

一种基于降维滤波的快速去雾算法

刘 杰, 谢正祥², 吕霞付¹, 陈学渊¹

(1 重庆邮电大学 工业物联网与网络化控制教育部重点实验室, 重庆 400065;

2 重庆医科大学 生物医学工程研究室, 重庆 400016)

摘 要: 为了达到单幅图像的快速去雾, 提出了一种基于降维滤波的快速去雾算法. 采用边界约束来求取透射率图, 然后用降维滤波对透射率图进行优化, 减少了计算量; 通过自适应调节大气光值和透射率下限值优化了天空区域的失真; 对去雾后的图像进行自适应 Zadeh-X 变换改善图像的视觉效果. 通过对多种带雾图像实验表明该算法可以达到高速和高质量的去雾, 并且适用于偏暗、偏亮和颜色不丰富的带雾场景.

关键词: 去雾; 透射率; 降维滤波; 自适应调节

中图分类号: TP317.4

文献标识码: A

文章编号: 1000-7180(2015)04-0044-06

A Fast Dehazing Algorithm Based on Dimension Reduction Filter

LIU Jie¹, XIE Zheng-xiang², LV Xia-fu¹, CHEN Xue-yuan¹

(1 Key Laboratory of Industrial Internet of Things & Network Control, MOE, Chongqing

University of Posts and Telecommunications, Chongqing 400065, China; 2 Biomedical

Engineering Lab, Chongqing Medical University, Chongqing 400016, China)

Abstract: In order to remove fog for single image fleetly, a defogging method based on dimension reduction filter is proposed in this paper. It uses boundary constraint to achieve transmission map, and then optimize the transmission map with dimension reduction filter. The complex computation is reduced considerably. Distortion of the image in the sky area is compensated by adaptively adjusting the atmospheric optical value and lower limit of transmission map. Then the visual effect of optimized image is improved by applying an adaptive Zadeh-X transformation. Experimental results on a variety of haze images demonstrate real-time performance and high-quality of the proposed algorithm, and it is applicable to foggy scenarios despite illumination is too low or too high and the color information is not rich enough

Key words: defogging; transmission; dimension reduction filter; adaptive control

1 引言

通常在带雾环境下获得的图像, 经常遭遇较差的能见度. 雾中远处的对象丢失了对比度变得模糊, 在图 1 中明显可以看到. 这是因为场景中对象的反射光到达摄像机的时候, 它受到了空气中悬浮微粒的散射和吸收而被衰减, 也是由于这个原因, 对象的颜色也会出现偏色, 变得与雾相似, 并且距相机的距离越远相似度越高.

早期的去雾算法主要依赖附加的深度信息或同一个场景的多个观察. 比如 Narasimhan 提出了一个基于散射物理模型的方法^[1]. 通过这个模型, 这个场景结构能通过同一场景两个或多个不同环境下的图像被恢复. Kopf^[2]提出了使用直接从数字地形和城市模型中获得场景深度信息用于去雾的方法.

单幅图像去雾的方法将更具挑战性, 因为只有很少的场景结构信息被获得. 近年来这些方法有了重大的进步, 如文献^[3-10]所述. 这些进步很大程度

收稿日期: 2014-06-20; 修回日期: 2014-07-22

基金项目: 国家自然科学基金项目(60975008)



