

行人视频检测中阴影检测与去除方法设计*

吕秋霞

(五邑大学 信息工程学院, 广东 江门 529020)

摘要: 设计了基于模糊神经网络的目标本体与阴影的自动分类器, 将光谱特征值作为神经网络输入, 用遗传算法对网络权值和模糊规则实时更新, 并结合图像的空间特性对去除结果进行修正。实验结果表明, 应用该方法能有效去除阴影, 并可提高行人视频检测的准确率。

关键词: 阴影去除; 模糊神经网络; 行人视频检测; 遗传算法

中图分类号: TP391

文献标识码: A

文章编号: 1674-7720(2010)19-0028-04

Design of the shadow detection and removal in pedestrian video detection

LV Qiu Xia

(School of Information Engineering, Wuyi University, Jiangmen 529020, China)

Abstract: A fuzzy-neural network auto-classifier of the moving target and its shadows is designed. The spectral feature value is input to the network. The weight and fuzzy rule is update real-time by genetic algorithm. And the spatial feature of image is taken into account to amend the result. Experimental results show that the application of this algorithm can improve the accuracy of pedestrian video detection.

Key words: shadow removal; fuzzy-neural network; pedestrian video detection; genetic algorithm

行人是城市交通系统的主要参与者, 保障行人安全和减少其对机动车的干扰是城市交通系统建设的重要目标, 因此对行人交通的研究也越来越受到重视。行人交通研究的主要问题^[1-2]包括行人检测、目标跟踪和行为分析。基于视频的行人检测与传统的红外检测、GPS检测、激光检测等方法相比, 具有不破坏路面、维护方便、实时性好、可检测的参数多等优点, 成为实时交通信息采集和处理技术的发展方向。

视频图像中的阴影会影响行人的检测与跟踪^[2-4], 因为阴影的存在会造成检测目标的变形、合并、甚至丢失, 使得目标定位及计数不准确。近年来, 科研工作者对图像中的阴影去除问题进行了大量研究, 在这些研究方法中, 考察的图像特征主要有三种: 光谱特征、空间特征和时间特征^[5]。光谱特征针对像素点, 如灰度值、颜色信息等^[6], 根据当前图与背景图的色差、亮度差值等判断像素点是否为阴影, 或者对图像进行变换得到光照无关

图^[7]进而去除阴影; 空间特征是针对某一区域或某一帧图像, 根据检测到的图像的轮廓、纹理、边缘等信息判断是否为阴影, 如利用图像的轮廓特征^[8-9], 找到目标与阴影的边界线, 对本体和阴影粗分, 再建立阴影像素的高斯模板进行细分, 既减少了计算量又能达到较好效果; 时间特征一般都是与前两种特征结合使用, 可以用于对阴影方向或运动速度的估算等, 以进一步提高阴影去除效果。

本文提出一种新的基于 YUV 颜色空间的阴影去除算法, 因为很多摄像头的输出信号采用 YUV 颜色空间, 与基于 RGB 颜色空间的处理方法相比, 省去了图像颜色空间转换的步骤, 能提高处理速度。在图像特征上, 本文结合像素点的光谱特征与图像整体的空间特征, 首先通过亮度差和色差对像素点进行判断, 再利用目标本体与阴影只相接不相交的空间特征, 对去除结果进行修正, 使其阴影去除效果更好。同时, 为了使算法适应光照、场景等的变化, 采用模糊神经网络进行目标本体与阴影的分类, 用遗传算法对网络参数和权值进行自适

* 基金项目: 广东省自然科学基金(8152902001000014); 广东省高等学校自然科学重点研究项目(052025)

应调整,以提高算法的鲁棒性。

1 YUV 颜色空间

在色彩学上,为了可以准确定量地描述颜色,将色彩定义为三大属性:“Y”表示明亮度,即灰度值;“U”和“V”表示色度,作用是描述图像色彩及饱和度,用于指定像素的颜色。根据美国国家电视制式委员会 NTSC 制式的标准,白光的亮度用 Y 来表示,色差 U、V 由 B-Y、R-Y 按不同比例压缩而成,与红、绿、蓝三色光的关系可用式(1)描述,这也是常用的转换公式。YUV 到 RGB 的转换公式则如式(2)所示。

