

用于机器人目标跟踪的压缩感知的改进算法

杨文勇

(厦门城市职业学院 电子与信息工程系, 福建 厦门 361008)

摘要: 针对机器人进行目标跟踪时的跟踪算法,对压缩感知在视频目标跟踪中的应用进行了改进,通过对图像进行伽马变换,以平滑图像的明亮程度,使其能够更好地适应复杂环境。研究结果表明,改进的算法有效地增加算法的鲁棒性和提高算法的效率。

关键词: 机器人; 目标跟踪; 压缩感知; 伽马变换; 算法

中图分类号: TP242 **文献标志码:** A **文章编号:** 1672-4348(2014)06-0573-04

Improved algorithm for compressive sensing in robot target tracking

Yang Wenyong

(Electronics and Information Engineering Department, Xiamen City University, Xiamen 361008, China)

Abstract: The application of compressive sensing in video target tracking was improved to update the tracking algorithm of robot target tracking. The images were transformed via gamma to smooth the brightness of the images to adapt to the complex situations. The experimental results show that the improved algorithm can effectively increase the algorithm robustness and improve the efficiency of the algorithm.

Keywords: robot; object tracking; compressive sensing; gamma transformation; algorithm

机器人对于运动目标的定位是机器人视频系统设计的重要组成部分,运动目标检测的目的是从背景中分割出运动物体的区域,这是许多系统设计的初始工作,其效果如何将直接影响后续的基于目标检测算法的目标跟踪、行为识别的效果。传统的基于静态背景的检测技术有4种:背景差分法、帧间差分法、光流法、统计模型分类法^[1-2]。文献[3]提出了一种简单高效的基于压缩感知的跟踪算法。首先利用符合压缩感知(restricted isometry property, RIP)条件的随机感知矩阵对多尺度图像特征进行降维,然后再利用贝叶斯分类器进行分类,具有较强的鲁棒性。但其在复杂环境中,跟踪效果并不理想,本文所述是在文献[3]的基础之上进行改进。

1 压缩感知

1.1 压缩感知理论

压缩感知理论(简称CS),开始出现于Cades, Doho, Tao 等人在2006年所发表的3篇文章^[4-6]。设 $\theta \in \mathbf{R}^N$,其中 θ 表示 N 个元素中只有 K 个非零元素,即 θ 为 K -稀疏向量。离散信号 $x \in \mathbf{R}^N$,测量信号 $y \in \mathbf{R}^n, n < N$,有如下公式:

$$\begin{cases} x = \Psi\theta \\ y = \Phi x \end{cases} \quad (1)$$

其中 Φ 为测量矩阵, Ψ 为稀疏基。如果 $A = \Phi\Psi$ 满足约束等距性(restricted isometry property, 简称RIP),则可通过如下的线性规划问题来重构 x :

$$\min \|\theta\|_1 \text{ 且 } y = \Phi\Psi\theta \quad (2)$$

收稿日期: 2014-10-30
作者简介: 杨文勇(1971-),男(汉),黑龙江齐齐哈尔人,讲师,硕士,研究方向:计算机软件与自动控制。

相当于寻找 $y = \Phi\Psi\theta$ 的最稀疏解。测量矩阵 Φ 与稀疏矩阵 Ψ 满足不相关性的要求是 RIP 准则一种等价的情况。矩阵的相关性定义如下:

$$\mu(\Phi, \Psi) = \sqrt{n} \max_{1 \leq i, j \leq n} \|\Phi_i^T \cdot \Psi_j\| \quad (3)$$

其中 Φ_i^T 与 Ψ_j 分别是 Φ 的第 i 行与 Ψ 的第 j 列。不相关性要求 μ 值越小越好, μ 值越小, 说明两个矩阵的就越不相关。实际应用中, 从概率上讲, 一维信号测量矩阵 Φ 选取服从高斯分布的基矢量能保证与任意稀疏基 Ψ 不相关的比较高。