图1 带雾原图

上得益于在新的图像模型上比较深入的研究或一些先验知识. 如 Fattal^[3] 假设这个透射率函数和表面阴影函数是局部不相关的, 所得算法的速度较快, 但由于这个方法提供了一个物理恒定的估计, 它不能用到这个两个函数变化不大的区域. Tan^[4] 通过最大化局部对比度来增强图像的能见度. 它的方法, 可以产生非常完好的结果, 尤其是在浓雾区域. 由于它不是基于物理模型的, 所以经常要遭遇偏色和明显的光晕. 2013 年 Kim^[5] 提出了另外一种优化对比度的方法虽然速度上相对前者有了大大提高但是有同样的缺陷. He^[6] 提出了一个非常有趣的图像先验(暗通道先验)来完成单幅图像去雾, 这个先验来自对无雾环境下图像的观察, 大部分局部块内经常存在低亮度像素, 这种方法在视觉效果上比前者有了很大提升, 但是由于这个先验的局限性导致对偏暗和颜色不丰富环境下的图像处理效果不佳并且处理速度慢. 2012 年 He^[7] 采用引导滤波的方法在代替了原有的 soft matting^[6] 优化透射率图, 速度得到了很大的提高. Tarel^[8] 提出了一种通过大气面纱值来推导透射率图的方法, 它们对彩色图像 3 通道的最小值进行快速的中值滤波得到大气面纱, 这个最大的优点就是它的处理速度有了进一步的提升, 可以进一步应用到实时处理, 但是去雾效果不完全. Gibson^[9-10] 对它进一步改进, 在处理速度和效果上都有很大的提高. 实验对比分析中, 本文选取文献[7]和文献[10]提出的 2 种快速有效的方法进行了对比.

本文提出了一个新的求透射率图、透射率图优化、估计大气光值和透射率下限值的方法, 然后增加了一个快速亮度自适应调节的方法, 每个方法相对现有的领先的方法速度和效果上都有较大的提高,

很好地弥补了现有的去雾算法对颜色偏暗和颜色不丰富的场景去雾效果不佳的特点.

2 大气散射模型

由于空气中微粒的散射和吸收, 造成接收到的图形会出现明显的衰减, 根据 He^[6] 中的大气散射模型对带雾图像进行研究, 它的表达式如式(1)所示,

$$I(x, y) = t \cdot J + (1 - t) \cdot A \quad (1)$$

式中, I 表示摄像机采集到的带雾图像, J 是去雾之后的图像, A 表示大气光值, t 表示场景的透射率, 它与场景的深度是相关的, 它的表达式如下:

$$t = e^{-\beta \cdot d(x)} \quad (2)$$

式中, β 表示衰减系数, d 表示景深.

图像去雾就是根据式(1) 求出 J , 它的表达式如下:

$$J = \frac{I(x, y) - A}{\max(t, t_0)} + A \quad (3)$$

式中, t_0 表示一个初值, 由于它的存在可以保证式(3) 中分母不为 0, 同时也用于微调去雾效果. 所以, 只要已知每一点透射率 t 、大气光值 A 和透射率下限值 t_0 就可以求出去雾后的图像 J .

3 新的快速去雾算法

3.1 边界约束求透射率图

根据式(1)可知, 由于受到雾的影响, 采集到的图像会随着大气光值变化. 如图 1 所示.

根据式(1) 我们可以得到关于 $t(x)$ 的表达式如下:

$$t(x) = \frac{I(x) - A}{J(x) - A} \quad (4)$$

考虑到自然无雾的图像 $J(x)$ 是有一个范围限制^[11],

$$C_1 < J(x) < C_2, \forall x \in \Omega \quad (5)$$

式中, Ω 表示整幅图像, C_1 和 C_2 表示与输入的 3 通道图像相关的 2 个矢量, 它表示的是 J 的边界约束. 假设这个大气光值 A 已给定, 根据式(4) 和式(5) 我们可以得出 $t(x)$ 的低的边界值 $t_b(x)$, 它满足以下条件.

$$0 < t_b(x) < t(x) < 1 \quad (6)$$

式中,

$$t_b(x) = \min\left\{\max_{c \in \{r, g, b\}} \left(\frac{A^c - I^c(x)}{A^c - C_1^c}, \frac{A^c - I^c(x)}{A^c - C_2^c}\right), 1\right\} \quad (7)$$

式中, c 表示 3 个颜色通道. 式(6) 的边界约束条件提供了一个新的求透射率图的方法. 经典的暗通道求透射率的方法^[6] 只是它的特例, 假设 $C_0 = 0$, A 的值比带雾

图像的所有亮度值都大时,首先,根据式(1)直接计算 t_b ,假设在一个局部块内,这个透射率 t 是恒定的,然后在局部块内对 t_b 进行最大值滤波,最后快速导出通过暗通道求得的基于块的透射率图 $\tilde{t}$ 。

