

基于改进双色模型的视程障碍类天气现象识别

郭佳^{1,2} 杨玲^{1,2} 吴可军³ 马舒庆³ 胡树贞^{1,2}

(1 成都信息工程学院电子工程学院, 成都 610225; 2 中国气象局大气探测重点开放实验室, 成都 610225; 3 中国气象局气象探测中心, 北京 100081)

摘要 针对各种视程障碍类天气现象,在现有双色大气散射模型的基础上,通过对大气散射光方向的修正、对矢量夹角算法的改进以及加入了湿度参数信息,提出了一种改进的双色模型算法,用于识别视程类天气现象。改进算法模型考虑了大气散射光方向的不完全一致性以及景物色度像素值在 RGB 空间中的三维性,因而具有更高的稳定性。该算法嵌入到视程障碍类天气现象识别系统中并计算出相应的特征值,最后通过经验阈值识别天气现象;通过对算法的改进并对哥伦比亚大学一个季度和甘肃张掖观象台 2 个月的数据进行处理,结果显示该算法对于雾、霾、扬沙、浮尘及沙尘暴等视程障碍类天气现象具有良好的判别效果,对上述两个地区的总体识别率分别为 73.2%、90.5%

关键词 双色模型 天气现象识别 特征值 大气散射光方向

引言

视程类天气现象主要是指空气中因存在水汽凝结物、干质悬浮物等而使空气变得混浊,并造成能见度下降的一类天气现象。包括雾、轻雾、雪暴、吹雪、沙尘暴、扬沙、浮尘、烟幕、霾等,对当地的生产、生活和交通运输都有极大的影响,再加之我国目前台站的天气现象业务观测仍未完全实现自动观测,主要靠人工观测来实现,从而迫切的希望具有自动化识别天气现象的方法出现。

国内外对视程障碍类天气现象的自动化观测主要依靠气象要素的综合判断,如可以通过能见度仪、湿度、风速等要素来综合判断^[1],并非从视程障碍类天气现象在大气中的形成机理为出发点,因而在分类上存在很多问题,国内很少有独立对视程障碍类天气进行观测的仪器,这样也就造成了能观察到的视程类天气现象较少,准确率不高等缺点^[2]。

Srinivasa G. Narasimhan 和 Shree K. Nayar 通过对不同天气条件下图像呈现不同信息的特点进行了研究,提出了双色模型^[3]的基本算法来获得不同恶劣天气情况下所匹配的场景属性。接着,他们

模拟了大气散射模型,并在雾和霾的情况下得到证实。基于这个模型,得出由大气状况变化所引起的场景色彩变化的几何约束条件,具有较强的应用性,此类课题在国内尚处发展阶段,具有较强的研究价值^[4]。

本文针对气象识别领域上的应用,在原双色模型的基础上,对算法中关于大气散射光方向进行了优化以减小其误差,并对矢量夹角做了更深层次的信息提取从而更好的利用原始信息。最终形成了一种专门针对视程障碍类天气现象的识别方法。

1 双色大气散射模型基本算法

1.1 直接辐射光分量和大气散射光分量

双色传输模型是由哥伦比亚大学的 Srinivasa G. Narasimhan 和 Shree K. Nayar 提出的一种大气散射彩色模型,它描述了恶劣天气下的场景点色彩是由直接传输光色彩和大气散射光色彩线性组合而成。其中直接传输分量是指光线穿过大气到达相机的敏感元件的光学视程过程中会收到并通过大气气溶胶粒子的阻碍以及散射发生衰减之后的色彩分量。这种直接传输可以由伯格指数公式^[3]得出。

