

文章编号:1001-9081(2009)08-2112-04

基于特征点的视频流拼接帧选择优化算法

王铁建^{1,2}, 赵红领^{1,3}, 王宗敏¹

(1. 郑州大学 河南省信息网络重点学科开放实验室, 郑州 450052; 2. 郑州大学 信息工程学院, 郑州 450001;

3. 郑州大学 水利与环境学院, 郑州 450001)

(znmwang@zzu.edu.cn)

摘要:提出了一种具有公共视野的视频流拼接中视频帧的选择优化算法,算法计算候选帧与基准帧特征匹配的特征向量差值的算术均值,选择均值最小的候选帧作为最佳拼接帧,与基准帧进行拼接。实验结果表明,基于该算法拼接的视频流全景视图效果较好,算法很好地解决了视频流拼接中最佳拼接帧的自动选择问题。

关键词:基于图像的绘制技术;尺度不变特征变换;视频流拼接;帧选择

中图分类号: TP391.41 **文献标志码:** A

Optimized frame selection algorithm for video stitching based on feature points

WANG Tie-jian^{1,2}, ZHAO Hong-ling^{1,3}, WANG Zong-min¹

(1. Henan Provincial Key Laboratory on Information Network, Zhengzhou University, Zhengzhou Henan 450052, China;

2. School of Information Engineering, Zhengzhou University, Zhengzhou Henan 450001, China;

3. School of Water Conservancy and Environmental Engineering, Zhengzhou University, Zhengzhou Henan 450001, China)

Abstract: This paper presented an optimized algorithm of frame selection on the stitching of videos with overlapped visions. Firstly, it compared the Scale Invariant Feature Transform (SIFT) value differences of all the match points of standard frame and candidacy frames, and then selected the point pairs with minimal SIFT value difference, and calculated the mean values of these differences. Finally, the frame with the minimum mean value was selected as the best frame, and stitched with the standard frame. This algorithm assures the best matching frame and can be used to choose the best stitching frame automatically in video stitching. Simulation shows the effectiveness of this algorithm.

Key words: Image Based Rendering (IBR); Scale Invariant Feature Transform (SIFT); video stitching; frame selection

0 引言

基于图像的绘制技术(Image Based Rendering, IBR)主要研究如何采用已知实景图像来直接生成虚拟视点下的相应图像^[1-2]。由于 IBR 从根本上解决了基于模型绘制技术(Model Based Rendering, MBR)在建模工作繁杂、实时渲染困难、真实感不足等方面的致命缺陷,所以在近年来得以迅速发展。IBR 成败的关键是采用图像拼接技术来构建能够观察到全局场景的全景图像^[3-4](Panoramic Image),而目前,大多数的图像拼接技术研究还集中于静态图像的拼接。

将图像拼接技术应用于动态的视频流,从而将两段有部分公共视野的视频流拼接为一段视野范围较大的全景视频流,使人们在浏览该段视频流时对场景内容有更直观、详尽的了解,具有重要的理论和实用价值。当然,动态视频流的拼接与静态图片的拼接相比具有其特殊性:动态视频流中图像帧的数目要比静态图像多得多^[5]。由于存在 DV 开机不同步等诸多问题,按照采集得到的帧序号直接将两段视频流进行拼接,并不一定能够得到较好的全景视图。如何从两段视频流中找出最为合适的视频帧进行拼接,得到最佳的全景视图,进而得到最佳的全景视频流,是视频流拼接的关键问题。本文在提取尺度不变特征变换^[6-7](Scale Invariant Feature Transform, SIFT)特征点的基础上,提出了基于 SIFT 特征向量的最佳拼接帧选择算法,很好地解决了这一问题。

1 总体设计

视频流拼接的过程大致分为视频流采集、视频帧提取、视频流拼接三步^[8]。为了得到最佳的全景视频流,必须在视频流拼接之前,加入最佳视频帧选择步骤,对两段视频流中提取到的视频帧进行分析选择,找出最为合适的视频帧进行拼接,从而得到最佳的全景视图(图 1)。