这个边界约束条件更为通用。在很多情况下,这个优化的全局大气光值 A ,比整幅图像中最亮的值要小,这些亮的像素经常是来自于亮的天空和白色的物体,在这些情况下,暗通道理论将会失效,然而本文的算法依然可行。另外,在一个局部区域内使用一个恒定的假设有点苛求,而在块区域内估计的透射率图往往是低于估计值。为了求出一个更加精确的透射率图,我们放松了这个假设,允许在局部区域内透射率图略微不同,这个新的基于块的透射率图 $\tilde{t}$ 可以通过对 t_b 进行一个形态学的闭合运算得到,如式(8)所示,

$$\tilde{t} = \text{imclose}(t_b, M) \quad (8)$$

式中, imclose 表示形态学中的闭合处理, M 表示一个正方形的模板。图2显示了与直接使用基于块的透射率图进行去雾处理效果比较,可以看出边界约束的方法有更少的光晕和更强的对比度。

基于块的透射率图,采用局部块内有相同的景深假设,经过去雾处理之后易在深度突变的边缘产生光晕现象。因此,需对透射率图进一步优化处理。

3.2 降维滤波优化透射率图

对比透射率图的优化方法,He^[6]采用了 soft matting 的方法,因为它在原理上遵循透射率图 t (alpha 通道)与输入图像是成线性关系的假设,这意味着 t 的小变化,得到的 J 将变化很大。比较透射率图发现它的透射率图更加丰富,主要是由于表面阴影和反射的改变,由于这个 t 的变化导致了在 J 上不想要的结果,明显对比度会减小。因为理想的透射率图可以更好地代表图像的深度信息,而在场景内这个深度的变化是分段平滑和缺乏细节的,这个细节是不需要的,因此需要去除它来得到一个更加精确的估计,否则最后恢复出来的图像会出现偏色。因此,要对求出的透射率图进一步边缘保持平滑处理,本文采用了一个快速降维滤波来实现。

一个边缘保持滤波可以用多维来解释,对一个2D的灰度图像可以在 R^3 空间中表示,这里 R 的指数表示空间的维数,这个 R^3 空间中单个点的坐标表示如下:

$$\hat{p} = (x_p, y_p, r_p) \quad (9)$$

式中, $\hat{p}$ 对应灰度图像 I 的空间坐标 $p = (x_p, y_p)$ 和

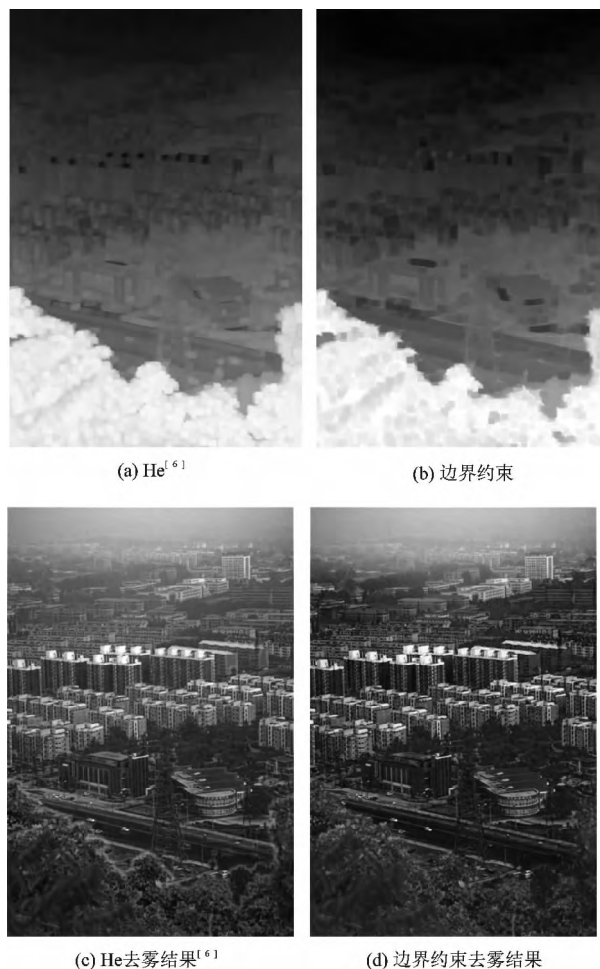


图2 基于块的透射率图和去雾结果对比

像素值为 $I(p)$ 的点。设 $F(\hat{p}, \hat{q})$ 为边缘滤波核,而 Z 为 I 经过 F 滤波之后的结果,表示如下:

$$Z(p) = \int_{\Omega} I(q) \cdot F(\hat{p}, \hat{q}) dq \quad (10)$$

例如双边滤波,其双边滤波核为

$$F(\hat{p}, \hat{q}) = G_{\sigma_s}(p - q) \cdot G_{\sigma_r}(r - r_q) \quad (11)$$

式中, G_{σ_s} 和 G_{σ_r} 分别表示典型的空间域和值域的高斯滤波。因为双边滤波工作在 R^3 空间,所以在实际应用中会出现处理速度过慢,由此我们考虑降维提速,提出把滤波核 F 从 R^3 降低到 R 的变换 T ,它满足如下式子:

$$\begin{aligned} Z(P) &= \int_{\Omega} I(q) \cdot F(\hat{p}, \hat{q}) dq \\ &= \int_{\Omega} I(q) \cdot H(T(\hat{p}), T(\hat{q})) dq \end{aligned} \quad (12)$$

式中, H 是一个 R 滤波核,同时 H 也具有边缘保持滤波的特性。由于在执行边缘保持滤波的时候2个像素之间的混合强度与它们之间的距离成反比,因为该降维变换 T 保持了原始 R^3 空间的像素距离,所

以 H 在 R 空间也保持边缘保持滤波特性。

为了寻找一个等距离的降维变换 T , 首先定义 I 为一个 1D 信号, 它是一个 R^2 空间内曲线 C , 它的横坐标为 x , 纵坐标为 $I(x)$, 而 $x \in \Omega = [0, +\infty)$, 我们的目的是找到一个 R^2 到 R 的变换 T 使 $|T(x_i, I(x_i)) - T(x_j, I(x_j))| = ||(x_i, I(x_i)) - (x_j, I(x_j))||$, 其中 $|\cdot|$ 表示求绝对值, $||\cdot||$ 表示的是某个距离的度量, 为了简化运算使用 2 个邻域像素 L_1 范数距离。假设 $x_{i+1} - x_i = w$, 为了保持等距性, 必须满足下面的式子:

$$T(x+w) - T(x) = w + |I(x+w) - I(x)| \quad (13)$$

上式表示在变换后的 R 空间里的相邻 2 点的欧几里得距离等于原来的 R^2 空间的 L_1 邻域范数距离。对式(13)两边求导得

$$T' = 1 + |I'| \quad (14)$$

式中, T' 和 I' 表示 T 和 I 的微分, 对上式两边积分可以得到 T , 如式(15)所示:

$$T(u) = \int_0^u 1 + |I'| dx, u \in \Omega \quad (15)$$

式中, T 是一个 R^2 到 R 的变换, 假设 u_1 和 u_2 是 T 上的 2 个点, 其距离为:

$$T(u_2) - T(u_1) = \int_{u_1}^{u_2} 1 + |I'| dx \quad (16)$$

从式(16)可知, 采用 L_1 范数距离经过变换之后的距离正好是原来 R^2 空间曲线 C 的侧地距离^[12]。而我们在降维的同时也失去了像双边滤波中通过 σ_s, σ_r 对空间域和值域的控制能力, 因此我们对式(15)进行变换得式(17):

$$T(u) = \int_0^u 1 + \frac{\sigma_s}{\sigma_r} |I'| dx, u \in \Omega \quad (17)$$