<http://www.qxkj.net.cn> 气象科技

公益性行业(气象)专项天气现象自动化观测技术研究(GYHY200906032)资助

作者简介:郭佳,男,1986 年生,硕士研究生在读,主要研究大气探测理论与技术,Email:297229058@qq.com

收稿日期:2011 年 12 月 31 日;定稿日期:2012 年 9 月 7 日

$$E_{\infty}(\lambda) = gL_0(\lambda)e^{-\beta(\lambda)} \quad (1)$$

$E_{\infty}(\lambda)$ 指的是无穷远处天空的亮度, $L_0(\lambda)$ 是场景点前向衰减幅度, g 表示了相机参数, $\beta(\lambda)$ 为大气散射系数,它是关于特定光线波长的一个函数。McCartney(1975)从建立的一种衰减模型中得出这种衰减基本上符合一种指数分布^[5]。相机获得景物的直接辐射光分量用公式(2)描述。

$$E_d(d, \lambda) = \frac{E_{\infty}(\lambda)r(\lambda)e^{-\beta(\lambda)d}}{d^2} \quad (2)$$

根据参考文献[6]可知式(2)可以简化为式(3)。

$$E_d(d) = \frac{E_{\infty}re^{-\beta d}}{d^2} \quad (3)$$

大气散射光分量是指太阳光或者说天空光经过大气的气溶胶粒子时发生散射和折射的作用,最终进入到 CCD 的那部分色彩光,这部分光线随着距离的增加而减小,可用下面的公式描述^[5]:

$$E_a(d, \lambda) = E_{\infty}(\lambda)(1 - e^{-\beta(\lambda)d}) \quad (4)$$

同理参考文献[6]可简写为

$$E_a(d) = E_{\infty}(1 - e^{-\beta d}) \quad (5)$$

1.2 双色模型和双色平面

根据双色大气散射彩色模型原理,可以把景物色度看成是大气散射光分量和直接辐射光分量之和, $\hat{D}$ 表示直接辐射光 P 方向上的单位向量。 $\hat{A}$ 表示大气散射光方向上的单位向量^[7],因此,可以写成

$$E = p \times \hat{D} + q \times \hat{A} \quad (6)$$

其中, p 表示直接传输光方向上($\hat{D}$)的幅值, q 表示大气散射光方向上($\hat{A}$)的幅值,如图1所示。

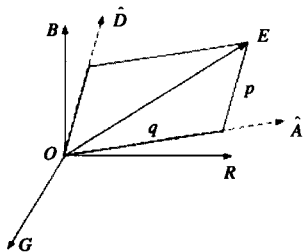


图1 双色大气散射模型

经过实验测试得出在天气条件不太恶劣的情况下,同一场景在不同天气条件下所成图像中,同一点 P 的各个像素矢量方向非常接近上述两个向量所组成的平面,我们把这个平面称为双色平面。因此,不同天气情况下观察所得的场景点 P 色彩,位于一个

单独的双色平面^[7],如图2所示。

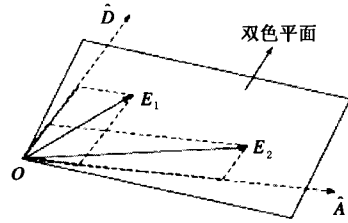


图2 双色平面

1.3 光学厚度差算法

根据双色平面的结论可知:在不同天气情况下同一场景点的像素值在同一个双色平面内,这样就可以假设 E_1 和 E_2 为不同天气现象的情况下同一场景点 P 的像素值大小和方向,如图2所示。根据参考文献[6]可知,大气散射光的方向是趋于稳定的,这样我们可以通过以下算法来确定大气散射光的方向,在两种不同天气条件下取同一场景中相同点的像素值分别为: $E_1^{(i)}$ 和 $E_2^{(i)}$,同理得到另外一组像素值为 $E_1^{(j)}$ 和 $E_2^{(j)}$,这4个像素点两两形成一个双色平面,如图3所示,这两个双色平面的公共交线即为大气光散射分量的方向。根据参考文献[6]可以得到大气散射光 $\hat{A}$ 的方向(如图3所示):