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.14 \\ 1 & -0.39 & -0.58 \\ 1 & 2.03 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix} \quad (2)$$

式中, R、G、B 的取值范围均为 0~255。通常摄像机的数据以 RGB、YUV 或 YCrCb 的格式输出。采用 YUV 颜色空间的重要性是它的亮度信号 Y 和色度信号 U、V 是分离的。目前有很多种颜色空间可以将图像的色度分量和亮度分量区分开来,如 HSV 颜色空间,但是这种转换较为复杂,对于大型图像非常耗时,并且在亮度值和饱和度较低的情况下,采用 HSV 颜色空间计算出来的 H 分量是不可靠的。

在 YUV 颜色空间中,如果只有 Y 信号分量而没有 U、V 信号分量,则这样表示的图像就是黑白灰度图像。除去亮度信号后,由 U 和 V 单纯表现出色度。因此,如果要将 U 与 V 色差信号用色相及饱和度来表示,必须从含有三维空间的色点 P 投影到 U-V 平面的 P' 点,如图 1(a)所示。U-V 平面投影法在受到不稳定光源亮度的扰动时,对于目标色度有较大的精确性且不易辨识错误,但是当光源色温变化过大时,其饱和度和色相的增减变化不易掌握。因此,如果需要判定两个任意色点是否为同一色度时,必须确定其色相与饱和度都是相等的。如图 1(b)所示,对两个色点 P₁ 与 P₂,当其与 U 轴的夹角 $\alpha_1 = \alpha_2$ 时,表示色相等;当其与原点的距离 $L_1 = L_2$ 时,表示饱和度相等。当两者都相等时,表示色度完全相同。

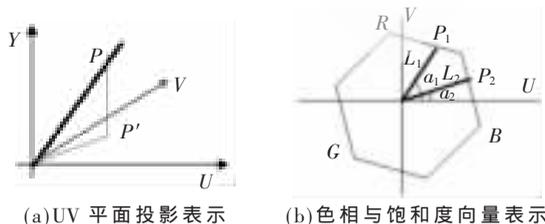


图 1 YUV 颜色空间模型

对于光源亮度的不稳定因素,只要光源亮度不是极值(极亮或极暗),对于相似颜色,如深蓝色和蓝色,就有

相近的色度关系。对运动目标本体和阴影,也有相近的色度,但亮度值差别较大,可通过计算当前图与背景图之间的亮度差值和色差来进行阴影去除。

2 阴影检测算法

2.1 运动区域提取

采用减背景法。假设 $I_k(x, y)$ 和 $B_k(x, y)$ 分别代表当前图像及其背景图的 Y 值,用式(3)获得的差分图像:

$$I_D^k(x, y) = \begin{cases} 0, & \text{当 } |I_k(x, y) - B_k(x, y)| \leq T \\ 1, & \text{其他} \end{cases} \quad (3)$$

式中, T 是预先设定的阈值。为克服光照变化所带来的影响,背景图像应实时更新,利用前 N 帧的背景图像和前一帧图像求和得到:

$$B_k = \frac{1}{N} \sum_{i=1}^N I_{k-i} \approx \frac{L-1}{L} B_{k-1} + \frac{1}{L} I_{k-1} \quad (4)$$

2.2 阴影的去除算法

在提取的运动区域中,包括运动目标及其阴影。根据前文所述在 YUV 颜色空间中阴影与背景的色度相似,而亮度低于背景的特征,计算当前图像与背景图像的 Y、U、V 差值 ΔY 、 ΔU 、 ΔV 以及当前图像与背景图像的色相差值 $\Delta\alpha$ 。

根据图 1(b), $\alpha = \arctan \frac{V}{U}$, 令 $x = \frac{V}{U}$, 为便于计算,将 $\arctan x$ 函数用一级泰勒级数如下展开:

$$\alpha = \arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} = x - \frac{1}{3} x^3$$