通过上述过程, 完成了把对 1D 信号的滤波从 R^2 降到 R , 而对 2D 信号进行滤波则沿着信号的每个维上执行分离处理^[13], 即先对行进行处理, 然后对它的结果进行列处理。又因为信号的滤波是通过与 R 空间的 H 卷积得到的, 所以下一步我们需要找到一个滤波核函数 H , 但由于执行卷积速度比较慢, 因此采用一阶边缘保持的循环滤波^[14]来代替, 如下所示:

$$Z[n] = (1 - a^d) \cdot I[n] + a^d \cdot Z[n-1] \quad (18)$$

式中, $d = T(x_n) - T(x_{n-1})$, $a = \exp(-\sqrt{2}/\sigma_H)$ 。由于通过式(18)处理后的结果依赖于前一次的输出, 因此它是不对称的, 而一个对称的处理是滤波 2 次, 先从左到右, 然后从右到左。图 3 是优化之后的效果对比图。

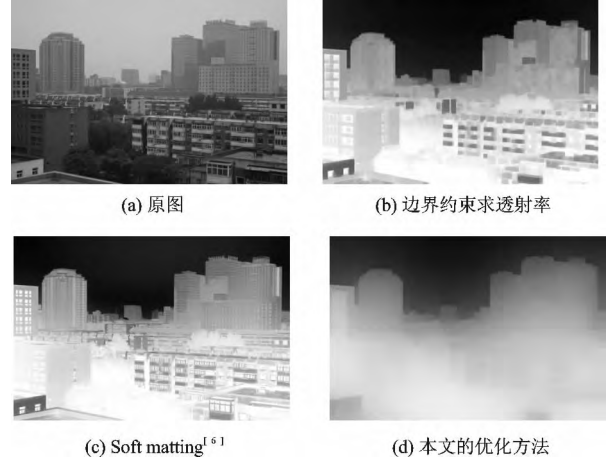


图3 透射率优化后的效果对比

3.3 A 和 t_0 的自适应估计

如果 A 的取值较大那么去雾后的图像会整体偏暗, 并且局部对比度也会降低, 反之则偏亮, 并且会造成天空区域失真。He^[6] 估计大气光值时, 当图像中有太亮的物体时会出现大气光值的错误估计。首先, 确定大气光值候选区域。由于大气光值应该是雾最浓的区域, 对应位置为天空或者无穷远处, 而天空区域常常位于图像的上部, 因此, 我们选择图像上方 10% 的区域作为大气光值的候选区域。其次, 通过提出的分块搜索的方法估计大气光值。第一步, 首先将候选块分成 4 个子块, 分别计算每个子块的平均值, 选择最亮的块再次划分为 4 个块, 重复执行直到块尺寸小于预先设置的下限值, 为了得到最终的大气光值, 我们选取最后得到子块中与理想的大气光值(255, 255, 255)距离最小的点进行大气光值的估计:

$$A = \min((255 - r_i) + (255 - g_i) + (255 - b_i)) \quad (19)$$

由于不同的雾气环境透射率下限值 t_0 的差别较大, 透射率下限值 t_0 的取值对含有天空区域的图像尤为重要, 取值过大会出现去雾不完全, 过小会出现严重失真。通过观察发现经过边界约束求得的透射率图, 天空区域和物体对象差别很大, 如图 3(c) 所示。为了使算法更具鲁棒性, 对 t_0 的选取, 选择边界约束得到的透射率图, 选取其中的较小的 0.5% 的像素点, 然后以它的最大值作为 t_0 的值。

图 4 是同一幅图像采用 He 的方法^[6] 和本文的自适应参数的效果对比图。

从图 4 可知, 经过 A 和 t_0 的自适应调整之后, 天空区域效果得到优化。由于带雾图像的亮度一般小于大气光值 A , 经过去雾之后图像会整体偏暗, 因此

(a) He的方法^[6]

(b) 本文的优化方法

图 4 A 和 t_0 的自适应估计效果对比

接下来需要对图像的亮度进行调节.

