡化方法产生的结果图要好。本文的均衡方法都是基于 8 阶灰度直方图来考虑的,至于更多灰度级的直方图处理,大家可以通过等区间均衡化方法的定义来推导,本文就不再一一详述了。

参考文献

- 1 章毓晋. 图像工程(上册): 图像处理和分[M]. 北京: 清华大学出版社, 1999.
- 2 章毓晋. 数字图像直方图处理中的映射规则——评“用于数字图像直方图处理的一种二值映射规则”一文[J]. 中国图像图形学报, 2004, 9(10): 1265-1268.
- 3 Rafael C G, Richard E W. Digital Image Processing (Second Edition) [M]. 北京: 电子工业出版社, 2004.
- 4 Zhang Y J. Improving the Accuracy of Direct Histogram Specification[J]. Electronics Letters, 1992, 28(3): 213-214.
- 5 Yang X, Xiao Q, Raafat H. Direct Mapping Between Histograms: An Improved Interactive Image Enhancement Method[C]. Proc. of IEEE Int. Conf. on Systems, Man and Cybernetics, Charlottesville, Virginia, 1991: 243-247.

(上接第 192 页)

便能解决时序的问题,引入时间因子 p ,用 p 来表示过去的范例和解决方案与当前问题的紧密程度, p 的值介于 0 和 1 之间,具体的方法是:把时间划分为若干时间段(不同的特征范畴依据不同的情况有不同的划分),每段上的 p 值不一样。离当前时间越久的范例重要性越低,它的 p 值也越小,离当前时间越近的范例重要性越高,它的 p 值也越大。系统建立守候进程,每隔一定时间进行 p 值的更新赋值操作。表 2 描述了一个 COSCBR 范例表示的例子。

2.4 COSCBR 相似度的计算

假设范例 X 的属性集: $X_a = \{A_1, A_2, \dots, A_n\}$,用 $val(A_i)$ 表示 A_i 的值域,对于任意已有范例 X_i ,设 $WX_i = \{wx_{i1}, wx_{i2}, \dots, wx_{im}\}$ 为其权向量,局部相似度用 $sim(A_1, A_2)$ 表示,时序因子用 p 表示。则新范例 P 与范例库中的范例 Q 的全局相似度 s 算法如下:

```

begin
  let s=0
  loop for any  $A_i^p$ 
    if exit  $A_j^q$  and  $A_i^p = A_j^q$ 
      then  $s = s + w_j^q * sim(val(A_j^q), val(A_i^p))$ 
    endloop
  let  $s = s * p$ 
endbegin

```

全局相似度 s 的值在 $[0, 1]$ 内,表示两个范例相似程度,当 $s = 0$ 时表示完全不相似, $s = 1$ 时表示完全相似。

3 结束语

COSCBR 模型构建了软件成本估算的框架,是一种解决实际问题的新途径的探索,其估算的准确程度依赖于软件成本属性数据和范例数据的提取。而范例数据提取是一个不断持续完善的过程,需要继续的实例验证和分析。软件成本估算对于软件开发者和用户都是相关的,这方面的研究已经在国内外逐渐形成新的研究领域,中国在该领域的研究处于起步阶段。软件成本估算研究的积累与成熟将对软件行业发展的规范化和经验积累起到重要作用。

参考文献

- 1 Boehm B W. Software Engineering Economics[M]. Prentice Hall, 1981.
- 2 Jorgensen M, Sjoberg D I K. Impact of Effort on Software Project Work[J]. Information and Software Technology, 2001, 43(15): 939.
- 3 Boehm B W. 李师贤译. 软件成本估算: COCOMO 模型方法[M]. 北京: 机械工业出版社, 2005-04.
- 4 Boehm B W. 李师贤译. 软件工程经济学[M]. 北京: 机械工业出版社, 2004-07.
- 5 王 平, 丁浩芳, 李 韬. 结构型软件成本估算模型的研究与改进[J]. 计算机工程, 2002, 28(12): 88-89.