当 $x < 1$ 时, $\alpha = \arctan x \approx x$ (这里计算差值 $\Delta\alpha$)。

阴影的判断准则如下:

(1) 若 $\Delta Y << Y_{\min}$, 则有可能为阴影, 进入(2)进一步判断; 若 $\Delta Y > Y_{\min}$, 表示像素差异较大, 应为运动目标。

(2) 若 $\Delta U < \Delta V < \varepsilon$ 或 $\Delta V < \Delta U < \varepsilon$, 则表示当前图像与背景图像的色度相近, 可视为阴影, 否则进入(3)进一步判断。

(3) $\Delta U < \Delta V$ 且 $\Delta(\frac{U}{V}) < \Delta\alpha$, 或 $\Delta V < \Delta U$ 且 $\Delta(\frac{V}{U}) < \Delta\alpha$,

选择 U 和 V 中较大者作为分母, 比值代表 $\Delta\alpha$ 的 \arctan 值; 若条件成立, 表示当前图像与背景图像的色度相近, 可视为阴影, 否则判定为运动目标。

上述准则在应用中, 要注意 Y_{\min} 、 ε 和 $\Delta\alpha$ 等阈值的选取, 因为这对判断结果的影响较大。要找到合适的阈值^[9], 需要对视频资料进行大量的仿真实验, 这需要花费很长时间, 而且根据现有资料得到的阈值不能根据场景、光照等的变化自适应进行调整, 实用价值不大。

针对上述问题, 将模糊神经网络^[10-11]融入到目标本体与阴影的分类中是很好的解决方法。它利用神经网络的自学习能力和自适应能力来调整模糊规则和隶属度函数, 通常对神经网络的训练采用 BP 算法, 但是 BP 算法具有收敛性依赖初始条件, 容易陷入局部极小值等问

题。因此,本文采用遗传算法优化模糊神经网络的结构和参数,并自动获得最优的模糊规则,使网络能自动适应场景与光照的变化。

2.3 模糊神经网络

模糊神经网络的结构如图2所示。

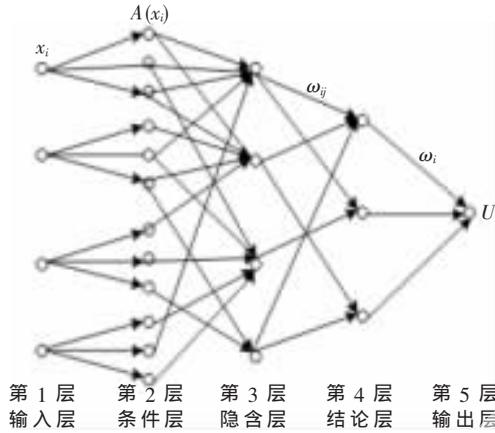


图2 模糊神经网络结构图

把提取的光谱特征值($\Delta Y, \Delta U, \Delta V, \Delta \alpha$)作为神经网络的输入(输入层4个节点)。第2层为条件层,进行“模糊化”处理,将输入的特征值用大、中、小三种语义来表示,采用如式(5)所示高斯型的隶属函数模糊化。

$$A(x_i) = \exp\left[-\left(\frac{x_i - C_{ij}}{b_j}\right)^2\right] \quad (5)$$

式中, x_i 表示输入变量, C_{ij} 、 b_j 分别表示隶属函数的中心和宽度,是全局性参数。第3层为隐含层(也称为可变规则层),每个节点代表一个模糊规则,模糊规则形式如下:

$$\text{OR}_{j=1}^m (\text{IF} (\text{AND}_{i=1}^n (x_i \text{ IS } A_{ij})) \text{ THEN } Y \text{ IS } \omega_j / z_i), \quad (6)$$

$$i=1, 2, \dots, l$$

$$\mu_i = \Pi A(x_i)^j$$

式中, m 为模糊规则数, n 为输入变量个数, l 为输出变量模糊子集个数, z_i 为输出结论的支集值, ω_j 为第 i 个模糊规则对于输出的作用权值。用遗传算法对模糊规则数优化,模糊规则数最大值 $M=3^4=81$ 。第4层为结论层,其节点数为输出变量的模糊子集个数,与第3层每个节点的连接系数为 ω_{ji} 。结论为输出对应每个模糊子集的隶属度

