

基于区域 CSIFT 特征的目标识别算法

罗 磊^a, 范彩霞^b

(西安理工大学 a. 后勤管理处; b. 印刷包装工程学院, 西安 710048)

摘 要: 无重叠多摄像机监控系统在使用局部形状特征进行目标识别时, 会忽略颜色信息, 且对光学变换具有不稳定性。为此, 提出一种基于区域彩色尺度不变特征变换的目标识别算法。通过修正的双色反射模型, 提取对光照特性和物体几何特性具有不变性的颜色特征, 得到尺度不变特征变换描述子, 利用颜色特征和形状特征建立目标模型。实验结果表明, 该算法对刚体和非刚体目标的识别都能取得较好的效果。

关键词: 多摄像机; 目标区域检测; 目标识别; 反射模型; 颜色不变量

Object Recognition Algorithm Based on Region CSIFT Feature

LUO Lei^a, FAN Cai-xia^b

(a. Department of Logistics Management; b. Faculty of Printing and Packaging Engineering, Xi'an University of Technology, Xi'an 710048, China)

【Abstract】 In non-overlapping multi-camera surveillance system, using the local shape features for target identification is unstable to optical transformation and ignores the color information. To solve this problem, an object recognition algorithm based on region Colored Scale Invariant Feature Transform(CSIFT) is proposed. The color feature is extracted by modifying dichromatic reflection model, which is invariant to light conditions and geometry character, the Scale Invariant Feature Transform(SIFT) descriptor is generated on the color invariant feature. The object appearance model is established by color and shape feature. Experimental results show that this algorithm can achieve a better match results for rigid and non-rigid object recognition.

【Key words】 multi-camera; object region detection; object recognition; reflection model; color invariant

DOI: 10.3969/j.issn.1000-3428.2012.19.049

1 概述

无重叠视域多摄像机监控环境下的目标跟踪成为视频监控研究中的重要研究内容。目标识别是多摄像机监控系统应用中最重要和基础的问题之一, 由于不同摄像机中目标的颜色、形状和观测距离、观测角度等方面的巨大差异, 以及光照等环境条件的不同, 因此目标识别需要在保障目标识别正确率的前提下, 尽可能减少约束条件。

文献[1]通过建立最大颜色光谱直方图表示模型构建表现模型, 计算相似度。文献[2]使用运动目标区域的尺度不变特征变换(Scale Invariant Feature Transform, SIFT)描述子进行多摄像机的目标匹配。文献[3]采用 SIFT 特征实现目标的检索, 取得较好的效果。目前对目标表现模型的建立仅通过颜色或形状特征。而局部特征的描述主要基于形状信息而忽略了颜色信息, 如 SIFT 特征是基于灰度图像提取的, 能够抵抗亮度和形变的影响。但是其对光学变换具有不稳定性。文献[4]采用高斯光谱模型下推导的颜色不变量, 建立彩色尺度不变特征变换(Colored Scale Invariant Feature Transform, CSIFT)特征, 但是所采用的特征对环境光敏感, 且由于使用非线性变换导致了不稳定性^[5]。

针对以上问题, 本文提出一种基于区域 CSIFT 特征的目标识别算法。利用背景差法检测运动目标, 得到运动区域, 在目标区域中, 提取对光照和几何都具有不变性的、稳定的颜色特征, 通过建立 CSIFT 特征进行目标识别, 并建立了多摄像机间的目标对应关系。

2 目标区域检测

考虑到背景中小的运动目标的扰动, 本文采用混合高斯法建立背景模型, 并使用基于像素邻域信息的改进的背景差法检测目标区域^[6]。该方法对噪声不敏感且能够很好的提取运动目标。设当前图像矩阵为 I , 混合高斯法建立的背景图像矩阵为 BG , 图像的行、列分别为 row 和 col , 则基于邻域的背景差法定义如下:

$$\zeta(i, j) = \sum_{l=-1}^1 \sum_{m=-1}^1 \text{abs}[I(x+i, y+j) - BG(x+i, y+j)] \quad (1)$$