测量矩阵的设计和信号的重构问题是 CS 理论中的难点。因为最优重构信号计算复杂度高, 求近似解就成为一个研究方向。一个求近似解的基本方法是贪婪算法, 利用选择适当的原子后经逐步递增方法来达到逼近信号矢量的目的(匹配跟踪算法)。另一个是基追踪算法, 原理是利用 l_p ($0 < p \leq 1$) 范数通过线性规划来求解, 这种方法比前一方法求解更精确, 缺点是计算复杂度高^[7]。

1.2 随机投影

一个随机矩阵 $R \in R^{n \times m}$, 将高维的图像空间 $X \in R^m$ 投影到一个低维空间 $V \in R^n$

$$V = R \times X \quad (4)$$

且 $n \ll m$ 。理想的情况下, 我们希望随机矩阵 R 在投影之后能够近似保留原信号的距离特征。Johnson-Lindenstrauss lemma^[8]阐述了如果高维图像空间被随机投影到合适的随机的高维子空间中, 那么在低维空间中高维两点之间的距离能够很大的概率被保留。Baraniuk 等^[9]证明了随机矩阵满足 Johnson-Lindenstrauss lemma 的要求并且在压缩感知中能够保持对 RIP 的正确性。因此如果得到一个随机矩阵 R 在式(2-4)满足 Johnson-Lindenstrauss lemma 的要求, 我们可以很大概率重构 x 。我们就可以确信 V 能够保留所有的信息关于 X 中。以上的理论支持让我们可以分析低维的随机投影来得到高维的信号。在下面介绍的算法中, 本文使用的稀疏矩阵不仅满足 Johnson-Lindenstrauss lemma, 而且能够有效地计算对于实时跟踪。

1.3 随机测量矩阵

典型的满足 RIP 的测量矩阵是随机高斯矩阵 $R \in R^{n \times m}$, $r_{ij} \sim N(0, 1)$ 。然而, 由于矩阵不是稀疏矩阵, 所以当 m 非常大的时候, 则矩阵存储和计

算的花费比较大。在本文中采用如下的稀疏矩阵:

$$r_{ixj} = \sqrt{s} \begin{cases} 1 \\ 0 \\ -1 \end{cases} \quad (5)$$

其中 1 的概率为 $\frac{1}{2s}$, 0 的概率为 $1 - \frac{1}{s}$, -1 的概率为 $\frac{1}{2s}$ 。Achlioptas^[8]证明这个类型的矩阵当 $s = 2$ 或 3 时, 满足 Johnson-Lindenstrauss lemma。这个矩阵可以很容易地得到, 只需一套随机产生器。更重要的, 当 $s = 3$ 时, 矩阵只剩下不多的数, 三分之二的计算将可以不用进行。另外, Li 等^[10]展示当 $s = O(m)$ ($x = R^m$), 这个矩阵是近似正常的。甚至当 $s = m/\log(m)$, 随机投影矩阵和传统的高斯随机投影一样正确。在这个算法中, 我们设置 $s = m/4$ 来构造这个稀疏矩阵, 这样可以只储存非零的数以减少内存的占用和计算时间的开支。对于矩阵 R 的每一行, 只需要计算 $c \leq 4$ 的数。因此计算复杂度非常低, 只要 $O(cn)$ 。甚至只需要存储矩阵 R 每行非零的数, 这样内存的消耗则非常低。

1.4 目标跟踪算法的描述

在进行描述之前先要进行低维空间特征值的提取和分类器的构建和更新。

本文所采用的特征提取是采用 Haar-like^[11]特征值的提取, Haar-like 特征已经广泛使用目标的检测并证明可行。但是其产生一个非常大的基本特征使得计算量非常大。Boosting 算法选择重要的特征有效地解决了这个问题。在本文中, 大量的 Haar-like 特征通过稀疏矩阵被有效地压缩, 其压缩感知理论保证了被压缩的特征能够保留原始图像的信息不被丢失。这样就避免了 Haar-like 的维数灾难。