$u_i = \sum_{j=1}^l \omega_{ji} \mu_i \quad i=1, 2, \dots, l$ 。第5层(最后一层)为输出层,采用式(7)的加权平均法对识别结论进行“去模糊”处理。

$$U = \frac{\sum_{i=1}^l u_i \cdot z_i}{\sum_{i=1}^l u_i} \quad (7)$$

式中, u_i 表示对第 i 个模糊子集的隶属度, z_i 表示输出结论的支集值。最后,对输出结果进行二值化表示,1表示

目标本体,0表示阴影。当结果小于0.05时,认定为阴影;结果大于0.95时,认定为目标本体,当结果在0.05~0.95之间时,认为无法判断。

2.4 网络自适应优化

用遗传算法对模糊神经网络的结构和参数进行优化。网络的结构优化指确定第3层节点数、第3层和第2层的连接数、以及第3层和第4层的连接数和连接权值。网络的参数优化包括输入变量的隶属度函数的中心参数和宽度参数、输出变量的隶属函数支集值。

种群的每个个体由网络结构和网络的输入隶属度函数参数和结论参数组成,其长度为结构基因长度+参数基因长度。结构基因中“连接”采用二值的编码,“0”表示没有连接,“1”表示有连接,连接权值 ω_{ji} 用(0~1)之间实数编码。输入的隶属度参数 C_{ji} 和 b_j 、结论参数 z_i 采用实数编码。一个染色体对应一种模糊神经网络结构及其参数。初始种群中包含着对应于最大节点数及输入变量和输出变量在其变化范围内均匀划分模糊子集的个体,其余个体随机产生。将根据经验得到的规则集及输入输出模糊划分对应的向量选入初始种群。

遗传操作包括复制、交叉、变异。为简化运算实现实时处理,本文仅采用变异操作。二值编码按一定的概率将控制基因串中的位从0变异为1,或者从1变异为0。实数编码按下式突变:

$$S_{ji} = S_{ji} + \alpha \frac{1}{f} N(0, 1)$$

式中, α 表示进化率, f 表示网络的适应度函数, $N(0, 1)$ 表示均值为0、方差为1的正态分布随机值。

网络的适应度应综合考虑网络误差和复杂度。网络误差函数采用均方差形式,网络复杂度用网络的第3层的节点个数以及第2、3层之间和第3、4层之间的连接个数之和来评价。在网络误差寻优的同时控制网络复杂度,从而得到最优的模糊规则。网络的适应度函数形式如下:

$$f(i, t) = \alpha E(i, t) + \beta H(i, t), \quad \alpha, \beta > 0 \text{ 且 } \alpha + \beta = 1$$

式中, $f(i, t)$ 表示第 t 代的第 i 个网络个体的适应度; $E(i, t)$ 表示第 t 代的第 i 个网络个体的网络误差; $H(i, t)$ 表示第 t 代的第 i 个网络个体的网络复杂度。

2.5 空间特征

考虑到图像中阴影和目标本体相接但互不相交,对于不能判断的像素及初步识别结果,按下述规则进行判断和修正:(1)如果周围像素点多数为“阴影”,则该点是“阴影”。(2)如果周围像素点多数为“目标”,则该点是“目标”。(3)如果周围像素点多数是目标而被判断为“阴影”,则改判断为“目标”。(4)如果周围像素点多数是阴影而被判断为“目标”,则改判断为“阴影”。这里的多数是指相邻8个像素点中5个以上。

