

基于梯度离散余弦变换的视频文字定位

颜俊华¹ 李 丹² 周亚同²

(西南石油大学计算机科学学院 成都 610500)¹ (河北工业大学信息工程学院 天津 300130)²

摘 要 视频文字信息在基于语义的视频分析、检索、提取中占有重要地位。根据视频中文字和背景的灰度变化程度不同,提出一种基于梯度离散余弦变换的视频文字定位方法:先对视频帧进行 $N \times N$ 分块,计算每一块的离散余弦变换系数,然后求出梯度算子的幅值,利用得到的幅值作为块强度进行平滑滤波以及形态学处理,最后对图像进行水平和垂直方向投影,统计字幕条数,并利用文本框标识文字区域,进而达到对视频文字定位的目的。仿真结果表明这种视频文字定位方法对于静态文字和滚动字幕的定位均是可行的,且其算法的运行速度快、效率高,特别是对于笔画较少的文字定位准确,不会出现遗漏现象。

关键词 文字定位,视频帧,梯度

中图分类号 TP391.43 **文献标识码** A

Approach to Video Text Localization Based on Gradient Discrete Cosines Transform

YAN Jun-hua¹ LI Dan² ZHOU Ya-tong²

(School of Computer Science, South West Petroleum University, Chengdu 610500, China)¹

(School of Information Engineering, Hebei University of Technology, Tianjin 300130, China)²

Abstract Video text information plays an important role in semantic-based video analysis, indexing and retrieval. The grade of gray in the video text and background is different, so an approach to video text localization using gradient discrete cosines transform(DCT) was proposed. Firstly, the video frame is divided into $N \times N$ macro-blocks, and we can get the DCT coefficients from every macro-block. Then the gradient operator value considered as block intensity is calculated and we deal with it by using of smooth filter and morphology processing. Lastly, the horizontal and vertical projection can be obtained, the number of text line can be got and the text region is marked with text box. The experimental results show that the proposed method is efficient on the localization of the static and rolling video text, especially the accuracy and integrality of character localization with the less strokes.

Keywords Text localization, Video frame, Gradient

1 前言

随着多媒体技术的广泛应用,视频已经成为了人们生活中不可或缺的一部分。面对海量的视频数据,人们需要方便快捷地从中获取所需要的信息,因此,视频文字信息在基于语义的视频分析、检索、提取中占有重要地位。视频文字分为两类:一类为视频场景本身包含的文字,称为场景文字;另一类为视频后期制作中加入的文字,称为人工文字。

目前,视频文字定位的方法甚多,如基于纹理的方法^[1-4]利用的是字幕区域的特殊纹理特征,在文字检测方面,如文献[1]中使用的是小波变换分解图像,然后提取其子带图像的均值、二阶矩和三阶矩作为纹理特征,最后将样本输入神经网络中进行训练,再将结果经过 16×16 的窗口扫描来确定文本,然而,此类方法计算量较大。为了提高运算速度,可用文献[3]中计算出的离散小波变换模板,快速提取图像的水平、垂直、斜向边缘特征,再与相应动态阈值比较,获得候选文字区

域。基于区域的方法^[5,6,11]是根据文本中的字符具有相同或相似的颜色,提取字符或字符的笔画,然后合并这些字符以形成文字区域,但对于文本字符较小、文字与背景的对比较小的视频帧,定位效果较差。而基于学习的方法^[7-9]将图像的特征作为分类器的输入来训练样本,从而将文字从背景中提取出来。文献[7,8]均使用支持向量机进行分类。文献[7]中将颜色纹理作为分类的特征,选择多项式核函数。文献[8]中为了达到降维的目的,对图像子块提取其“米”字型区域的像素灰度值作为特征,然后利用支持向量机进行分类;对于尺寸较大的文字,采用“金字塔”模型,分为3级,各级分别运用支持向量机进行分类,再恢复得到原尺寸的定位结果,其中选择的是径向基核函数。虽然此类方法具有通用性,但是用学习机对图像进行分割的结果受训练样本集与测试样本集的相似程度的影响。基于字符形状特征的方法^[10,11]能够用于文本字符较小、文本与背景对比较度较小、分辨率较低的图像。文献[11]中运用角点检测来提取候选文本区域,角点检测使用的

收稿日期:2010-11-29 返修日期:2011-04-06 本文受中国博士后科学基金(20090450750)资助。

颜俊华(1971-),女,讲师,主要研究方向为计算机应用等,E-mail: pengwen@vip.sina.com;李 丹(1986-),女,硕士,工程师,主要研究方向为图像和视频处理;周亚同(1973-),男,博士,副教授,主要研究方向为模式识别、图像处理。