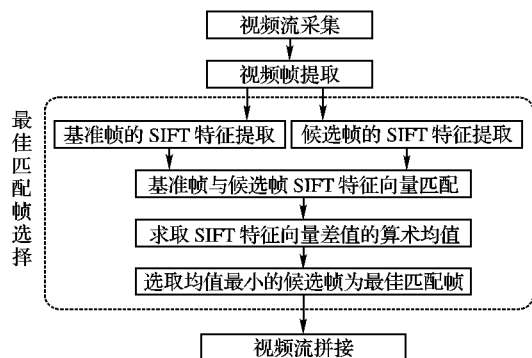


图 1 视频流拼接流程

2 视频帧的采集

2.1 采集视频流

视频流的采集过程如下:把两台 DV 固定在操作台上,使 DV₁ 和 DV₂ 有部分公共视野,匀速推动操作台前行,则两台

收稿日期:2009-02-23;修回日期:2009-04-08。 基金项目:河南省自然科学基金资助项目(0611051900)。

作者简介:王铁建(1982-),男,河南周口人,硕士研究生,主要研究方向:虚拟现实; 赵红领(1978-),男,河南郑州人,讲师,博士研究生,主要研究方向:虚拟现实、水利信息技术; 王宗敏(1964-),男,河南郑州人,教授,博士生导师,主要研究方向:虚拟现实、网络技术与应用、水利工程信息技术。

DV 分别采集到相邻场景的视频流(图2)。

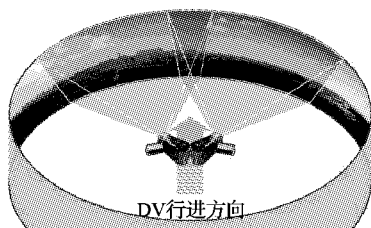


图2 视频帧的采集

2.2 提取视频帧

采集得到视频流以后,还需要将视频流提取为视频帧序列。通过程序实现视频流的提取,DV 的帧速率是 25 帧/秒,可以根据需要,决定每隔多少帧提取一帧^[9]。

3 SIFT 特征提取

提取得到的视频帧实际上是一幅幅的静态图像。接下来就要通过 SIFT 算法,找出每幅图像的 SIFT 特征,为将来的图像配准做准备。SIFT 算法分三个阶段实现。

3.1 检测尺度空间极值点

检测高斯差分(Difference of Gaussian, DOG)尺度空间的极值点,目的是把提取到的极值点作为候选 SIFT 特征点。这一步分三个阶段完成,不仅保证了特征点的缩放不变性,而且保证了特征点不受亮度差异和噪声影响。

1)构建高斯尺度空间。高斯图像金字塔共 O 阶($O = 4$),每阶有 S 层尺度图像($S = 5$),第一阶的第一层是放大 2 倍的原始图像,下一阶第一层由上一阶的中间层尺度图像进行亚采样得到;同一阶中相邻两层的尺度因子比例系数为 k 。

2)构建 DOG 尺度空间。高斯金字塔中相邻尺度空间函数相减即得到 DOG 金字塔。DOG 金字塔第 1 层的尺度因子与高斯金字塔的第 1 层一致,其他阶也一样(图3)。

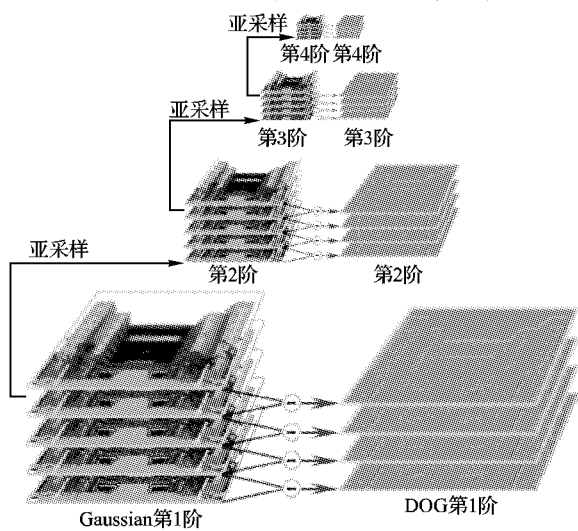


图3 高斯差分尺度空间

3)在 DOG 空间中找出极值点作为候选特征点。为寻找尺度空间的极值点,DOG 尺度空间中间层的每一个采样点都要和它同尺度的 8 个相邻点和邻尺度的 9×2 个相邻点(共 26 个点)做比较(图4),以确保在尺度空间和图像空间都检测到局部极值点,标记下局部极值点的位置和对应尺度。