$$\hat{A} = \frac{N_i \times N_j}{\|N_i \times N_j\|} \quad (7)$$

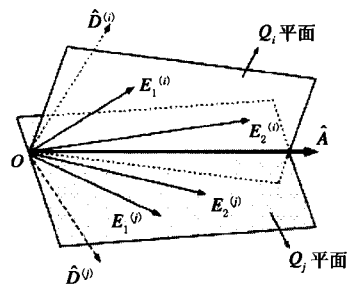


图3 $\hat{A}$ 方向确定

根据双色模型的理论可以知道,图像中任意一点的像素值在 RGB 空间中都可以表示成大气散射光分量和直接辐射光分量之和,又由不同天气现象情况下同一场景点的像素值满足同一个双色平面,这样就可以把不同天气情况下同一场景处的两个像素值分别分解在两个矢量方向上,如图4所示,其详细内容见参考文献[6]。

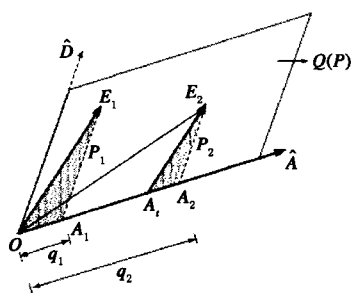


图4 双色模型的分解

E_1 和 E_2 在直接辐射分量方向上的大小

$$P_1 = \frac{E_{\infty 1} r_1 e^{-\beta_1 d_1}}{d_1^2} \quad (8)$$

$$P_2 = \frac{E_{\infty 2} r_2 e^{-\beta_2 d_2}}{d_2^2} \quad (9)$$

由于选取的是同一景物点,因此上式中的距离 d_1 和 d_2 相等,景物的反射属性 r_1 和 r_2 相同,故有

$$\frac{P_2}{P_1} = \frac{E_{\infty 2}}{E_{\infty 1}} e^{-(\beta_2 - \beta_1)d} \quad (10)$$

上式经过变换,取对数可得

$$(\beta_2 - \beta_1)d = \ln\left(\frac{E_{\infty 2}}{E_{\infty 1}}\right) - \ln\left(\frac{P_2}{P_1}\right) \quad (11)$$

式中 d 在实际的测量中为一个常数,如果知道两种天气条件下无穷远处天空的亮度信息就可以求出这两幅图像在该场景点处的大气散射系数之间的差值,我们把这种差值称为大气光学厚度差,简称 DOT^[8],在不同的天气情况下,大气光学厚度差也不一样,这样我们就可以根据不同的光学厚度差作为天气识别的一个特征参数,用于天气识别^[9]。

1.4 向量夹角算法描述

所谓向量夹角法即是指不同天气现象条件下同一场景中的两个像素值,在 RGB 空间中矢量方向上的夹角,同一场景中不同天气现象时同一点的像素值所形成的向量夹角不同。

图5表示同一场景不同天气条件下3个像素点的像素值在双色平面上的分布情况。如图所示, E_1 、 E_2 、 E_3 分别表示晴天、霾和雾3种天气下场景点的像素值,设晴朗天气时的像素值为基准值,其具体细节参见参考文献[6],则基准图片各像素点与待测图片对应像素点之间的夹角关系,可由以下公式^[9]求得:

$$\cos\langle \overrightarrow{OE_1}, \overrightarrow{OE_i} \rangle = \frac{\overrightarrow{OE_1} \cdot \overrightarrow{OE_i}}{|\overrightarrow{OE_1}| \cdot |\overrightarrow{OE_i}|}$$

$$= \frac{x_1 x_i + y_1 y_i + z_1 z_i}{\sqrt{x_1^2 + y_1^2 + z_1^2} \sqrt{x_i^2 + y_i^2 + z_i^2}} \quad (12)$$