分类器的构造和更新。对于每个样本 $z \in R^m$, 它的低维空间表示为 $v = (v_1, \dots, v_n)^T \in R^n$, 且 $m \gg n$, 我们假设所有元素在 v 是独立分布的并且对它们进行一个朴素的贝叶斯建模。

$$H(v) = \log \left(\frac{\prod_{i=1}^n p(v_i | y = 1) p(y = 1)}{\prod_{i=1}^n p(v_i | y = 0) p(y = 0)} \right) = \sum_{i=1}^n \log \left(\frac{p(v_i | y = 1)}{p(v_i | y = 0)} \right) \quad (6)$$

这里建设概率均匀分布, $p(y = 1) = p(y = 0)$ 。

其中 $p(y = 1)$ 表示为目标框附近采样样本的概率, $p(y = 0)$ 表示目标框较远采样样本的概率。

Diaconis 和 Freedman^[12] 表明高维度的随机向量的随机投影总是满足高斯分布的。因此在分类器 $H(v)$ 中, $p(y = 1)$ 和 $p(y = 0)$ 的状态分布是符合高斯分布的 $(\mu_i^1, \sigma_i^1, \mu_i^0, \sigma_i^0)$

$$\begin{aligned} p(v_i | y = 1) &\sim N(\mu_i^1, \sigma_i^1), \\ p(v_i | y = 0) &\sim N(\mu_i^0, \sigma_i^0) \end{aligned} \quad (7)$$

高斯参数的更新如下公式:

$$\begin{cases} \mu_i^1 \leftarrow \lambda u_i^1 + (1 - \lambda) u^1 \\ \sigma_i^1 \leftarrow \sqrt{\lambda (\sigma_i^1)^2 + (1 + \lambda) (\sigma^1)^2 + \lambda (1 - \lambda) (\mu_i^1 - \mu^1)^2} \\ \mu^1 = \frac{1}{n} \sum_{k=0}^{n-1} v_i(k) \\ \sigma^2 = \sqrt{\frac{1}{n} \sum_{k=0}^{n-1} (v_i(k) - \mu^1)^2} \end{cases} \quad (8)$$

其中, 当 $\lambda > 0$ 表示学习速率的参数。

压缩感知的目标跟踪的描述(假设开始目标的位置已知): 输入第 t 帧视频帧,

1) 采样一系列的矩形图形块, $D^y = \{z | \|l(z) - l_{t-1}\| < \gamma\}$, 其中 l_{t-1} 是 $(t-1)$ 帧中目标的位置, 之后提取低维空间的特征值;

2) 使用朴素贝叶斯分类器 H 对提取的特征值进行分类, 找到分类最大值的跟踪位置 l_t ;

3) 采样两个一系列的矩形图像块(一个是前景图像目标区域周围图像块和一个距离目标区域较远的背景图像周围图像块) $D^\alpha = \{z | \|l(z) - l_t\| < \alpha\}$ 和 $D^{\beta, v} = \{z | \beta < \|l(z) - l_t\| < v\}$ 且 $\alpha < \beta < v$;

4) 提取采集的两个样品特征并更新分类参数依据公式(8)。

输出为跟踪的位置 l_t 和分类参数。

跟踪的效果见图 1, 分别为视频第 4、34 帧的截图。

从以上的结果中可以看出, 本算法可以很好地对目标进行跟踪, 具有较强的鲁棒性。

2 目标跟踪算法的问题和改进

压缩感知算法适合背景较简单, 光线明暗并没有太大变化, 且目标物无遮掩的情况下跟踪效果较好。对光线有较大的变化或有遮挡的情况下, 效果则很差甚至无法跟踪。如图 2 和图 3。