3.2 确定特征点的位置、尺度和方向

3.2.1 确定特征点的位置和尺度

由于 DOG 值对噪声和边缘较敏感,因此还要对 DOG 尺度空间中的局部极值点进行三维二次函数拟和才能精确确定

特征点的位置和缩放尺度;同时,还应去除那些低对比度的特征点和不稳定的边缘响应点,以提高抗噪声能力,增强匹配稳定性。

1)确定哪些极值点为特征点。局部极值点 (x_0, y_0, σ_0) 处的三维二次函数拟和函数(尺度空间函数在局部极值点的泰勒展开式)为: $D(x, y, \sigma) = D(x_0, y_0, \sigma_0) + \frac{\partial D^T}{\partial X} + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X$;对这个式子求导并令导数为 0,求得精确的特征点:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X}, \text{即可获取特征点的精确位置和缩放尺度。}$$

2)去除低对比度的特征点。特征点对应的极值为: $D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X}$;剔除极值小于 0.03 的低对比度特征点。

3)去除不稳定的边缘响应点。特征点的 Hessian 矩阵 $H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$ 的特征值 α 和 β 分别代表特征点 X 方向和 Y 方向的梯度,剔除仅对一个方向梯度大的边缘相应点。

3.2.2 确定特征点的主方向

利用特征点邻域像素的梯度方向分布特性为每个特征点指定主方向,使特征点具备旋转不变性。

计算出特征点邻域内的像素点 $L(x, y)$ 的梯度大小 $m(x, y)$ 和方向 $\theta(x, y)$,进行高斯加权求和,把邻域内所有像素点的投票方向作为该特征点的主方向。

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}$$

$$\theta(x, y) = \alpha \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right)$$

实际计算时在以关键点为中心的邻域窗口内采样,并用方位直方图统计邻域像素的梯度方向,方位直方图的最高峰值点对应处的方向就是该关键点的主方向(存在高于最高峰值 80% 的高峰值时,将这个方向认为关键点的辅方向,可以增强匹配鲁棒性)。

3.3 生成 SIFT 特征向量

我们把特征点邻域的方向信息进行联合,生成 SIFT 特征向量。这样可以大大增强特征点的抗噪声能力和容错能力,从而增强图像匹配的稳健性(图4)。

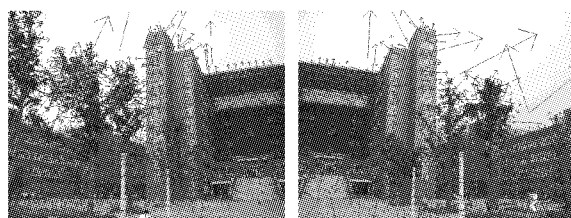


图4 SIFT 特征向量

4 视频流拼接

4.1 匹配 SIFT 特征向量

对于两幅视频帧,分别获取 SIFT 特征向量后,建立 K-D 树^[10]进行优先搜索,查找每个 SIFT 特征点的最邻近特征点,进行 SIFT 特征点对的匹配。

在特征点匹配过程中,只保留两幅拼接图像的对应匹配特征点对,对这 N 对匹配特征点对,按照它们的 SIFT 特征向量差值进行从小到大排列,只保留前 M 对匹配特征点对,删除差值过大的匹配特征点对,这样就能保证保留的特征点对

匹配得相当精确(图5)。

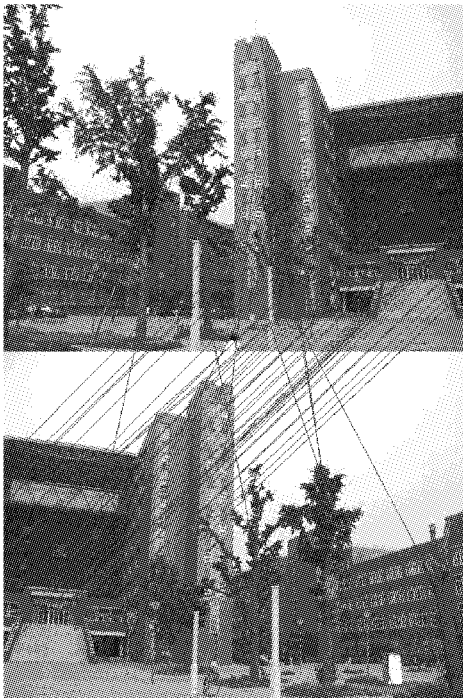


图5 SIFT特征向量的匹配

4.2 选择最佳拼接帧

由于存在DV开机不同步等问题,直接按照帧序号进行拼接,即:把DV₁采集到的第 a_i 帧与DV₂采集到第 b_i 帧拼接,并不一定能够得到最佳全景。为了得到最佳全景,采取如下最佳拼接帧自动选择算法:

设DV₁采集到的视频序列为 $a_1, a_2, a_3, \dots, a_i, \dots, a_n$, DV₂采集到的视频序列为 $b_1, b_2, b_3, \dots, b_n$ 。

1) 取DV₁采集的第 a_i 帧作为基准帧;取DV₂采集的第 b_{i-m} 帧至第 b_{i+m} 帧共 $2m$ 帧作为候选帧。

2) SIFT配准。用DV₁的基准帧 a_i 与DV₂的候选帧 b_{i-m} 进行SIFT特征匹配,得到一组匹配特征点对。

3) 计算特征点对差值并排序。对该组匹配特征点对,计算出每对匹配特征点的差值 $Distance(p)$,并按差值 $Distance$ 从小到大排列,保留差值 $Distance$ 小的前 N 对。

4) 计算算术均值。算出基准帧 a_i 与候选帧 b_{i-m} 的前 N 对匹配特征点对差值 $Distance$ 的算术均值 $\bar{D}_{(a_i, b_{i-m})} = \frac{1}{N} \sum_{p=1}^N Distance(p)$ 。

5) 对基准帧 a_i 与其他候选帧,依次进行2)~4)处理,求出各自的算术均值 \bar{D} ,最小均值 $\min(\bar{D})$ 对应的第 b_j 帧为基准帧 a_i 的最佳拼接帧。

通过上述算法,即可从 $2m$ 帧候选帧里选出基准帧的最佳拼接帧。将基准帧与该候选帧进行拼接,即可得到最佳全景图,进而得到最佳全景视频。

5 实验结果

实验采用的DV型号为:Sony HDR-HC1E,场景为郑州大学新校区计算中心教学楼,提取视频帧时设定每隔6帧抽取1帧,即抽取每秒内的第1、6、11、16、21帧图像进行实验。

取DV₁采集的第 a_{11} 帧图像作为基准帧。根据匹配算法,应该取DV₂采集的第 b_1 、 b_6 、 b_{11} 、 b_{16} 、 b_{21} 帧作为候选帧(图6)。

根据SIFT算法,将基准帧 a_{11} 分别与候选帧 b_1 、 b_6 、 b_{11} 、

b_{16} 、 b_{21} 进行SIFT特征匹配。依次执行最佳匹配帧筛选算法的第3)步,分别计算出匹配特征点对的差值 $Distance$,只保留差值 $Distance$ 最小的前15对(若第15对差值大于2.5,则将其去掉)。然后依次执行最佳拼接帧自动选择算法的第4)步,计算出各自SIFT特征向量差值 $Distance$ 的均值 \bar{D} (表1)。