通过这个公式算出所选区域各点的向量夹角之和在求平均即可得到两幅图的一个平均的向量夹角值,最终作为判定天气现象的一个特征。

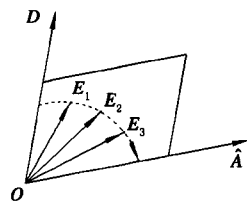


图5 不同天气条件下同一点处的双色模型

2 改进双色模型算法

2.1 对大气散射光方向($\hat{A}$)的修正

根据光学厚度差算法的基本原理,可知由于短时间内气溶胶粒子是稳定的,所以 $\hat{A}$ 的方向是趋于稳定的^[9],这样也就造成原算法在此就忽略了对大气散射光方向不一致性的研究,虽然说整个图片中所有像素点分解在大气光散射方向上的分量是趋于稳定的,但是向量值并非完全相等,这样也就带来了一定的误差,虽然说大气散射光方向是稳定并且一致,但是在色彩空间中也并非完全重合,这点可以通过图6得以证明,图6中不同颜色的平面表示了不同的双色平面,而这些双色平面的交线即是这里所定义的大气散射光的方向,由图6可以看出这些双色平面交线的方向确实是一致并且趋于稳定的,但是也可以看出他们并非重合的^[10],也就是说原算法在这个地方是有误差的,由此可知这些双色平面的任意一个交线都不能代表大气散射光的方向,如果用其中一个方向来代表大气散射光的方向就会造成不同程度的误差,也就是算法的稳定度降低,于是,需要找出一个使误差最小的大气散射光的方向,在此可以通过使得方向 $\hat{A}$ 和法平面叉乘的平方和最小来求最佳的 $\hat{A}$ 方向^[11],由如下公式所示:

$$\epsilon = \sum_i (N_i, \hat{A})^2 \quad (13)$$

如果整个平面的像素点的 $\hat{A}$ 方向完全一致,则 ϵ 的理论值就应该是零,但由于有一定的误差,所以使得这个平方和最小即为最佳的 $\hat{A}$ 方向。

图6中有色平面为图像中所取点形成的双色平面,可以看出平面的交线没有重合但是具有可拟合

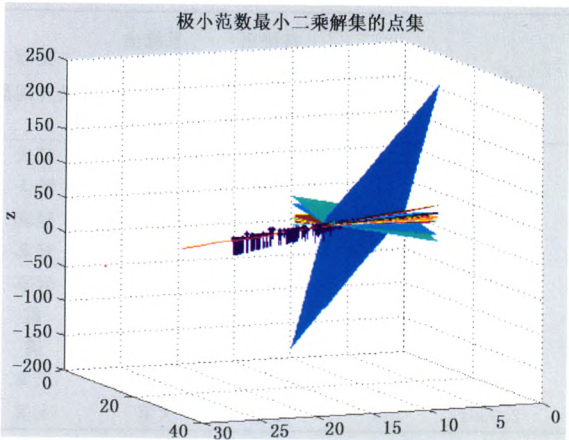


图 6 大气散射光方向最小二乘拟合

性,其中蓝颜色的点为各交线延长线上的点,红色线条为拟合的大气散射光方向 $\hat{A}$ 。通过后续的测试结果可以知道通过拟合得到的大气散射光的方向具有误差小更接近实际情况的特点^[12]。

2.2 对矢量夹角算法的改进和湿度参数的利用

通过对原始双色模型中矢量夹角数据的分析发现,矢量夹角的结果是一个角度或者角度余弦值的标量,并没有得到一个有效反映空间向量位置关系的量,在此通过把色彩向量投影到 RGB 的每个平面上再计算差值,这样得到的结果就会是一个矢量值,从而比较有效的得出颜色向量和每个平面之间的关系,如图 7 所示。

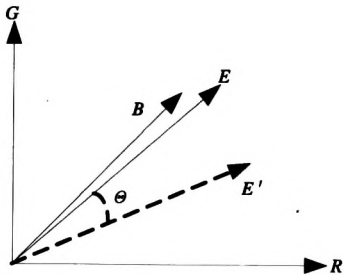


图 7 景物色度在 RB 上的投影

图 7 中 E 为景物色度, E' 为景物色度在 RG 平面上的投影, θ 为景物色度与 RB 平面上投影的夹角,同理可以得到景物色度与 RG 和 GB 平面投影的夹角值,在通过与晴天条件下景物色度与各平面夹角值之差来作为特征值,最终参与天气现象的判别,这样就能利用向量空间夹角差值这个矢量来改进向量夹角这个标量值,已达到更好利用原始信息的目的。