是圆形窗口,将窗口内的像素与中心点的像素强度进行比较,如果差值小于门限值,则认为当前像素与中心像素具有相似的强度,所有满足这一条件的像素组成的区域称为核值相似区(USAN),再计算 USAN 的核子与圆形区域的“重心”,将其作为参数来检测角点是否存在。若背景中含有丰富的角点信息,比如一些包含了大量树枝的背景,文献中的方法就会出现误检。因此,该方法受噪声的影响较大。此外,基于离散余弦变换的方法^[12]直接作用于压缩域,处理速度快,但是对于笔画简单的文字,定位效果不理想。

为此,本文研究基于梯度离散余弦变换的方法,即针对人工文字,将梯度信息与离散余弦变换相结合,其中引入梯度计算的目的是为了突出文字的边缘。对于文本字符较小、文字与背景具有相似纹理或颜色、分辨率较低的视频帧,定位效果较好;对于笔画较少的文字定位准确,没有遗漏,保证了文字定位的完整性;对于视频中的动、静态文本都达到了显著的定位效果,而且算法的定位速度快,可以进行实时处理。

2 方法原理

2.1 图像的梯度离散余弦变换

图像函数 $f(x, y)$ 的梯度定义为一个向量:

$$\begin{pmatrix} G_x \\ G_y \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} \quad (1)$$

式中, G_x 和 G_y 为梯度算子,其幅值为:

$$\sqrt{G_x^2 + G_y^2} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (2)$$

为简化运算,可使用式(3)近似计算幅值。

$$\sqrt{G_x^2 + G_y^2} \approx |G_x| + |G_y| \quad (3)$$

下面利用离散余弦变换系数来求解图像函数的梯度。

图像函数为 $f(x, y)$, 其 $N \times N$ 块的二维离散余弦变换系数为 $F(u, v)$, 由离散余弦变换的逆变换得:

$$\begin{aligned} f(x, y) &= \frac{2}{N} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} C(u) C(v) \cos \frac{(2x+1)u\pi}{2N} \cdot \\ &\quad \cos \frac{(2y+1)v\pi}{2N} F(u, v) \\ &= \frac{2}{N} \sum_{x=0}^{N-1} C(u) \cos \frac{(2x+1)u\pi}{2N} \sum_{y=0}^{N-1} C(v) \cdot \\ &\quad \cos \frac{(2y+1)v\pi}{2N} F(u, v) \end{aligned} \quad (4)$$

式中, u 和 v 的取值分别为 $0, 1, \dots, N-1$ 。

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}}, & u, v = 0 \\ 1, & \text{其它} \end{cases}$$

对式(4)求梯度得:

$$\begin{aligned} G_x &= -\frac{2\pi}{N^2} \sum_{x=0}^{N-1} C(u) u \sin \frac{(2x+1)u\pi}{2N} \sum_{y=0}^{N-1} C(v) \cdot \\ &\quad \cos \frac{(2y+1)v\pi}{2N} F(u, v) \end{aligned} \quad (5)$$

$$\begin{aligned} G_y &= -\frac{2\pi}{N^2} \sum_{x=0}^{N-1} C(u) \cos \frac{(2x+1)u\pi}{2N} \sum_{y=0}^{N-1} C(v) \cdot \\ &\quad v \sin \frac{(2y+1)v\pi}{2N} F(u, v) \end{aligned} \quad (6)$$

其中,

$$\begin{aligned} F(u, v) &= \frac{2}{N} C(u) C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cdot \\ &\quad \cos \frac{(2x+1)u\pi}{2N} \cos \frac{(2y+1)v\pi}{2N} \end{aligned} \quad (7)$$

2.2 视频文字定位方法步骤

视频文字定位的流程如图1所示,主要步骤如下。

(1)图像分块计算:现有的压缩视频标准中,大多使用了离散余弦变换,如 MPEG、MPEG-2 和 H. 264/AVC 标准。本文对 MPEG-2 编码数据进行变长码解码、反扫描和反量化,可得到底层像块层的 8×8 二维离散余弦变换系数;由式(5)和式(6)求图像函数的梯度算子,再根据式(3)求得梯度幅值。并将梯度幅值作为图像的块强度信息 I 。

$$I = \sum_{u=0}^7 \sum_{v=0}^7 |G_{x_{u,v}}(i, j)| + |G_{y_{u,v}}(i, j)| \quad (8)$$