3.4 亮度的自适应调节

去雾公式(3)是逐点的线性变换,其斜率为 $\frac{1}{t}$,由于 $t \in [0, 1]$,所以它的斜率大于1,又因为这条直线必经过 (A, A) 这个点,这就从数学的角度上解释了去雾后图像对比度增大的同时,整体亮度会降低的现象.而我们发现文献[15]中的 Zadeh-X 变换后亮度提升的方式与式(3)得到的 J 亮度降低的方式存在惊人的相似,为了更好地复原到原始的无雾场景,因此我们用 Zadeh-X 变换对去雾后的图像进行整体亮度增强同时又不降低图像的局部对比度, Zadeh-X 变换的公式如下:

$$T(c, x, y) = K \cdot \frac{J(c, x, y) - \theta}{\Delta} + \theta \quad (20)$$

式中, $T(c, x, y)$ 为目标图像任一点的灰度 / 色度值, c 表示 3 个颜色通道, $J(c, x, y)$ 是根据式(3)得到的去雾后的值; $\theta \in [0, 255]$ 为设定的变换起始灰度 / 色度值, $\Delta \in [1, 255]$ 为设定的灰度 / 色度范围, 决定变换后的灰度 / 色度层次. $K \in [1, 255]$ 为设定的空间膨胀系数, 为了有最大的对比度分辨率, 设定 $K = 255$. 为了让亮度调整具有自适应性, 经过大量的样本研究, 获得参数 Δ 与平均亮度 L_A 的经验模型:

$$\Delta = 5 \cdot \text{abs}(L_A - \theta) 0.811 \quad (21)$$

式中,

$$\theta = \begin{cases} 0, & L_A < 127.5 \\ 255, & L_A \geq 127.5 \end{cases}$$

图 5 是亮度自适应调节前后的效果图, 从图中看出图像的整体亮度和对比度都得到了很好的提升, 具有更好的视觉效果. 而且它处理一幅 576×760 的图片仅仅只要 0.05 s.

4 实验对比分析

实验中, 本文针对偏暗、颜色不丰富、和偏亮的带雾场景进行处理, 如图 6 所示, 其中, 选取的窗口大小为 15×15 , 在一台 1.86 GHz 的双核电脑上用 MAT-



图 5 亮度自适应调节之后的结果

BLE10.0 平台进行处理. 实验给出了通过引导滤波改进的快速暗通道的方法^[7]和快速的 Wiener 滤波去雾的方法^[10]与本文的算法的效果对比结果.



(a) 偏暗带雾场景

(b) 颜色不丰富带雾场景

(c) 偏亮带雾场景

图 6 带雾原图



(a) 引导滤波改进的暗通道



(b) Wiener滤波去雾



(c) 本文算法

图 7 不同算法下处理效果图对比

图 6 中分别为 3 种不同场景下的原始图片, 而图 7 为不同算法下对原图的处理效果图. 其中, 引导

滤波改进的快速暗通道的算法处理效果如图 7(a) 所示,对偏亮的场景中高楼处理效果较好,但由于在偏暗场景和颜色不丰富的场景中,山和高速路不能很好地适用暗通道理论,因而处理效果较差.与此同时,改进的快速暗通道的方法采用的是引导滤波对透射率进行优化,但是原始通过暗通道求透射率图的方法并没有改变,处理一幅 576×768 的图片,通过暗通道求透射率图需要 0.48 s,引导滤波优化透射率图需 0.43 s,而最终完成去雾需要 1.04 s. 而 wiener 滤波去雾算法^[10] 由于采用的直接对图像的 3 通道最小值作滤波处理最终得到去雾结果,省略了分块求透射率图的步骤,在速度上有了显著的提升,它完成整个去雾流程仅需 0.54 s,但从图 7(b) 中可以发现处理效果图的视觉效果差,因而同样不适合偏暗和颜色不丰富的场景.而本文算法处理效果图如图 7(c) 所示,不仅同时适用于偏暗、颜色不丰富和偏亮的带雾场景而且视觉效果较好,同时本文算法从求透射率图和对透射率图优化上都降低了算法的复杂度,处理同样大小的图像用边界约束求透射率图需 0.19 s,用降维滤波优化透射率图需 0.23 s,最终完成整个去雾流程只需 0.58 s,其速度大约是引导滤波改进的快速暗通道的算法的 2 倍.与 wiener 滤波去雾的算法相比,最终的处理时间相差不多.因此,本文的算法除了适合多个场景去雾,同时提高了去雾处理效果和时间.