在特征值参数中我们通过对轻雾、霾天气现象的分析发现,这两类天气现象在色度空间中的信号表现非常接近,很难分辨,但是二者在湿度上有较大的差异,所以可以通过湿度信息来区分。

3 识别结果分析

3.1 实验结果

为了验证改进算法的优越性以及系统稳定性的验证,我们借助哥伦比亚大学提供的图片和张掖观象台的图片来验证算法的稳定性和优越性,分别对哥伦比亚大学提供图片(图 8)和张掖观象台的图片(图 9)通过原双色模型和改进后双色模型进行参数运算和判别,表 1 为原双色模型和改进后双色模型对哥伦比亚提供图片处理后的特征值及结果对照,表 2 为原双色模型和改进后双色模型对张掖观象台提供图片处理后的特征值及结果对照。表中参数 DOT 代表光学厚度差值, θ 为矢量夹角值, θ_R 表示

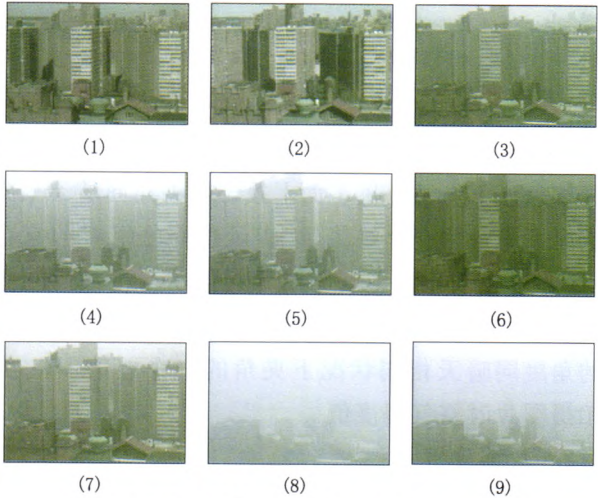


图 8 各种天气现象(哥伦比亚)

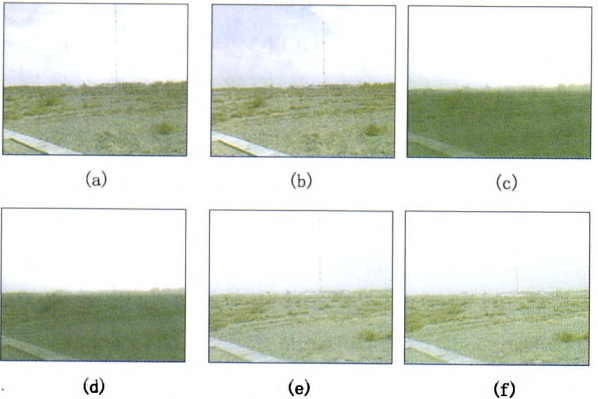


图 9 各种天气现象(甘肃省张掖观象台)

表 1 (哥伦比亚)参数值及天气现象结果(原双色模型和改进后模型对照)

图片号	DOT 原模型	DOT 改进模型	θ	θ_R	θ_B	θ_G	$H_{um}/\%$	系统出 天气现象 (原模型)	系统出 天气现象 (改进模型)	实际 天气现象
1	0.10425	0.04098	0.46732	0.34715	-0.3471	-0.7742	31	晴天	晴天	晴天
2	0.23086	0.11038	0.37855	0.89313	-0.5631	-0.8932	41	晴天	晴天	晴天
3	0.44952	0.59152	1.69464	-2.1758	1.07579	0.92314	96	轻雾	轻雾	轻雾
4	1.09821	0.99976	1.42428	-2.4758	1.37586	1.43167	88	轻雾	轻雾	轻雾
5	1.06732	0.73974	1.71941	-2.4486	1.27595	1.29631	87	轻雾	轻雾	轻雾
6	0.54318	0.32840	0.85566	-0.9682	0.86831	0.77219	72	霾	霾	霾
7	1.03729	0.61439	1.85456	-1.2748	0.93176	0.81679	65	轻雾	霾	霾
8	5.83751	5.36865	4.99707	-4.3412	2.34573	2.37593	99	大雾	大雾	大雾
9	3.14529	3.61896	5.15749	-5.0593	2.09583	2.98613	100	大雾	大雾	大雾