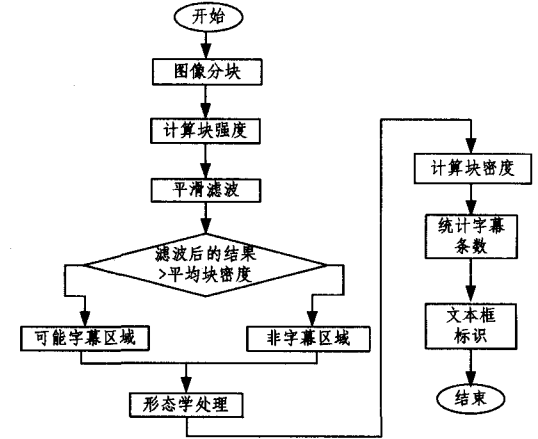


图1 软件流程图

特别对文字与背景相对运动的视频而言,还需要考虑到滚动字幕的出现和消失位置的定位。

第一,在相邻的视频帧之间字幕的相对位置是不变的,可以计算相邻帧之间文本行块强度的相关性,即两帧之间的块强度匹配误差:

$$\begin{aligned} ME(i, j) &= |I_{Prev}(i, j) - I_{Next}(i, j)| \\ i &= 1, 2, \dots, m, j = 1, 2, \dots, n \end{aligned} \quad (9)$$

式中, $I_{Prev}(i, j)$ 和 $I_{Next}(i, j)$ 分别为相邻两帧字幕的对应位置的块强度, m 和 n 为字幕区域在水平和垂直方向上的块数。

第二,将相邻两帧文字区域的匹配误差与平均匹配误差:

$$ME_{Aver} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n ME(i, j) \quad (10)$$

相比较,判断是否有文字的出现或者消失。

(2)图像平滑滤波:对图像的块强度信息进行基于区域的平滑滤波,可以用空间域卷积运算方式来描述,把平均化处理看作一个作用于图像上的低通滤波器,本文利用均值滤波器对图像进行处理,其模板大小由所选邻域大小决定,一般来说, 3×3 的小邻域效果就很好了,邻域取得过大,会使灰度突变的边缘图像变得模糊,从而不利于文字的定位。将滤波后的结果与平均块密度对比,初步划分出可能的字幕区域和非字幕区域,便于后续对文字的标识。

(3)形态学处理:为了剔除可能的文字区域的背景和将漏掉的文字信息补充回来,对可能的文字区域利用结构元素进行形态学处理,本文使用的是水平线性结构元素,这种结构元素对于视频中水平排列的文本行具有较好的处理效果。结构

元素的设置均由实验结果所决定,详见 3.1 节。

(4)块密度计算:对图像进行水平和垂直方向的投影,求得块密度。以水平方向为例,其块密度为:

Hor_{dens}(i)=∑_{i,j∈Area}I'(i,j)/Num(i) (11)

式中,I'表示滤波之后的块强度,即可能的文字区域的块强度;Num表示水平方向的块数;Area表示可能的文字区域。

(5)字幕条数统计与文本框标识:由于文字的块密度大于背景的块密度,因此可将文字与背景更准确地分开,即划分文本行,找到精确的文字位置,统计字幕的行数,并用文本框进行标识,最终获得理想的文字定位结果。

本文在离散余弦变换的基础上引入了梯度算子,因此,在定位的过程中能更加突出文字的边缘信息,进而可以减少复杂背景中出现文字的遗漏以及错误的定位。

3 仿真与应用

3.1 结构元素的设置对定位结果的影响

视频中存在的人工文字大多是以水平方向排列的,因此在对视频文字进行形态学处理时,本文选择了水平方向的线性结构元素。使用了行数小于列数的全 1 矩阵作为其结构元素,来体现水平线性的思想。

表 1 为一段含有字幕的视频,经过处理之后得到 476 个含有文字的视频帧,选择不同尺寸的结构元素对其进行定位。可以看出,腐蚀和膨胀过程的结构元素形式相同,但由于需要将可能文字区域中的背景尽量剔除,同时又不丢失文字信息,因此在尺寸上膨胀操作的结构元素要比腐蚀操作的略大。