其中, ζ 为背景差分后的图像矩阵; $\forall x \in \{2, 3, \dots, row-1\}$, $\forall y \in \{2, 3, \dots, col-1\}$; 采用 3×3 的邻域像素。

如果输入的图像为彩色图像, 则可分 R 、 G 、 B 三通道分别按式(1)进行背景差。得到的差分图像包含了可能为运动区域的像素, 之后, 采用 Ostu 算法获得图像的阈值,

作者简介: 罗 磊(1982—), 男, 工程师, 主研方向: 图像处理; 范彩霞, 讲师、博士研究生

收稿日期: 2011-12-08 **修回日期:** 2012-02-07 **E-mail:** luolei@xaut.edu.cn

并对图像进行二值化,将得到的二值图像进行形态学运算,去除小区域的影响。邻域背景差法与背景差法的结果比较如图 1 所示。

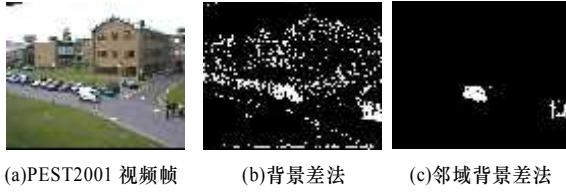


图 1 邻域背景差法与背景差法的结果比较

3 区域 CSIFT 描述子

3.1 颜色不变量特征

考虑室内外环境光的影响,根据 Shafer 双色反射模型,场景的光谱反射模型可表示为^[5,7]:

$$I^k = m^b(x) \int_{\lambda} f_k(\lambda) e(\lambda) b(\lambda, x) d\lambda + m^s(x) \int_{\lambda} f_k(\lambda) e(\lambda) d\lambda + \int_{\lambda} a(\lambda) f_k(\lambda) d\lambda \quad (2)$$

其中, $k = \{R, G, B\}$; $f_k(\lambda)$ 表示成像设备的感光函数; $e(\lambda)$ 为光源的光谱分布; $b(\lambda, x)$ 表示物体的漫反射; m^b 、 m^s 表示漫反射和镜面反射的几何参数; $a(\lambda)$ 表示环境光。

设成像设备感光函数 $f_k(\lambda)$ 为 δ 函数且在中性界面反射(Neutral Interface Reflection, NIR)模型和白光条件下,即 $e(\lambda) = e$, 则上式可简化为:

$$I^k = em^b(x)b^k(x) + em^s(x) + a^k \quad (3)$$

对其求导可去除环境光的影响,同时为了抵抗镜面反射的影响,定义对立色的一阶导数 $O1_x$ 、 $O2_x$ 如下:

$$\begin{aligned} O1_x &= \frac{1}{\sqrt{2}}(R_x - G_x) = \frac{1}{\sqrt{2}}(e(m_x^b(x)(b_x^R(x) - b_x^G(x))) + \\ &\quad m^b(x)(b_x^R(x) - b_x^G(x))) \\ O2_x &= \frac{1}{\sqrt{6}}(R_x + G_x - 2B_x) \end{aligned} \quad (4)$$

得到的对立色 $O1_x$ 、 $O2_x$ 对镜面反射具有不变性,但是其仍依赖于光源和物体自身反射的几何特性,因此,为了得到既对环境光和镜面反射具有不变性,又对光源和物体几何参数具有不变性的颜色不变量,定义色相如下:

$$\begin{aligned} hue_x &= \tan^{-1} \left(\frac{O1_x}{O2_x} \right) = \\ &\quad \tan^{-1} \left(\frac{\sqrt{3}((b_x^R(x) - b_x^G(x)) + (b_x^R(x) - b_x^G(x)))}{((b_x^R(x) + b_x^G(x) - 2b_x^B(x)) + (b_x^R(x) + b_x^G(x) - 2b_x^B(x)))} \right) \end{aligned} \quad (5)$$