表 2 (张掖)参数值及天气现象结果(原双色模型和改进后模型对照)

图片号	DOT 原模型	DOT 改进模型	θ	θ_R	θ_B	θ_G	$H_{um}/\%$	系统出 天气现象 (原模型)	系统出 天气现象 (改进模型)	实际 天气现象
a	0.417618	0.26942	0.517544	0.297417	0.328619	-0.811874	35	晴天	晴天	晴天
b	0.429157	0.37821	0.518190	0.238619	0.226913	-0.298135	39	晴天	晴天	晴天
c	1.382967	1.14376	2.459213	-1.07529	3.482196	0.923147	33	浮尘	浮尘	浮尘
d	1.374912	1.32186	2.437180	-1.32761	3.631894	1.184195	41	浮尘	浮尘	浮尘
e	1.343836	1.33845	1.427825	-2.54861	1.495291	1.784143	36	扬沙	扬沙	扬沙
f	1.583120	1.54000	2.291304	-2.76824	1.531875	1.772197	42	浮尘	扬沙	扬沙

景物色度 E 与在 GB 平面上投影的夹角同晴天相同状况下夹角的一个差值, θ_G 表示景物色度 E 与它在 RB 平面上投影的角度同晴天相同状况下夹角的一个差值, θ_B 表示景物色度 E 与它在 RG 平面上投影的角度同晴天相同状况下夹角的一个差值, H_{um} 为拍摄图片时空气湿度值。

3.2 数据分析

通过原双色模型算法和改进型的双色模型算法对两个不同地方天气现象图片的处理,可以发现在以下 2 个方面得到了改进:

(1)在光学厚度差(DOT)参数在改进双色模型算法中具有更好的稳定性,原算法得出的光学厚度差值(DOT)

不稳定的主要原因在于大气散射光方向的误差较大,造成了在后期计算中误差扩大和结果的不稳定。

(2)通过改进后的矢量夹角算法得到了更多的有用信息,对于人们视觉难以判断的轻雾和霾天气现象具有一定的作用,同时湿度对于轻雾和霾天气现象的判别具有很好的辅助作用,虽然在张掖对于浮尘和扬沙的作用并不明显,但是对于轻雾和霾这样肉

眼很难分辨的天气现象来说是一个非常有用的参数。

虽然通过算法的改进可以提高对视程障碍类天气现象的识别,但是还是存在一些问题,主要问题如下:

(1)对于肉眼难以区别的浮尘和扬沙湿度并没起到太好的辅助作用,所以在以后的算法改进中可以考虑风速的影响来进一步提高识别率。

(2)在本算法中用的是阈值范围判别法来判别天气现象的,在以后可以考虑运用神经网络模型的自学习方式判断天气现象,使得系统更加优化。

本次研究对哥伦比亚大学提供的一个季度的 141 张各类天气现象图片和张掖观象台采集连续 2 个月的 132 张图片做了测试,原双色对哥伦比亚图片的整体识别率为 69.2%,通过改进后算法对哥伦比亚图片的整体识别率为 73.2%,原模型对张掖图片的整体识别率为 89.3%,改进后算法模型对张掖的识别率为 90.5%,算法模型对哥伦比亚大学所提供图片的识别率有所偏高,其主要原因是由于图像采集时间偏短且天气现象类型较少。改进后模型对哥伦比亚图片的识别率提高较多的原因是湿度起到了比较关键的作用,而张掖图片中湿度所起到的效果并不明显。