3 实验结果和分析

图3、图4是室外拍摄的视频序列的处理结果,视



图3 视频1第154帧图像处理结果



图4 视频1第363帧图像处理结果

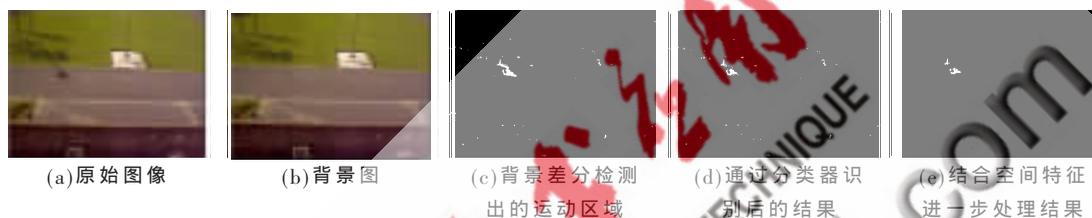


图5 视频2第503帧图像处理结果

频序列共 2 571 帧,单帧图像大小为 354×288 ,图 3 是第 154 帧图像,图 4 是第 363 帧图像。

童车在图 3 中作为背景被提取出来,而在图 4 中成为前景。与图 3 相比,图 4 中光照有较大变化,图 3(d)、图 4(d)、图 5(d)表明模糊神经网络分类器能有效地进行阴影去除。由图 5(d)可见,通过阴影去除,行人能被分隔开来,这样有利于提高视频检测的准确率。

表 1 是对在不同路口拍摄的行人视频进行行人检测的结果,进行阴影去除后视频检测的平均准确率由 61.52% 提高到 80.15%。

表 1 阴影处理前后视频检测准确率比较

视频	实际行人数	阴影去除后视频检测		阴影去除前视频检测	
		检测结果	准确率/(%)	检测结果	准确率/(%)
视频 1	519	438	84.39	345	66.47
视频 2	2 685	2 309	85.99	1 796	66.89
视频 3	5 046	3 865	76.60	2 934	58.15
总计	8 250	6 612	80.15	5 075	61.52

本文给出了一种新的阴影去除算法,该算法以 YUV 颜色空间为基础,用模糊神经网络分类器识别对像素点提取的光谱特征是否为阴影,网络的结构和参数采用遗传算法进行实时更新,最后结合运动目标与阴影的空间特征对分类结果进行修正。实验表明,该方法能适应光照、场景的变化,通过阴影去除能明显提高行人视频检测的准确率。

本文关于行人视频检测的研究尚处于起步阶段,对视频检测中的遮挡问题、运动描述和行为理解问题还在

进一步研究中。

参考文献

- [1] 邵春福.行人交通的视频检测方法综述[J].交通运输系统工程与信息,2008,8(4):23-29.
- [2] SEITNER F H. Fast pedestrian tracking based on spatial features and color [R]. Computer Vision Winter Workshop, 2006:1-6.
- [3] 马国峰.基于 YUV 颜色空间的视频运动检测[J].计算机工程与设计,2008,29(14):3700-3703.
- [4] BROGGI A. A multi-resolution approach for infrared vision-based pedestrian detection [R]. IEEE Intelligent Vehicles Symposium, 2004.
- [5] PRATI A. Detecting moving shadows: algorithms and evaluation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(7):918-923.
- [6] YOON J J. Shadow flash: an approach for shadow removal in an active illumination environment[C]. BMVC, 2003:636-645.
- [7] 刘富强.基于光照无关图的阴影去除方法[J].中国图像图形学报,2007,10(12):1837-1840.
- [8] HSIEH J W. Shadow elimination for effective moving object detection by Gaussian shadow modeling [J]. Image and Vision Computing, 2003,21:505-516.
- [9] SCHREE O. Fast and robust shadow detection in videoconference applications [C]. Symposium on Video Proces. and Multimedia Comm, 2002.

- [10] RAFIEE A. Neuro-fuzzy system in sonographical image enhancement [C]. the 6th WSEAS Int. Conference. on signal Processing Computational Geometry & Artificial Vision, 2006:95-100.
- [11] LO B P L. Neuro-fuzzy shadow filter[M].LNCS, 2002:

381-392.

(收稿日期: 2010-05-11)

作者简介:

吕秋霞,女,1978年生,硕士,讲师,主要研究方向:智能交通控制。