可见通过式(5)提取的颜色不变量,对光照环境和物体的几何参数具有不变性,仅与物体的自身反射特性相关。

3.2 CSIFT 描述子

为了得到综合颜色和形状特征,对几何变换和光照变换都具有鲁棒性的描述子,本文在 3.1 节提取颜色不变量的基础上,采用 SIFT 算法提取目标的尺度不变特征。SIFT 算法的步骤如下^[8]:

(1)在尺度空间检测极值

SIFT 特征通过在 DoG(Difference of Gaussian)尺度空

间检测局部极值作为特征点。DoG 算子利用不同尺度的高斯差分核与图像卷积生成:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = \\ &\quad L(x, y, k\sigma) - L(x, y, \sigma) \\ G(x, y, \sigma) &= \frac{1}{2\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \end{aligned} \quad (6)$$

其中, $I(x, y)$ 为输入图像,本文为上节提取的颜色不变量 hue_x ; σ 为尺度空间因子; k 为相邻尺度空间因子的比例系数。

在高斯卷积图像两两相减得差分图像后,将高斯差分图像中的每个极值点与它的邻域、对应的上一层图像的邻域、对应的下一层图像的邻域共 26 个邻域值进行比较,若该点均大于或者均小于这 26 个值,则记录此点的位置和它所在的尺度作为候选关键点。

(2)关键点筛选

在候选的特征点中,去除对比度低且对噪声敏感的点,用特征点周围差分图像的 2×2 Hessian 矩阵排除边缘干扰点。

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (7)$$

如果 $\frac{(Tr(H))^2}{Det(H)} \leq \frac{(\gamma+1)}{\gamma}$ (取 $\gamma=10$), 则保留该特征点,否则丢弃。在式(7)中, $Tr(H) = D_{xx} + D_{yy}$, $Det(H) = D_{xx}D_{yy} - (D_{xy})^2$ 。

(3)确定主方向

每个关键点的方向参数可通过计算关键点的梯度大小和方向来获得。梯度大小可用下式表示:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (8)$$

梯度方向可用下式表示:

$$\theta(,) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \quad (9)$$

在实际计算中,在以关键点为中心的邻域上使用高斯函数进行加权,获得对应的梯度幅度大小和梯度方向。梯度直方图的范围是 $0^\circ \sim 360^\circ$, 其中,每 10° 一个立柱,共 36 个。直方图的峰值作为该关键点的方向。若有其他局部峰值在最大峰值的 80% 内,则把该方向也作为该点的一个方向,即一个特征点可以有多个方向,这样可以增强匹配的鲁棒性。

(4)特征点描述

在颜色不变量表示的图像上,首先将图像所在的坐标轴旋转到主方向,以特征点为中心,选取周围 16×16 大小的区域,然后将此区域分成 4×4 的子区域,每个子区域内计算 8 个方向($0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ, 360^\circ$)的梯度累加和,得到梯度方向直方图,形成 128 维的特征向量。

4 目标识别

4.1 CSIFT 描述子匹配

本文 CSIFT 特征点的匹配采用最近邻方法,即寻找

图像中每个特征点在另一幅图像中的最近邻点。取摄像机 1 采集的图像中的某个特征点, 通过遍历找到其与摄像机 2 采集的当前图像帧中欧氏距离最小的 2 个特征点。如果最近邻距离与次近邻距离的比值小于某个阈值 α (本文中取 0.7), 则判定为一对匹配点。

CSIFT 和 SIFT 2 种算法的基本步骤相同, 本文算法并没有改变 SIFT 特征的结构, 只是使用各子区域内彩色模型来形成特征向量, 建立了对颜色变化和几何变化都具有鲁棒性的描述子, 弥补了 SIFT 算法丢失彩色信息的缺陷。图 2 为 CSIFT 与 SIFT 特征匹配的结果, 其中, CSIFT 成功匹配 79 对, SIFT 成功匹配 120 对, 线条表示特征点的匹配对。



