

基于信息熵的 Mean-shift 目标跟踪^①

周尚波^{1,2} 蒋镇波¹ 谢显中² (1. 重庆大学 计算机学院 重庆 400044;
2. 重庆市计算机网络与通信技术重点实验室 重庆 400065)

摘要: 视频目标跟踪在交通、军事等领域具有重要的应用价值。基于信息熵理论,提出了一种视频特征相关匹配的视频目标跟踪算法。首先引入信息熵概念,以信息熵描述视频目标特性,结合 Mean-shift 算法,针对不同的两个颜色空间 RGB 与 HSV,用特征相关匹配法设计跟踪算法。试验结果表明,所提跟踪算法跟踪具有较好的实时性,取得较好的跟踪效果。

关键词: 信息熵; 特征匹配; mean-shift; 目标跟踪; 特征直方图

Object Tracking Using Mean-Shift Based on Comentropy

ZHOU Shang-Bo^{1,2}, JIANG Zhen-Bo¹, XIE Xian-Zhong²

(1. Computer College of Chongqing University, Chongqing 400044, China; 2. Chongqing Key Lab of Computer Network and Communication Technology, Chongqing 400065, China)

Abstract: Video object tracking technology has great applied value in the fields of traffic and military. In this paper, a video matching object tracking algorithm, which is comentropy-based, is proposed. First of all, the concept of comentropy is introduced. Secondly, the feature of the video object is described by comentropy, and the algorithm which combined with mean-shift is designed with feature match in two different feature spaces, RGB and HSV. Finally, the experiment simulation is presented. The experiment results show that the proposed algorithm can track the object steady and real-time. This algorithm can achieve the desired effects and it is a practical algorithm.

Keywords: comentropy; feature match; mean-shift; object tracking; feature histograms

近年来,视频研究渐渐成为计算机研究领域的重点和热点,而其中视频跟踪技术更是得到大多研究者的关注。随着社会的进步与科技的发展,视频跟踪技术的运用越来越广泛,诸如医学探测,生物分析,军事制导等方面。而如何能够让视频跟踪更加准确与实时是当前最关注的问题。

就目前跟踪算法而言,大部分采用了基于颜色特征直方图的跟踪方法,该方法具有采集特征方便直接以及相似度计算简便的特点,同时它受目标物体大小

变化扭曲的影响也不是很大。但是运用颜色特征直方图代表目标的特征会存在特征维数过高的问题,而造成存储空间过大而降低了算法的实时性。John Zachary 提出了运用信息熵的算法来进行图像匹配算法^[1],这样可以将特征直方图的维数由多维降到一维,这样大大提高了算法的效率。然而信息熵其实是代表了某个概率的分布,可能存在两个物体的特征概率相同,但分布的顺序不同,而他们的熵值却是相同的。根据 John Zachary 的算法思想,这两个物体是可以匹配的,但事实却与之相反。在之后的研究中,还有

① 基金项目:重庆市计算机网络与通信技术重点实验室项目(CY-CNCL-2008-02)

收稿日期:2010-03-23;收到修改稿时间:2010-05-04

学者将熵的概念运用到了高斯模糊类聚的方法中^[2],通过将初始得到的信息熵构建矩阵,再通过高斯类聚变换以及卡尔曼滤波得到目标的位置。此外,由于小波直方图统计跟踪的算法稳定,跟踪准确,但复杂度太大,部分学者也倾向于运用小波熵来进行图像检索^[3]以及其他图像优化算法^[4-6]。而在本文的算法中,我们借助于 John Zachary 提出的信息熵图像匹配的思想进行跟踪,将局部信息熵加入到 mean-shift 算法的目标函数中去,通过迭代的方法求得最匹配目标的区域;同时,我们又将特征直方图的反向投影值运用于相关性的计算,这样,新算法即能保持信息熵图像匹配思想的高效率性又能体现特征直方图对于目标特征的准确性。