4 总结

本系统通过对双色模型算法中相关参数的修正,和改进来获取更稳定的判别效果,并将算法融入到系统当中,这样对于算法的结果有较强的可读性。原算法对于人眼难以辨别的天气现象的区分度也比较差,如对轻雾和霾的判别上效果不太好,在对浮尘和扬沙的判别上有一定的误差,在对算法的改进上面我们加入湿度作为辅助手段。同时图像处理方法在气象探测方面的应用也具有很大的潜力,所以希望在以后的发展中能有更多的图像处理算法融入到气象探测当中。

参考文献

- [1] 陈冬冬,魏国栓,马启明,等.多台国内天气现象仪对比分析.气象水文海洋仪器,2010,(4):13-17.
- [2] 陈东东,施丽,娟,李肖霞,等.天气现象自动化观测现状调研[J].气象科技,2011,39(5):596-602.
- [3] Narasimhan S G, Nayar S K. Chromatic Framework for Vision in Bad Weather [C]//IEEE Conference on Computer Vision and Pattern Recognition, 2000:598-605.
- [4] 顾大权,范茵,许屏,等.建立天空场景的模拟算法[J].气象科技,2007,35(2):292-296.
- [5] Narasimhan S G, Nayar S K. Chromatic Framework for Vision in Bad Weather [R]. IEEE, New York; Department of Computer Science, Columbia University, 2000.
- [6] 胡树贞,杨玲,马舒庆,等.基于双色大气散射模型的视程障碍类天气现象识别[C]//第十二届中国科协年会,2010.
- [7] Narasimhan S G, Nayar S K. Removing Weather Effects from Monochrome Images [R]//IEEE. New York; Department of Computer Science, Columbia University, 2001.
- [8] Narasimhan S G, Nayar S K. Interactive (De)Weathering of an Image using Physical Models [R]. ICCV, New York; Department of Computer Science, Columbia University, 2003.
- [9] Shechner Y Y, Narasimhan S G, Nayar S K. Polarization-based vision through haze [R]. Optical Society of America, New York; Department of Computer Science, Columbia University, 2003.
- [10] Narasimhan S G. Models and Algorithms for Vision through the Atmosphere [D]. Graduate School of Arts and Sciences, 2003.
- [11] Shechner Y Y, Narasimhan S G, Nayar S K. Polarization-based vision through haze [R]. Optical Society of America, New York; Department of Computer Science, Columbia University, 2003.
- [12] 田宏伟,郑友飞,陈怀亮,等.郑州地区气溶胶光学厚度反演与分析[J].气象科技,2010,38(4):515-520.

Recognition of Weather Phenomena with Low Visibility on Basis of Improved Dichromatic Atmospheric Scattering Model

Guo Jia^{1,2} Yang Ling^{1,2} Wu Kejun³ Ma Shuqing³ Hu Shuzhen^{1,2}

(1 Electronic Engineering College, Chengdu University of Information Technology, Chengdu 610225;

2 Key Lab of Atmospheric Sounding, China Meteorological Administration, Sichuan, Chengdu 610225;

3 CMA Meteorological Observation Centre, Beijing 100081)

Abstract: Concerning various kinds of low visibility weather phenomena and based on the existing dichromatic atmospheric scattering model, a modified calculating method of dichromatic model is proposed, to range weather phenomenon via revising the atmospheric scattering direction, improving the calculating method of vector angles and adding humidity parameter information. The improved algorithm model has higher stability, considering that the scatter light direction is not all consistent and pixels have three-dimensional character in the RGB space. Embedding the algorithm into the model and computing the corresponding eigenvalues, these weather phenomena can be recognized through the empirical thresholds. By improving the calculating method and processing the seasonal data of Columbia University and two-month data of Zhangye Observatory, it is shown that the improved method is excellent in identifying various kinds of low visibility weather phenomena, such as fog, haze, dust blowing, dust floating, and sand storm, whose total identification rates for the above two regions are 73.2%, 90.5%, respectively.

Key words: dichromatic model, weather phenomenon recognition, eigenvalue, atmospheric scattering direction