(a)CSIFT 特征匹配结果



(b)SIFT 特征匹配结果

图 2 特征匹配结果比较

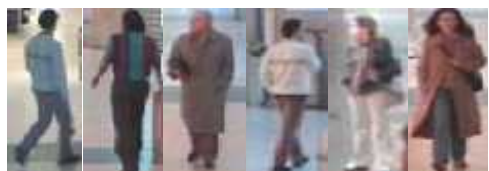
由图 2 可知, CSIFT 算法特征匹配的精度高于 SIFT 算法特征, 但是误匹配的情况依然很多。

4.2 基于区域 CSIFT 的目标识别

从上述分析可知, 直接使用整幅图像的全部 CSIFT 特征进行匹配, 在多摄像机监控中是不适合的。因此, 本文采用基于区域的目标识别。对于摄像机采集的图像序列, 首先使用第 1 节的运动目标区域检测方法提取运动目标区域, 基于区域 CSIFT 的目标识别如图 3 所示。



(a)CAVIAR 视频帧



(b)运动目标区域



(c)目标匹配结果

图 3 基于区域 CSIFT 的目标识别

图 3(c)从右到左 2 幅图为一组, 匹配的特征点个数分别为 9、1、0、0、0。最后, 选择最大匹配度的目标为匹配结果。

5 实验结果与分析

本文采用 PEST2001 和 CAVIAR 测试视频分别测试算法对刚体目标和非刚体目标的识别效果。两者均满足多摄像机对同一场景拍摄的要求。本文比较了 SIFT 描述子^[2]和 CSIFT 描述子的识别效果。刚体目标 CSIFT 和 SIFT 特征匹配结果如图 4 所示。



(a)CSIFT 特征目标匹配结果



(b)SIFT 特征目标匹配结果

图 4 刚体目标 CSIFT 和 SIFT 特征匹配结果

图 4 中所有的目标区域归一化为 180×140 像素, 在图 4(a)中, 从右至左, 2 幅图为一组, 第 1 行匹配的特征点个数分别为 9、5、4, 第 2 行匹配的特征点个数分别为 11、1、4, 在图 4(b)中, 匹配的特征点个数分别为 13、8、6、20、27、23, 具有最大匹配度的目标为匹配结果。匹配结果每一行的第 1 列为同一目标的匹配结果, 可见, 本文提出的算法的匹配效果优于 SIFT 特征^[2](限于篇幅, 选取部分数据罗列)。

非刚体目标 CSIFT 和 SIFT 特征匹配结果如图 5 所示。



(a)CSIFT 特征目标匹配结果



(b)SIFT 特征目标匹配结果

图 5 非刚体目标 CSIFT 和 SIFT 特征匹配结果

由图 5 可知, 目标的姿态和颜色发生了变化。图 5(a)匹配的特征点个数分别为 10、2、1、2、2, 图 5(b)匹配的特征点个数分别为 11、5、4、8、5。基于 CSIFT 和 SIFT 特征匹配均能得到实际应该匹配的结果。本文算法正确匹配的特征点个数与其他的匹配点个数相比相差较大, 而 SIFT 特征的正确匹配个数为 11, 而其他的匹配结果最大

为 8, 可见, 本文算法有更强的鲁棒性。

图 6 为表 1(2 种算法识别率比较)对应的识别目标, 其中, 从左至右依次对应表 1 的 A~F。



图 6 识别目标

表 1 2 种算法识别率比较 (%)						
算法	A	B	C	D	E	F
本文算法	92.3	91.1	93.1	90.2	91.7	92.0
文献[2]算法	80.5	81.8	79.7	83.3	78.4	83.5

实验结果表明, 本文算法能够达到较高的识别率, 但是运动目标的分割结果以及目标间的遮挡和粘连是影响目标识别效果的主要原因。

6 结束语

针对无重叠视域多摄像机联合监控中的目标识别问题, 本文提出一种基于区域 CSIFT 特征的目标识别算法。在检测得到的目标区域中, 提取与光照和几何特性无关的颜色不变量, 通过尺度不变描述子抵抗形变的影响, 得到 CSIFT 特征进行目标识别。实验结果表明, 本文算法的匹配效果较优。今后将研究目标的遮挡、粘连对目标识别效果的影响及其处理方法。

参考文献

[1] Madden C, Eric D C, Piccardi M. Tracking People Across Disjoint

Camera Views by an Illumination Tolerant Appearance Representation[J]. Machine Vision and Applications, 2007, 18(3): 233-247.