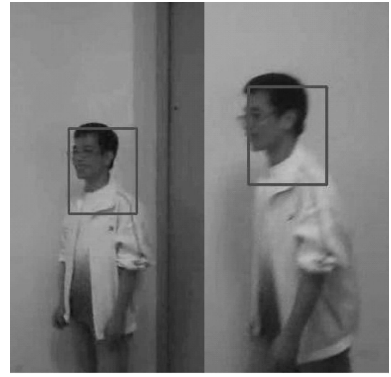


图 1 视频第 4、34 帧的截图

Fig. 1 Truncated image of video phase 4, phase 34



图 2 视频第 1 帧的截图

Fig. 2 Truncated image of video phase 1



图 3 视频第 62 帧的截图

Fig. 3 Truncated image of video phase 62

由上述面对比中可以看出当目标物的背景由亮度较大变为亮度较暗时, 发生了较大的错误。基于以上问题, 本文提出了图像基于伽马变换进行灰度值修改, 使图像的明暗光线变化趋于平缓。

伽马变换的基本形式为

$$s = cr^\gamma \quad (9)$$

其中 c 和 γ 为正常量。式(9)中, 随着 γ 取值的不同, 伽马变换的效果也不一样。当 $\gamma = 1$ 时, 则为恒等变换, 当 $\gamma < 1$ 时, 则把较窄范围的暗色输入值映射为较宽的输出值, 则较暗的图像转换成较明亮的图像, 当 $\gamma > 1$ 时, 则相反。但是并不是每一段视频都需要进行伽马变换, 且伽马变换所需

要的计算时间也较长,本文随机抽取图像的若干个图像块,并以图像块的平均值作为图像块的值,将灰度空间 0 ~ 255 平均划分 8 个空间区域,即: [0, 31] [32, 63] [64, 95] [96, 127] [128, 159] [160, 191] [192, 233] [234, 255]。统计落入各个区间的数目,并进行各个区间的概率计算。如果区间 1 和 2 的概率超过某个暗阈值 σ ,则认为图像偏暗,需进行伽马变换,如果区间 7 和 8 的概率超过了某个亮阈值 ω ,则认为图像偏亮,需要进行伽马变换。如图 4 灰度图所示。



图 4 伽马变换前的灰度图

Fig. 4 Grey degree figure pre gamma transformation

图 4 的灰度值统计如下:

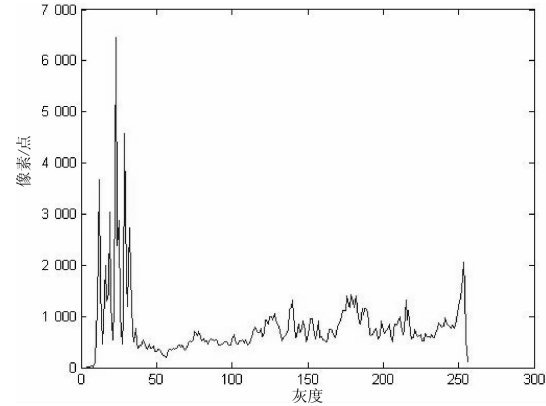


图 5 统计直方图(变换前)

Fig. 5 Statistical histogram (pre gamma transformation)

由上统计图可以获得在区间 1,2 所获得的概率较大,所以需要进行把较窄范围的暗色输入值映射为较宽的输出值,得到的灰度图和直方图如图 6 和统计图 7 所示。

由以上的对比结果可知图像的明暗对比度被缩小了,可以减少压缩感知的误差。实验结果也能够较好地适应图片对比度较大的环境。修改之后所得到的结果,见图 8。

除此之外,也可以通过路径拟合判断来增加跟踪的准确性。



图 6 伽马变换后的灰度图

Fig. 6 Grey degree figure post gamma transformation

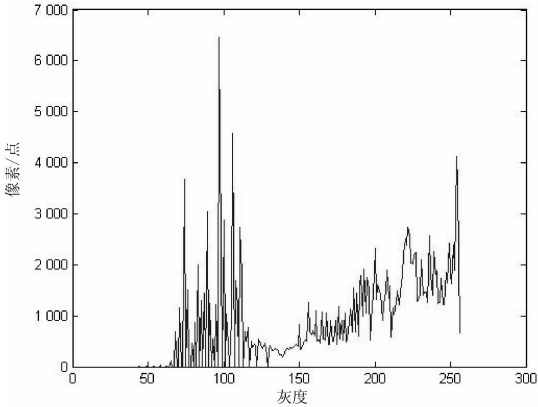


图 7 统计直方图(变换后)