5 结束语

本文提出了一种快速的去雾算法,针对 He^[6] 的暗通道求透射率速度慢、soft matting^[6] 容易引入原图的噪声的问题,采用一个更为通用的边界约束条件来求透射率图,用一个降维滤波进一步优化透射率图和对模型中参数的自适应调节达到高质量的去雾,然后对去雾后的图像进一步自适应亮度调节达到更好的视觉效果.实验结果分析表明:本文算法很好地弥补了现有的算法仅仅针对偏亮的场景而对偏暗和颜色不丰富的场景去雾效果不佳的缺点,同时提升了速度,对视频实时去雾有重要的参考价值.

参考文献:

- [1] Narasimhan S G, Nayar S K. Contrast restoration of weather degraded images[J]. IEEE TPAMI, 2003, 25(6):713-724.
- [2] Kopf J, Neubert B, Chen B, et al. Deep photo: model-

- based photograph enhancement and viewing[J]. ACM Transactions on Graphics, 2008, 27(5):116-126.
- [3] Fattal R. Single image dehazing[J]. ACM Transactions on Graphics, 2008 27(3):721-729.
- [4] Tan R T. Visibility in bad weather from a single image [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Alaska, USA: IEEE Computer Society, 2008:1-8.
- [5] Jin Hwan Kima, Won Dong Janga, Jae Young Simb, et al. Optimized contrast enhancement for real-time image and video dehazing[J]. Journal of Visual Communication and Image Representation, 2013, 24(3):203-438.
- [6] He Kaiming, Sun jian, Tang Xiaou. Single image haze removal using dark channel prior[J]. Pattern Analysis and Machine Intelligence, 2011, 33(12):2341-2353.
- [7] He Kaiming, Jian Sun, Tang Xiaou. Guided image filtering[C] // Computer Vision - ECCV 2010. Springer, Berlin, Heidelberg, 2010:1-14.
- [8] Tarel J P, Hautiere M. Fast visibility restoration from a single color or gray level image[C] // IEEE 12th International Conference on Computer Vision, Boston, 2009:2201-2208.
- [9] Gibson K B, Vo D T, Nguyen T Q. An investigation of dehazing effects on image and video coding[J]. IEEE Transactions, 2012, 21(2):662-673.
- [10] Gibson K B, Nguyen D T. Fast single image fog removal using the adaptive wiener filter[C] // Image Processing (ICIP), 2013 20th IEEE International Conference on, China: Sanya, 2013:714-718.
- [11] Meng Gaofeng, Wang Ying, Duan Jiangyong, et al. Efficient image dehazing with boundary constraint and contextual regularization[C] // Computer Vision (ICCV), 2013 IEEE International Conference on Australia, Sydney, 2013:617-624.
- [12] Antonio Criminisi, Toby Sharp, Carsten Rother, et al. Geodesic image and video editing[J]. ACM Transactions on Graphics, 2010, 29(5):1-15.
- [13] Manuel M Oliveira, Gary Bishop, David McAllister. Relief texture mapping[C] // Proceedings of the 27th annual conference on Computer graphics and interactive techniques. USA: New York, 2000:359-368.
- [14] Eduardo S L Gastal, Manuel M Oliveira. Domain transform for edge-aware image and video processing [J]. ACM Transactions on Graphics (TOG), 2011, 30(4):69-71.
- [15] 谢正祥, 王志芳, 熊兴良, 等. 基于视觉感知噪声模型的彩色图像质量评价和彩色图像质量最佳化[J]. 中国图象图形学报, 2010, 15(10):1454-1464.