[2] 明安龙, 马华东. 多摄像机之间基于区域 SIFT 描述子的目标匹配[J]. 计算机学报, 2008, 31(4): 650-661.

[3] 林传力, 赵宇明. 基于 Sift 特征的商标检索算法[J]. 计算机工程, 2008, 34(23): 275-277.

[4] Abdel H A E, Farag A A. CSIFT: A SIFT Descriptor with Color Invariant Characteristics[C]//Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE Press, 2006.

[5] Joostvan D W, Beigpour S. The Dichromatic Reflection Model: Future Research Directions and Application[EB/OL]. (2011-11-21). <http://www.cat.uab.cat/Public/Publications/2011/VaB2011/>.

[6] Saeid F, Hamed M P, Hamed B. Multiple Object Tracking Using Improved GMM-based Motion Segmentation[C]//Proc. of the 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. [S. l.]: IEEE Press, 2009.

[7] Sande K E A, Gevers T, Snoek C G M. Evaluation of Color Descriptors for Object and Scene Recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(9): 1582-1596.

[8] Lowe D G. Distinctive Image Features from Scale-invariant Keypoints[J]. Computer Vision, 2004, 60(2): 91-110.

编辑 刘 冰

(上接第 190 页)

参考文献

[1] Donald A N. Emotional Design: Why We Love(or Hate) Everyday Things[M]. New York, USA: Basic Books, 2004.

[2] Donald A N. The Design of Everyday Things[M]. New York, USA: MIT Press, 1998.

[3] Richard E M, Roxana M. Nine Ways to Reduce Cognitive Load in Multimedia Learning[J]. Educational Psychologist, 2003, 38(1): 43-52.

[4] Richard E M. The Cambridge Handbook of Multimedia Learning[M]. [S. l.]: Cambridge University Press, 2005.

[5] Ben S, Catherine P. Designing the User Interface: Strategies for Effective Human-computer Interaction[M]. 5th ed. [S. l.]: Addison-Wesley, 2011.

[6] Card S, Thomas P M, Allen N. The Psychology of Human Computer Interaction[M]. New York, USA: Routledge Press, 1986.

[7] Allison D, Elizabeth F, Hilary H. Children's Roles Using Keyword Search Interfaces at Home[C]//Proc. of the 28th International Conference on Human Factors in Computing Systems. New York, USA: [s. n.], 2010.

[8] Michael S H, Crouser R J, Marina U B. Tangible Interaction and Learning: The Case for a Hybrid Approach[EB/OL]. (2010-11-21). http://scienceindex.com/stories/1697570/Tangible_interaction_and_learning_the_case_for_a_hybrid_approach.html.

[9] Alan D, Janet F, Gregory D A. Human-computer Interaction[M]. 3rd ed. [S. l.]: Pearson/Prentice-Hall, 2004.

[10] Norman D A. The Psychology of Everyday Things[M]. New York, USA: Basic Books, 1988.

[11] Xu Rui, Donald C W. Survey of Clustering Algorithms[J]. IEEE Transactions on Neural Networks, 2005, 16(3): 645-678.

[12] 刘夫涛, 张 雷. 多重系统聚类挖掘算法及其实现[J]. 计算机工程与应用, 2000, 36(10): 41-42.

[13] 袁 方, 孟增辉, 于 戈. 对 K-means 聚类算法的改进[J]. 计算机工程与应用, 2004, 40(36): 177-179.

[14] Davies D L, Bouldin D W. A Cluster Separation Measure[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1979, 1(2): 224-227.

编辑 刘 冰