1 图像信息熵

自香农将物理中熵的概念引入信息论以来,熵已被广泛地应用于信号处理、图像处理等领域。为了将香农在信息论中定义熵的概念应用于图像匹配,需要定义图像的熵。图像信息熵定义:假定图像函数具有非负值,即 $f(x,y) \geq 0$, 对于一幅大小 $M \times N$ 的图像,定义

$$H_f = - \sum_{i=1}^M \sum_{j=1}^N p_{ij} \log p_{ij}$$

其中, $p_{ij} = f(x,y) / \sum_{i=1}^M \sum_{j=1}^N f(x,y)$ 。称 H_f 为该图像的

熵。如果只对图像的某一局部定义熵,则称其为局部熵^[7]。从局部信息熵的定义来看,显然它是由一块图像的所有信息统计而来,也就是说仅变化单个像素点的信息对于局部熵的变化来说意义不大,这样,它可以很好的避免了干扰。

交叉熵^[7-8]是一种用来度量两个概率分布 $P = \{p_1, p_2, \dots, p_N\}$, $Q = \{q_1, q_2, \dots, q_N\}$ 之间信息量差异的量,其定义为:

$$D(P, Q) = \sum_{i=1}^N p_i \log \frac{p_i}{q_i}$$

它的对称形式称为对称交叉熵,其定义为:

$$D(P : Q) = \sum_{i=1}^N (p_i \log \frac{p_i}{q_i} + q_i \log \frac{q_i}{p_i})$$

如果只对图像的某一局部计算交叉熵,则成为图

像局部交叉熵。在跟踪的过程中,我们可以把 p_i 看成是目标特征模板的概率分布,而 q_i 则可以看成是跟踪过程中每一帧图像的实时特征模板的概率分布。由于交叉熵是一类下凸函数,当 P 和 Q 完全一致时可取得最小的零值,说明模板图像与实时图像之间的信息量差异最小。所以,我们可以用熵或局部熵来构造 mean-shift 算法的目标函数,通过图像匹配的方法进行跟踪。

2 视频目标跟踪模型

2.1 Mean-shift 算法

目标函数的构造

设图像某一颜色特征 C 的取值范围为 $0 \sim K$ 。我们对特征 C , 基于 Mean-shift 算法思想,构造目标匹配相关函数,并以此相关函数为目标跟踪的目标函数。其步骤如下:

1) 选取初始帧

初始帧为目标已知的一帧图像。以目标的某一窗口为对象,将特征 C 取值范围分成 u 个特征(例如 $u=1, 2, \dots, m, m=16$ 。 $u=1$ 表示特征 C 值为 0 至 15 , $u=2$ 表示特征 C 值为 16 至 31), 选定目标初始窗口(设窗口含 n 个像素点),对每个特征 u , 计算其直方图值,即根据各个像素点的 C 特征,判断其属于哪个特征区间,哪个特征区间的值就相应增加。所有像素点都处理后,就得到了特征直方图,并将其标准化,使其中最大的值对应于 255 。而这个最大值的选取也有待商榷,它可能也会影响跟踪的效果,在之后的实验仿真过程中将会提到。接着,我们参考 Bradski 提出的反向投影思想^[9],将所得的特征直方图 H_u 反向投影回原图,这样整个图像就变成了一幅目标特征的概率分布图,上面的每一个像素点就可以理解为与目标的相关性系数,可以当作某个像素点的权值。现在,我们可以计算出这个目标的概率分布值 q_u 。

$$q_u = C \sum_{i=1}^n k \left(\left\| \frac{x_0 - x_i}{h} \right\|^2 \right) H_u d[b(x_i) - u] \tag{1}$$

其中, x_0 为目标中心点,目标含有像素点 $\{x_i\}_{i=1, 2, \dots, n}$, $k(\cdot)$ 为核函数(可取 $k(t) = \frac{1}{\sqrt{2\pi}} e^{-t}$), h 为窗口带宽(可取为窗口半径), $d[b(x_i) - u]$ 为分类函数,只有当 u

属于该区间时函数值为1，否则为0。

$$C = 1 / \sum_{i=1}^n k \left(\left\| \frac{x_0 - x_i}{h} \right\|^2 \right) H_u \quad (2)$$

2) 第1帧(l>1)中的目标匹配

根据第l-1帧的目标位置 y_0 ，以为搜索窗口中心坐标计算当前帧的候选图像直方图 H'_u 。第u个特征值的概率分布值为

$$p_u = C \sum_{i=1}^n k \left(\left\| \frac{y_0 - x_i}{h} \right\|^2 \right) H'_u d[b(x_i) - u] \quad (3)$$