Fig. 7 Statistical histogram (post gamma transformation)



图 8 视频第 63 帧截图

Fig. 8 Truncated image of video phase 63

3 结论与展望

基于压缩感知的视频目标跟踪算法能够有效地进行实时的目标跟踪,且效率有很大的提高,但在明暗变化较大的环境中,跟踪的效果会比较差。本文提出了一种先随机的抽取视频画面中的图像块,并将灰度值分为 8 个区间,将抽样中落入区间的概率算出,继而得出是否需要视频图像进行预处理。如果需要对视频图像作预处理,则进行伽马变换,以平滑图像的明亮程度,使得压缩感知的目标跟踪能更好地适应更复杂的环境。但其对

(下转第 594 页)

表 2 4 × 4 交换机下带宽保证率

Tab.2 Bandwidth guaranteeing rate with 4 × 4 exchanger

约定带宽/b	IEDF/b	EDF/b
1	75	55
2	60	43
3	57	33

如表 2 所示,在相同的约定带宽下,IEDF 算法的带宽保证率较原 EDF 算法平均提高 19.7%,有效保证了带宽。

参考文献:

[1] 林闯,单志广,任丰原. 计算机网络的服务质量(QoS)[M]. 北京:清华大学出版社,2004;210 – 213.

[2] 王强,王宏安,金宏,等. 实时系统中的非定期任务调度算法综述[J]. 计算机研究与发展,2004,41(3):385 – 392.

[3] 郭艾寅,钱光明,徐云剑. 基于多种服务并存网络的实时队列时延研究[J]. 计算机工程与应用,2009,44(17):101 – 103.

(责任编辑:肖锡湘)

(上接第 576 页)

目标有遮挡的过程,其跟踪效果不佳,因此还可以考虑通过运动光流法对目标跟踪算法进行改进。

参考文献:

[1] Magee D. Tracking multiple vehicle using foreground, background and motion models[J]. Image and Vision Computation, 2004,22(2):143 – 155.

[2] Jia L H, J H Z. Improved background subtraction based on novel thresholding[J]. Acta Automatica Sinica,2009,35(4):467 – 473.

[3] Zhang K, Zhang L, Yang M H. Real-time compressive tracking[C]// Proceedings of the 12th European Computer Vision(Ⅲ)ECCV 2012. Berlin:Springer,2012:864 – 877.

[4] Candes E J, Romberg J, Tao T. Robust uncertainty principles: Exact signal recognition from highly incomplete frequency information[J]. IEEE Trans on Info Theory, 2006,52(2):489 – 509.

[5] Donoho D L. Compressed sensing[J]. IEEE Trans on Info Theory, 2006,52(4):1289 – 1306.

[6] Candes E J, Tao T. Near optimal signal recovery from random projections: Universal encoding strategies[J]. IEEE Trans on Info Theory, 2006,52(12):5406 – 5425.

[7] 朱明,高文,郭立强. 压缩感知理论在图像处理领域的应用[J]. 中国光学,2011,4(5):441 – 447.

[8] Achlioptas D. Database-friendly random projections: Johnson-Lindenstrauss with binary coins[J]. Journal of Computer and System Sciences,2003,66(4):671 – 687.

[9] Baraniuk R, Davenport M, Devore R, et al. A simple proof of the restricted isometry property for random matrices[J]. Constructive Approximation,2008,28(3):253 – 263.

[10] Li P, Hastie T J, Church K W. Very sparse random projections[C]// Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2006;287 – 296.

[11] Babenko B, Yang M H, Belongie S. Robust object tracking with online multiple instance learning[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on Pattern Analysis and Machine Intelligence,2011,33(8):1619 – 1632.

[12] Diaconis P, Freedman D. Asymptotics of graphical projection pursuit[J]. The Annals of Statistics,1984,12(3):793 – 815.

(责任编辑:肖锡湘)