表 1 结构元素尺寸对定位结果的影响

		腐蚀结构元素尺寸					
		1×2	1×4	1×6	1×8	1×10	1×12
膨胀 结构 元素 尺寸	1×2	遗	2.5	—	—	—	—
		虚	45.4	—	—	—	—
	1×4	遗	1.3	1.3	—	—	—
		虚	40.3	34.0	—	—	—
	1×6	遗	1.3	0.6	1.3	—	—
		虚	29.0	34.0	26.5	—	—
	1×8	遗	1.3	0.6	0.6	2.5	—
		虚	23.9	31.5	27.7	26.5	—
	1×10	遗	1.3	1.3	0.6	1.3	3.8
		虚	22.7	27.7	29.0	27.7	26.5
	1×12	遗	1.3	0.6	0.6	1.3	3.8
		虚	21.4	25.2	27.7	27.7	27.2

注:以上数据均为所占百分比

由于篇幅所限,这里仅对腐蚀与膨胀结构元素分别为行数为 1、列数大于行数且逐渐递增的尺寸给出了其对视频文字定位的影响。经过实验可得,当腐蚀结构元素为 1×7,且膨胀结构元素为 3×11 的全 1 矩阵时,遗漏率为 2.5%,虚警率为 6.3%,定位效果较好。

表 1 中,“遗”表示遗漏率,“虚”表示虚警率,“—”表示在尺寸上膨胀结构元素尺寸小于腐蚀结构元素的部分,不予讨论。

3.2 视频文字定位的仿真分析

为了验证本文方法的有效性,下面针对人工文字定位进行仿真分析。

在视频中,提取一帧带有人工文字的彩色图像作为待定位的原始图像,如图 2 所示。此图像的特点是文字相对于背

景来说是静止的,字幕区域与背景灰度对比明显,且文字不仅仅是单行,还包括标点符号和笔画较少的文字。



图 2 原始图像

首先对此原始图像得到的 8×8 离散余弦变换系数,利用式(5)和式(6)计算出梯度算子,进而根据式(8)求得块强度。然后进行基于区域的平滑滤波,利用块强度信息与滤波器模板算子卷积运算,目的是滤除大部分背景,得到可能的文字区域,滤波后结果如图 3 所示。由于中心点和邻域重要程度的不同,同时,为了保持输出图像仍在原来的灰度值范围内,应使整个模板算子的平均数为 1,且系数均为正数。因此,滤波器模板算子为

Filter=1/16 * [1 2 1; 2 4 2; 1 2 1]

图 4 为对可能的文字区域进行形态学处理^[13],即腐蚀和膨胀操作,去除了可能的文字区域中的背景,且更大程度地包含了所要定位的文字信息。



图 3 滤波处理后的可能文字区域



图 4 形态学处理

经过水平和垂直方向的投影得到块密度后,根据块密度可以准确定位文字。由于选取的图像有两行文字,因此需要划分文本行,用文本框进行标识,最终达到图 5 中的效果。由图 5 可以看到,对文字和标点的定位均不存在遗漏的现象,且文本行划分清晰、明确,效果较好。因而,验证了本文方法的有效性。

为了验证本文方法的通用性,下面针对各种情形进行仿真验证。

图 6 为对英文进行处理的效果图,可见本文方法对多行的英文字幕定位也较为准确。



图 5 最终文字定位图



图 6 对英文定位

图 7 为对位于复杂背景中的文字定位结果图,从图中看到,背景与文字的颜色相近,对比不十分强烈,但定位效果较好。



图7 复杂背景下的定位

图8为3种方法的对比图,图8(a)为引入梯度计算与图8(b)中仅使用离散余弦变换方法^[12]进行对比,由图可知,梯度计算更强调文字边缘的变化,因此,对位于文字开头位置的简单笔画文字的定位准确,且除去了小部分位于文字末尾的背景区域。

图8(c)为利用文献[8]中的方法提取图像子块“米”字型区域的像素灰度值之后,采用支持向量机对其进行分类所得到的定位图。为了减少计算量,降低了图像的分辨率,因此,显示出来的结果较为模糊。但可以清楚地看到,图中存在着漏检问题。



(a)采取本文方法的文字定位图

(b)离散余弦变换方法的定位图

(c)支持向量机方法的定位图

图8 3种方法的对比图

表2为对文字与背景相对静止的一段视频进行处理得到665帧图像后,分别采用本文方法、仅使用离散余弦变换方法^[12]和采用支持向量机方法^[8]对其进行定位的结果。