3) 目标与候选区域的相似度计算

使用信息熵来构造目标函数，定义为

$$r(y) = \sum_{u=1}^m [p_u(y) \ln q_u + q_u \ln p_u(y)] \quad (4)$$

4) 目标搜索

根据之前所想的图像匹配，想要求出当前帧目标的位置，即是求式(4)的最小值处，而在这个情况下求得的位置即为当前目标的位置。现在用泰勒公式在点 y_0 处展开(4)式，通过一些等量变换后得到的线性表达式为

$$r[p(y), q] \approx \sum_{u=1}^m [q_u (\ln p_u(y_0) - 1)] + \sum_{u=1}^m \left[\ln q_u + \frac{q_u}{p(y_0)} \right] p(y) \quad (5)$$

假设在连续帧的跟踪下，如果候选特征 $p(y)$ 与之前的 $p(y_0)$ 特征相似，我们就认为图像特征匹配，跟踪也就可以继续了。将(1)带入(5)后得到：

$$r[p(y), q] \approx \sum_{u=1}^m [q_u (\ln p_u(y_0) - 1)] + C \sum_{i=1}^n w_i k \left(\left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (6)$$

其中，

$$w_i = \sum_{u=1}^m \left[\left(\ln q_u + \frac{q_u}{p(y_0)} \right) H_u d[b(x_i) - u] \right] \quad (7)$$

为了使得局部匹配，就希望信息熵最小，即(6)式最小。观察表达式(6)，第一项和y没有关系，不影响极小值；而第二项中，我们可以看成是当前帧中y点以 w_i 为权重，通过核函数 $k(t)$ 计算的概率密度估计。而这种方式密度估计可以用在mean-shift算法中计算最大匹配的熵^[10]。在这个跟踪过程中，目标的位

置可以根据式(8)从 y_0 处移动到 y_1 处。

$$y_1 = \frac{\sum_{i=1}^n x_i w_i g \left(\left\| \frac{y_0 - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n w_i g \left(\left\| \frac{y_0 - x_i}{h} \right\|^2 \right)} \quad (8)$$

其中， $g(x) = -k'(x)$ ， $k(x)$ 是对于所有非负数都成立的递减函数。

通过(8)式的计算后，得到一个当前的匹配点，通过与原来点 y_0 的距离计算，如果 $\|y_1 - y_0\| < \epsilon$ ，即达到收敛，则当前的 y_1 即是下一帧开始时的 y_0 ；而 $\|y_1 - y_0\| > \epsilon$ 时，则表示尚未收敛，继续进行mean-shift的迭代运算，并用替换成 y_0 ，直到收敛为止。

对于在实际的仿真运算过程中，若取

$$k(x) = \frac{1}{\sqrt{2p}} \exp(-x), \text{ 则式(3)为}$$

$$p_u = C \sum_{i=1}^n \exp \left(- \left\| \frac{y_0 - x_i}{h} \right\|^2 \right) d[b(x_i) - u] / \sqrt{2p}$$

对于核函数 $e^{-\frac{|x_0-x|}{h}}$ 和对数函数，为了简化计算与节省计算时间，用泰勒级数近似，即

$$e^{-\frac{|x_0-x|}{h}} \approx 1 - \frac{|x_0-x|}{h} \left[1 - \frac{1}{2} \frac{|x_0-x|}{h} \left(1 - \frac{1}{3} \frac{|x_0-x|}{h} \left(1 - \frac{1}{4} \frac{|x_0-x|}{h} \right) \right) \right] \text{ 显然，}$$

$e^{-\frac{|x_0-x|}{h}} \geq 0$ ，为了使 $e^{-\frac{|x_0-x|}{h}}$ 单调，我们取 $h \geq \sup_x \|x_0 - x\|$ 。