(下转第 55 页)

实验结果对比分析如下:

(1)图3表明,对于小规模数据库来说,两种方法的执行时间基本相同,而随着数据库规模的增大,GMNAR方法的执行时间明显少于Apriori方法的执行时间。

GMNAR方法仅需扫描一次数据库即可获得数据矩阵,一般情况下遗传算法发现关联规则所需的时间会远远少于获取数据矩阵所消耗的时间,因此,GMNAR方法的时间主要消耗在数据矩阵的构建过程中;而Apriori方法需要多次扫描数据库。当数据库规模比较小时,两者运行时间差距不明显;随着数据规模的增大,扫描数据库的时间远远多于生成数据矩阵的时间。因此,GMNAR方法比Apriori方法运行速率高一些。

(2)图4表明,GMNAR方法可以实现负关联规则的提取,且挖掘的负关联规则数量有所缩减,一定程度上解决了包含负项目的频繁项集爆炸问题。

通过引入相关度 φ_{xy} 这一评价依据,根据规则的支持度值和置信度值,实现了负关联规则的有效提取,并且通过调整置信度阈值的取值范围,可以剔除部分冗余信息,一定程度上缩小负关联规则的数量规模,尽快发现最有效的规则集。

5 结束语

本文提出了一种面向软件缺陷数据的负关联规则挖掘方法—GMNAR方法,并以XXX装备信息管理软件部分缺陷数据为例,进行了详细阐述^[10]。该方法引入了数据矩阵的思想以及改进的遗传算法,实现了负关联规则的提取,减少了数据库扫描次数,缩减了负关联规则数量规模,一定程度上提高了挖掘的效率。

参考文献:

[1] Agrawal R. Mining association rules between sets of i-

tems in large databases[C]// Proceeding of the 1993 ACM SIGMOD International Conference on Management of Data. Washington, ACM Press, 1993: 207-216.

[2] Brin S. Beyond market baskets: Generalizing association rules to correlations[C]// Proceeding of the 1997 ACM SIGMOD International Conference on Management of Data. USA: Arizona, ACM Press, 1997: 265-276.

[3] 董祥军,王淑静,宋瀚涛,等.负关联规则的研究[J].北京理工大学学报,2004,24(11):978-981.

[4] 朱红蕾,徐志刚,李明.负关联规则挖掘算法研究[J].微电子学与计算机,2010,27(5):167-169.

[5] 马占欣,陆玉昌.负关联规则挖掘中的频繁项集爆炸问题[J].清华大学学报:自然科学版,2007,47(7):1212-1215.

[6] 张卫国,高贤强,武晓宏.基于支持度-置信度框架的负关联规则研究[J].微电子学与计算机,2009,26(4):102-105.

[7] 朱玉全,陈耿,杨鹤标.正负关联规则挖掘算法研究[J].计算机科学,2006,3(3):188-190.

[8] 王国胤. Rough 集理论与知识获取[M]. 西安:西安交通大学出版社,2001.

[9] Xiong H. Exploiting a support-based upper bound of pearson's correlation coefficient for efficiently identifying strongly correlated pairs[C]// The 10th Intl. Conf. on Knowledge Discovery and Data Mining. Seattle, WA, 2004 8:334-343.

[10] Cohen J. Statistical power analysis for the behavioral sciences[M]. 2nd ed. New Jersey, Lawrence Erlbaum, 1988.

作者简介:

万琳女,(1974-),博士,副教授.研究方向为军事信息系统、软件工程、软件测试。

范秋灵(通讯作者)女,(1989-),硕士研究生.研究方向为军事信息系统、软件工程、软件测试。E-mail: fanqiuling1989@163.com.

(上接第49页)

作者简介:

刘杰男,(1988-),硕士研究生.研究方向为视频图像处理及应用。E-mail: 466914350@qq.com.

谢正祥男,(1938-),教授,博士生导师.研究方向为生物医

学信号和图像处理。

吕霞付男,(1966-),教授,研究生导师.研究方向为视频图像处理及应用。

陈学渊男,(1988-),硕士研究生.研究方向为视频图像处理及应用。