表2 文字定位算法的结果对比

方法	准确率	遗漏率	虚警率	运行时间
支持向量机 ^[8]	86.1%	5.3%	9.0%	6.4s
离散余弦变换 ^[12]	93.3%	5.5%	1.2%	3.5s
本文方法	95.6%	3.1%	1.1%	1.3s

其中运行时间是在 Intel(R) Core(TM)2 CPU, 512M 内存上测得的数据。可见,本文的方法对静态文字的定位准确率高,即尽量保存了文本的完整性,在处理速度上,文献[8]中支持向量机的分类及优化过程耗时较多,因此与其相比,本文方法运行时间较短;与文献[12]相比,本文方法漏检率有所降低,但虚警率改善不大。

为了验证本文方法在动态文字定位中的应用效果,选取一段具有动态字幕的实际视频,共322帧,提取了其中的连续帧,利用本文的方法对其进行了文字定位,结果显示如图9所示。



图9 动态字幕定位结果

由图9可知,连续帧中的滚动文字均用文本框进行了标识,效果显著。值得注意的是这段视频中不仅仅包含动态文字,同时还包括了部分静态文字,表3统计了两类文字定位的准确性。

表3 视频中两种文字的定位结果

文字类型	准确率	遗漏率	虚警率
静态文字	93.5%	4.7%	1.8%
动态文字	83.9%	9.6%	6.5%

由表3可知,本文的方法对于视频中的静态文字和动态文字均具有较好的定位效果,但对于动态文字定位的准确率还有待提高。

结束语 本文研究的基于梯度离散余弦变换的视频文字定位方法是将梯度计算引入离散余弦变换中,通过各种仿真实验与对比分析表明了这种方法在视频文字定位中效果较好,对于不同的文字类型具有通用性,在文字的完整性与处理速度方面较文献[8,12]有所提高。

参考文献

- [1] Li H, Doermann D, Kia O. Automatic text detection and tracking in digital video[J]. IEEE Transactions on image processing, 2000, 9(1):147-156
- [2] Mao W G, Chung F L, Lam K K M, et al. Hybrid Chinese/English text detection in images and video frames[C]//Proceedings of 16th International Conference on Pattern Recognition, 2002 (3):1015-1018
- [3] Liang Chung-wei, Chen P-Y. DWT Based Text Localization[J]. International Journal of Applied Science and Engineering, 2004, 2(1):105-116
- [4] Bouaziz B, Mahdi W, Ardabilain M, et al. A New Approach for Texture Features Extraction; Application for Text Localization in Video Images[C]//2006 IEEE International Conference on Multimedia and Expo, 2006:1737-1740
- [5] Pei S C, Chuang Y T. Automatic text detection using multi-layer color quantization in complex color images[C]//2004 IEEE International Conference on Multimedia and Expo, 2004:619-622
- [6] 何家颖,黎绍发.一种复杂背景图像文字分割算法[J].模式识别与人工智能,2005,18(2):148-153
- [7] Lee C. Automatic text detection and removal in video sequences [J]. Pattern Recognition Letters, 2003, 24(15):2607-2623
- [8] Cheng Zhi-guo, Liu Yun-cai. An automatic system for text location and extraction in digital video based using SVM[A]//International Conference on Signal Processing Proceedings (ICSP2004)[C]. Vol 3, 2004:2592-2595

(下转第293页)

深入讨论。

4 实验结果

采用本文的循环分布算法,可以充分地将循环内可向量化与不可向量化的语句放在不同的循环中。而当前主流的编译器都做不到这一点,也就无法开发这种部分语句可向量化的循环。上一节讨论的例 9,在 alpha 处理器上(256bit)进行试验,采用本文的分布方法与采用传统的分布算法,加速比可达到 1.56。

图 9 是基于 open64 实现的传统循环分布和基于语句依赖的循环分布在 alpha 机器上的测试结果,显示了循环分布算法改进前后的性能对比。改进后程序平均加速比为 1.11,相对于原来的向量化方法对测试集的性能提高平均达到 4%。从测试可以看出,使用本文的方法通过分析程序中语句的数据依赖性,可分布出更多可向量化的循环,生成向量化代码,可更充分地利用向量化功能部件。但是在分布时可能会引入一些额外开销,造成部分测试用例效果不明显。