2.2 基于HSV颜色空间与灰度的跟踪

1) 通过把RGB特征空间转为HSV特征空间，我们可以运用特征H来绘制直方图，并通过3.1介绍的mean-shift算法进行跟踪。

2) 将RBG颜色空间转换为灰度特征空间，其变换公式如下：

$$GRAY = 0.3 * R + 0.6 * G + 0.1 * B$$

同样，转换后将GRAY视为跟踪的特征，并绘制直方图，进行跟踪。

3 实验结果及分析

在实验仿真过程中，我们运用了第三节所说的两

种特征空间。通过实验，与传统的 Mean-shift 算法相比，我们发现，由于增加了反向投影后的相关性参数，使得像素点之间的关系判断更加明确，再加上运用了局部信息熵的图像匹配思想，使得本文提出的新算法更具实时性与鲁棒性。

实验过程中，我们对 HSV 特征与灰度特征分别进行了跟踪仿真，实验结果表明两种特征空间都可以运用新算法进行很好的跟踪，跟踪效果如图 1 所示。



图 1 对于快速移动车辆的跟踪

在图 1 的实验中，我们从第 115 帧开始跟踪白色车辆，直到最后车辆离开摄像头的第 160 帧，期间共跟踪了 45 帧。在之后的 20 次类似的跟踪实验中，我们都可以得到将近 50 帧的跟踪效果。通过实验可以看出，在选择了目标后，本文的算法可以非常稳定的进行跟踪，这是由于 Mean-shift 本身的迭代算法，使目标位置的计算趋于收敛，再加上反向投影后的特征信息，所以跟踪能够非常准确与及时。同时，我们可以看到，与目标车辆相同的车辆有两辆，但跟踪却准确的定位在目标上，那是由于我们运用了局部信息熵，使得跟踪仅仅从附近的图像区域进行匹配运算。

在之后的实验过程中，我们选择了多种视频进行跟踪实验，发现对于小目标的跟踪，例如图 2 中，跟踪球场上球员的实验，运用灰度特征进行跟踪大约可以跟踪 50 帧，而运用 HSV 特征则只能跟踪 30 帧左右。通过分析，我们发现由于灰度特征是直接通过 RGB

特征的线性组合而成，对于图像特征的提取较为完整，而在这种情况下，H 特征的提取则需要进行一定转换，这样就不可避免的丢失了一些目标的特征。而相对于大物体的跟踪，HSV 特征模型则相对要好一些，因为 H 特征代表的是目标物体的色调特征，特别是对人脸的识别跟踪，因为不管黑人白人，皮肤的色调都是一样的，只是饱和度不同而已。

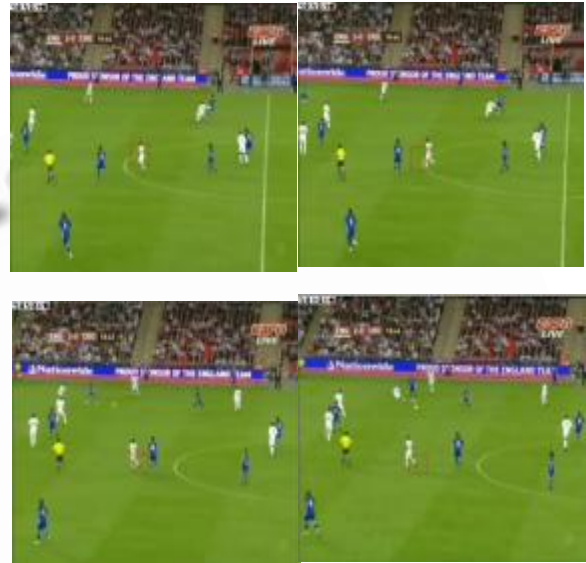


图 2 对于足球场中球员的跟踪

表 1 为不同参数对实验结果的影响。我们对同一快速移动的物体在不同参数下进行 20 次相同的实验，将特征分为 8 个，16 个，32 个特征区间，从实验结果看出 16 个特征区间效果相对最好的，收敛速度最快，不管是 HSV 还是灰度特征空间，都能对目标进行很好的跟踪。特征区间太少了，不利于目标特征的提取；而特征区间太多，又显得目标特征过于平凡，与背景差异不大，特别是在灰度空间中，几乎很难跟踪物体。

表 1 不同参数实验结果

特征区间	特征空间	跟踪帧数	迭代次数	收敛速度
8 个	HSV	34	165	4.85 次/帧
	GRAY	32	159	4.97 次/帧
16 个	HSV	35	160	4.57 次/帧
	GRAY	30	139	4.63 次/帧
32 个	HSV	25	137	5.48 次/帧
	GRAY	10	105	10.5 次/帧