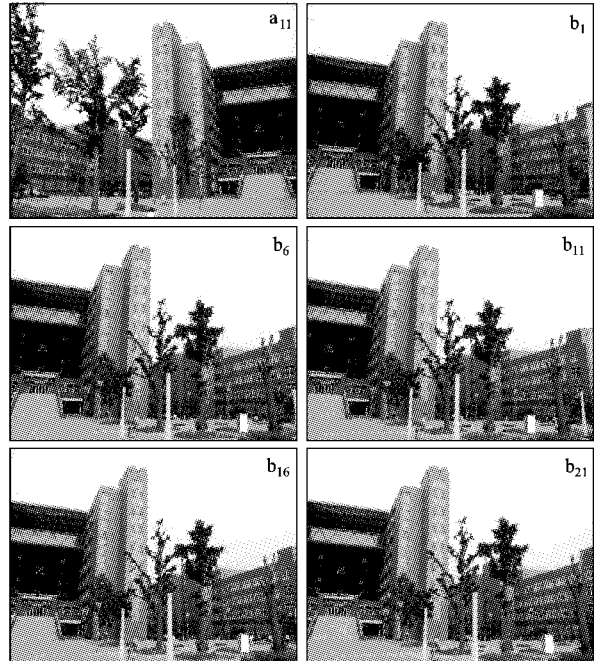


图6 待匹配视频帧

表1 SIFT特征向量差值的算术均值

序号	$a_{11} - b_1$	$a_{11} - b_6$	$a_{11} - b_{11}$	$a_{11} - b_{16}$	$a_{11} - b_{21}$
1	0.091 938	0.308 723	0.299 203	0.064 299	0.147 316
2	0.347 091	0.531 603	0.404 780	0.804 815	0.200 982
3	0.609 051	0.577 963	0.523 089	0.807 055	0.230 429
4	0.804 198	0.626 502	0.902 006	0.810 030	0.340 334
5	0.823 194	0.661 800	1.007 400	1.035 410	0.389 338
6	1.088 480	0.813 615	1.011 580	1.081 830	0.449 707
7	1.129 560	1.310 100	1.020 980	1.126 450	0.539 687
8	1.137 700	1.354 720	1.035 190	1.133 950	0.641 157
9	1.218 340	1.398 400	1.169 390	1.140 460	0.757 360
10	1.478 510	1.483 370	1.190 120	1.186 610	0.771 434
11	1.490 730	1.764 120	1.391 760	1.205 740	0.955 618
12	1.697 140	1.764 220	1.479 830	1.333 660	0.959 722
13	2.266 420	1.900 530	1.677 420	1.423 420	1.150 140
14	2.728 360	2.048 920	3.171 930	1.578 970	1.241 510
15		2.203 990		1.675 710	1.466 630
均值	1.2079 080	1.249 905	1.163 191	1.093 894	0.682 758

从结果看,最小均值 $\min(\bar{D})$ 对应的候选帧为第 b_{21} 帧。根据最佳拼接帧自动选择算法,第 b_{21} 帧候选帧应为基准帧 a_{11} 的最佳拼接帧。图7为 a_{11} 与 b_{21} 拼接得到的全景 $a_{11}-b_{21}$ 。

为了比较拼接的效果,将基准帧 a_{11} 分别与候选帧 b_1 、 b_6 、 b_{11} 、 b_{16} 、 b_{21} 进行拼接,得到5幅全景,图8给出了这5幅全景的拼接处截图。

从拼接处的截图看(图中用线框标志出拼缝的拼接痕迹),全景 $a_{11}-b_{21}$ 的确是5幅全景里的最佳全景,也就说明采用最佳拼接帧自动选择算法筛选出的候选帧 b_{21} 的确是基准帧 a_{11} 的最佳匹配帧。

采取相同方法,对另外两组场景进行了实验。最小均值对应的候选帧分别为第 d_{16} 帧和第 f_{16} 帧。实验结果如图 9 ~ 12 所示。

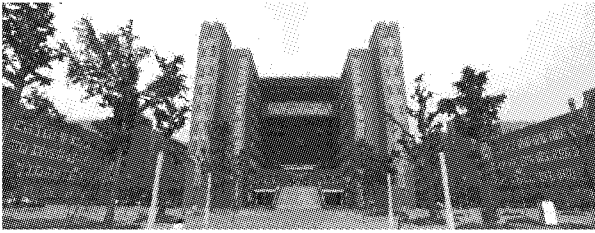


图7 最佳全景图 $a_{11}-b_{21}$

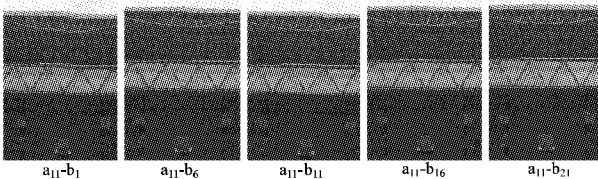


图8 $a_{11}-b_{21}$ 全景图拼接处比较



图9 最佳全景图 $c_{11}-d_{16}$

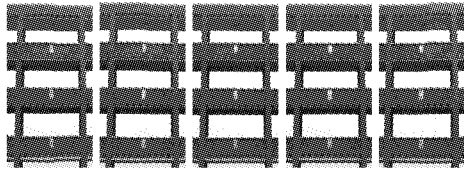


图10 $c_{11}-d_{16}$ 全景图拼接处比较

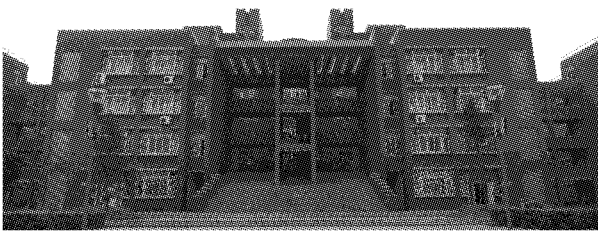


图11 最佳全景图 $e_{11}-f_{16}$

度和砧区特征,设计的基于八方向链码的累积导数和差码特别有利于突显图像轮廓上不规则的凹凸形态。实验证明,本文提出的提取图像层次类特征和图像轮廓的形态特征思想和算法合理有效,类间区分度高,为与传统统计特征一起共同训练高质量分类器奠定了坚实的基础,对同类图像识别问题具有借鉴作用。