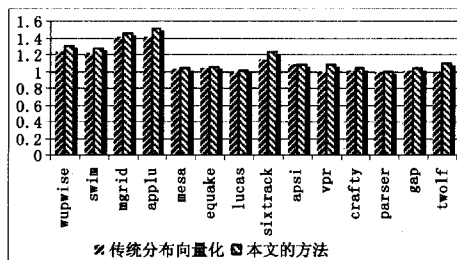


图 9 传统分布算法和基于语句的分布算法对 spec2000 程序的加速比

结束语 本文分析了传统向量化中循环分布方法的不足之处,从依赖关系角度来分析语句可否向量化,又从语句能否向量化的角度来进行循环分布。此循环分布方法主要是对传统的循环分布方法的改进,增加了对单个语句可否向量化和依赖环语句可否向量化的分析,对可向量化的依赖环采用语句分裂的方法来实现,这一改进可以从依赖关系的角度进行彻底的循环分布。虽然此循环分布方法是对传统的循环分布的改进,但还存在一些问题值得研究,主要是以下两个方面:

1)通过本文的循环分布方法对可向量化的依赖环进行分析时,为了去除反依赖边,需要对反依赖边的源点语句进行语句分裂,此过程会增加语句的数目。若整个依赖环内的全部语句均可向量化,增加的语句在进行 SIMD 时会消除掉,不会带来附加的操作。但是,若依赖环中有可向量化的语句又有不可向量化的语句,进行循环分布时,那些为了消除反依赖边

进行语句分裂而产生的临时语句会被分布在不同的循环中,这样就无法不执行这些临时操作,会产生额外的开销,这也是有些例子测试效果不好的原因。而如果产生的开销大于 SIMD 向量化的收益,循环分布就不应该做,因此需要对此进行循环分布的收益分析。

2)此循环分布方法目前只是从依赖关系的角度来考虑的,在实际应用中,循环能否向量化的影响因素有很多,比如循环内有无函数调用、指令跳转、循环内语句是否连续、对齐等,这些方面均可影响循环的向量化能力。循环分布应该从多个方面和角度出发,且在实际应用中自动分析这些情况均很复杂,自动进行循环分布可能并不能达到我们预期的目标,因此可以考虑采用交互式的循环分布方法,目前该方法还在进一步的研究中。

参考文献

- [1] Allen R, Kennedy K. Optimizing Compilers for Modern Architectures—A Dependence-Based Approach[M]. US: Morgan Kaufmann Publishers, 2001
- [2] Larsen S, Amarasinghe S. Exploiting superword level parallelism with multimedia instruction sets[C]//Proc of the ACM SIGPLAN Conference on Programming Language Design and Implementation. 2000;145-156
- [3] Stewart J. An investigation of SIMD instruction sets[R]. University of Ballarat School of Information Technology and Mathematical Sciences. 2005
- [4] Kennedy K, Mckinley K S. Loop distribution with arbitrary control flow[C]//Proceedings of the 1990 Conference on Supercomputing. 1990;407-416
- [5] Kennedy K, Mckinley K S. Loop distribution with multiple exits[C]//Proceedings of the 1992 ACM/IEEE Conference on Supercomputing. Minneapolis, Minnesota, United States, 1992; 204-213
- [6] Lengauer T, Tarjan R E. A fast algorithm for finding dominators in a flowgraph[J]. ACM Transactions on Programming Languages and Systems, 1979;1(1):121-141
- [7] Tarjan R E. Depth first search and liner graph algorithms[J]. SIAM Journal of Computing, 1972,1(2):146-160
- [8] 曾扬. 循环分布及依赖关系破除的优化问题[J]. 计算机学报, 1993,16(6)
- [9] 沈志宇,胡子昂,等. 并行编译方法[M]. 北京:国防工业出版社, 2000
- [10] Guo Ge, Jin Jin, Ping Xi-jian, et al. Automatic Video Text Localization and Recognition[A]//Proceedings of Fourth International Conference on Image and Graphics (ICIG 2007)[C]. Chengdu, China, August 2007;484-489
- [11] Qian Xue-ming, Liu Gui-zhong, Wang Huan, et al. Text detection, localization, and tracking in compressed video[J]. Signal Processing: Image Communication, 2007,22(9):752-768
- [12] 胡学龙,许开宇. 数字图像处理[M]. 北京:电子工业出版社, 2006;169-179
- [9] Gllavaj J, Ewerth R, Freislebe B. Text detection in images based on unsupervised classification of high-frequency wavelet coefficients[C]//Proceedings of the 17th International Conference on Pattern Recognition. 2004(1):425-428
- [10] Hontani H, Koga T. Character extraction method without prior knowledge on size and position information[A]//Proceedings of the IEEE International Vehicle Electronics Conference [C]. 2001;67-72

(上接第 274 页)