最后,我们将本文提出的算法同运用 Bhattacharyya 系数作为相似度函数的 Mean-shift 算法进行比较。我们选取同一个飞驰的汽车作为跟踪目标,分别进行 20 次实验,得出的迭代结果如图 3。从图中我们可以看出,基于信息熵的跟踪算法与基于 Bhattacharyya 的算法相比,跟踪的更平稳,并且平均迭代次数更少。从实验结果可以看出,本文提出的算法有很好的鲁棒性和实时性。

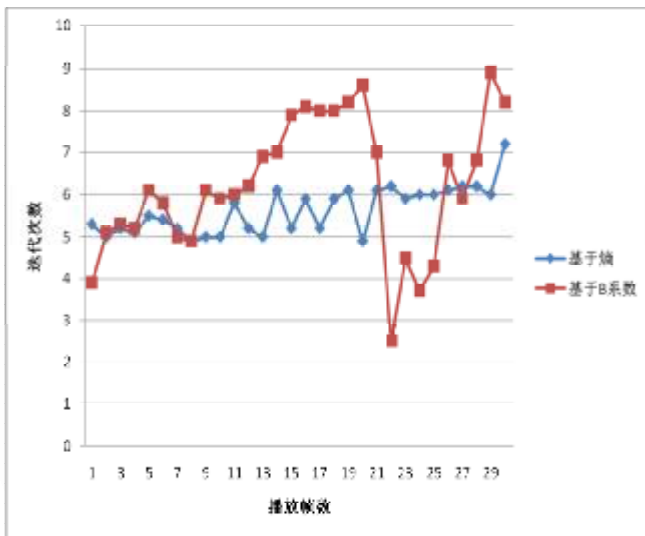


图3 不同算法迭代次数的比较

4 结论

本文将信息熵的概念引入到图像匹配中,构造了基于图像信息熵的优化目标函数,提出了基于图像信息熵的 mean-shift 算法,并对 mean-shift 算法进行了优化,加入了有关目标特征的反向投影相关系数。实验结果表明,该算法具有良好的实时性与鲁棒性。在之后的研究中,为了实现更加实用的跟踪算法,可以加入诸如卡尔曼滤波,预测估计等算法来优化算法。

参考文献

- 1 John ZM. An Information Theoretic Approach to Content Based Image Retrieval. Louisiana State University and Agricultural and Mechanical College. Phd. Thesis, 2000:45 – 62.
- 2 Lian XK. A maximal fuzzy entropy based Gaussian clustering algorithm for tracking dim moving point targets in image sequences, International Conference on Computer Science and Software Engineering, CSSE 2008, v6:54 – 57.
- 3 曾智勇,张学军,周利华,快速小波熵在图像检索中的应用.红外技术, 2005,27(6):469 – 472.
- 4 Zhuang XH, Ostevold E, Haralick R. A differential equation approach to maximum entropy image reconstruction. IEEE Trans. Speech and Signal Processing, 1987,35(2):208 – 218.
- 5 Di Zenzo S, Cinque L, Levaldi S. Image thresholding using fuzzy entropies. IEEE Transactions on. Man and Cybernetics, Part B, 1998,28(1):15 – 23.
- 6 Song Y, Li QL, Sun FC. Shannon Entropy-Based Adaptive Fusion Particle Filter for Visual Tracking. Pattern Recognition, CCPR, 2009:1 – 5.
- 7 田金文,苏康,柳健. 基于局部熵差的图像匹配方法——算法及计算机仿真.宇航学报, 1999,20(1):28 – 32.
- 8 江和平,沈振康. 基于局部交叉熵的图像匹配跟踪算法.红外与激光工程, 2005,34(6):729 – 732.
- 9 Bradski GR. Computer vision face tracking for use in a perceptual user interface. Intel Technology Journal, 1998:1 – 15.
- 10 Comaniciu D, Meer P. Mean Shift: A Robust Approach Toward Feature Space Analysis. IEEE Trans. Pattern Analysis and Machine Intelligence, 2002,24(5):603 – 619.