参考文献:

- [1] 曾雅靖,曾昌军. 从化一次短时暴雨过程的雷达回波演变特征[J]. 广东气象, 2008, 30(4): 72-74.
- [2] 李庆,黄成亮. 盆地东北部特大暴雨雷达回波分析[J]. 成都信息工程学院学报, 2003, 18(3): 2-6.
- [3] 许焕斌,段英,刘海月. 雹云物理与防雷的原理和设计[M]. 北京: 气象出版社, 2006: 249-263.
- [4] 俞小鼎,周小刚,等. 新一代天气雷达讲义[Z]. 北京: 中国气象局培训中心, 2000: 1-50.

6 结语

本文根据 SIFT 算法,对获得的 SIFT 特征点匹配按照其差值的均值大小进行排序,提出了视频流拼接中基于 SIFT 的最佳拼接帧自动选择算法,解决了具有公共视野的视频流拼接中最佳拼接帧的自动选择问题。实验结果表明,该算法有效。



图12 $e_{11}-f_{16}$ 全景图拼接处比较

参考文献:

- [1] KANG S B. A survey of image-based rendering techniques, CRL 97/4 [R]. Cambridge: Cambridge Research Laboratory, 1997.
- [2] 王俊杰,刘家茂,胡运发,等. 图像拼接技术[J]. 计算机科学, 2003, 30(6): 41-45.
- [3] BROWN M, LOWE D G. Automatic panoramic image stitching using invariant features [J]. International Journal of Computer Vision, 2007, 74(1): 59-73.
- [4] BROWN M, LOWE D G. Recognising panoramas [C]// Proceedings of the Ninth IEEE International Conference on Computer Vision, Washington, DC: IEEE Computer Society, 2003: 1218.
- [5] 朱云芳,叶秀清,顾伟康. 视频序列的全景图拼接技术[J]. 中国图象图形学报, 2006, 11(8): 1150-1155.
- [6] LOWE D G. Object recognition from local scale-invariant features [C]// Proceedings of the 7th IEEE International Conference on Computer Vision. Washington, DC: IEEE Computer Society, 1999: 1150-1157.
- [7] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] 张朝伟,周焰,吴思励,等. 基于 SIFT 特征匹配的监控图像自动拼接[J]. 计算机应用, 2008, 28(1): 191-194.
- [9] 郭三华,方贤勇,罗斌. 一种视频序列的拼接算法[J]. 计算机应用, 2007, 27(11): 2786-2788.
- [10] MOORE A W. An introductory tutorial on KD-trees [R]. Cambridge: University of Cambridge, 1991.

(上接第 2067 页)

- [5] 王萍,董晓凯,贾惠珍. 基于雷达回波反射率图的雹云特征提取[J]. 天津大学学报, 2007, 40(10): 1241-1246.
- [6] Intel Corporations. Open Source Computer Vision Library [Z]. USA: Intel Corporation, 2001: 91-92.
- [7] WANG XIAOLING, XIE KANGLIN. A novel direction chain code-based image retrieval [C]// Proceedings of the Fourth International Conference on Computer and Information Technology (CIT' 04). Washington, DC: IEEE Computer Society, 2004: 190-193.
- [8] LIN H-J, KAO J-T, YEN S-H, et al. A study of shape-based image retrieval [C]// Proceedings of the 24th International Conference on Distributed Computing Systems Workshops (ICDCSW' 04). Washington, DC: IEEE Computer Society, 2004, 7: 118-123.
- [9] 陆宗骥,童韬. 链码和在边界形状分析中的应用[J]. 中国图象图形学报, 2002, 7(12): 22-26.
- [10] 冈萨雷斯. 数字图像处理[M]. 2版. 阮秋琦,译. 北京: 电子工业出版社, 2003: 50-400.