

基于自适应分块和 SSIM 的运动目标检测

田洪金 战荫伟

(广东工业大学计算机学院 广州 510006)

摘 要 研究了目标检测方法。针对传统背景更新方法易受噪声干扰、算法执行速度慢等弊端,对背景差分法予以改进,提出一种基于自适应图像分块和结构相似性(SSIM)的运动目标检测方法。根据视频最初几帧得到初始背景模型,再对视频后续的每帧进行自适应分块处理,利用相邻帧对应分块的结构相似性计算局部更新率,建立背景模型,将背景与当前帧差分即得到运动目标。实验结果表明,与传统的背景差分法相比,改进后的方法具有更好的检测效果。

关键词 运动目标检测,自适应分块,结构相似性

中图分类号 TP391.4 **文献标识码** A

Moving Object Detection Based on Adaptive Image Blocking and SSIM

TIAN Hong-jin ZHAN Yin-wei

(School of Computer, Guangdong University of Technology, Guangzhou 510006, China)

Abstract This paper focused on object detection. Motivated by the drawbacks of existing background update algorithms that are noise sensitive and slow in execution, an improvement on moving object detection method was proposed by image adaptive blocking and block-wise structure similarity of inter-frames. An initial background model was obtained with a few beginning frames and every successive frame was divided into blocks. Over corresponding blocks of two neighboring frames, a similarity was defined in order to update the background model. The moving objects were then obtained by subtracting the background model from the current frame. Experimental results demonstrate that the improved method has better performance than traditional methods.

Keywords Moving object detection, Adaptive blocking, Structure similarity of inter-frames

当前,实时运动目标检测中最常用的就是背景差分法^[1-8],其出发点是维持监控场景下的背景模型,将每个当前视频帧和背景模型相比较,如果同位置的像素点或区域存在一定程度的差别,则对这些像素点或区域作进一步处理,即可得到运动目标。然而实际中,场景会随时发生变化,如光照条件的渐变、光线的突变(室内环境的开关灯等)、非静止背景物体的扰动(树叶的晃动等)。因此,在运动目标检测过程中,需要构造一个好的背景模型,并对其实时更新,消除噪声的干扰,使模型能较好地表示当前场景中的变化,同时还要考虑能否达到实时性要求。文献[2,3]使用 GMM(高斯混合模型)模拟每个像素点值的变化,建立背景模型,并针对当前帧中的每一个像素点与模型进行对比分析,采用统一的学习率对背景进行更新。该方法适用于复杂环境,但在处理快速运动目标时,由于受噪声干扰,学习率很难控制在既不漏检又不过检的理想状态。之后,针对 GMM 建模,很多学者提出了改进方法:文献[4]针对固定的高斯混合模型个数,建立模型产生和退出的机制,使模型数量能自适应场景;文献[5]提出基于贝叶斯分类的混合背景建模方法,该方法根据像素值变化得到特征向量,并在此基础上利用贝叶斯准则来区分前景和背景,适用于做周期运动的复杂环境,如树叶摇晃、水面等。

以上所列举的背景模型多数是以像素为单位建立的,这

些像素被看作是相互独立的随机变量,并被单独地决策成背景或前景(目标)。尽管单个像素形式的模型具有准确、灵活等优点,但单个像素本身并不能被判为背景或前景,比如噪声点。事实上,对没有目标的区域,图像本身的结构具有相对的稳定性。例如,当外界的光线发生变化时,一般来说,背景中的像素是同时变亮或变暗。统计表明,结构稳定的背景在整个图像中占有绝对多的比例,而移动目标在图像中只占较小的比例。这样,在目标提取过程中仅对单像素进行分析会产生大量冗余信息,不可避免地会对算法的实际执行效率产生影响,为了解决这些问题,很多学者尝试将图像分块,利用图像块的特征建立背景模型。

文献[6]提出一种基于块处理技术的背景重构算法,该算法将整帧图像分成均匀的子块,对每一子块计算四阶中心距,进而判别背景区域,但四阶中心距的计算相对复杂;文献[7]给出了图像分块的方式以及几种常用的图像块特征,并且利用图像块的特征来构建自适应的高斯混合模型,但是分块大小固定,获取的运动目标边缘不够精确;文献[9]提出一种基于图像分块和 Hausdorff 距离的背景更新方法,即先采用高斯混合模型初步得到背景,再对图像进行分块,借助图像匹配的思想计算当前图像块和背景图像块间的 Hausdorff 距离,判断是否满足背景更新条件,算法能较好地克服背景变化的

到稿日期:2013-05-20 返修日期:2013-07-26

田洪金 硕士生,主要研究方向为视频监控及机器视觉;战荫伟 教授,主要研究方向为小波分析、压缩感知等。

干扰,但距离计算依然复杂,且计算结果易受单个噪声点的影响;文献[10]也在分块的基础上,利用局部自相似算子(LSS)比较当前帧与背景模型的差别,提取出前景目标的大致区域,再利用色彩信息和形态学处理提取出的运动目标,算法能较好地避免相机噪声和亮度渐变的干扰,但不能很好地移除阴影,在背景发生突变时,也很难取得较好的效果。

上述分块都是固定大小,很难适应不同大小的运动目标。因此,本文尝试设计一种可变大小的自适应分块策略,并采用图像的结构相似(SSIM)获取更新率。

SSIM 是 Wang 等人^[11,12]提出的,其原理基于自然图像具有高度的结构化,特别是相邻像素点之间具有强烈的相关性。人眼的主要功能是从视觉区域提取图像的结构信息,人类视觉系统(HVS)具有高度的适应此特性的功能。他们认为光照对于物体结构是独立的,而光照主要影响亮度和对比度,所以将亮度和对比度从图像的结构信息中分离出来,并结合结构信息对图像质量进行评价。之后,Artur Lozad 等人^[13]对此进行拓展,将 SSIM 用于视频运动目标跟踪,取代粒子滤波中的色彩和边缘直方图计算,文献[14]进一步将 SSIM 用于多目标跟踪,实验证明,它们都获得了较好的效果。

本文在上述基础上提出一种改进算法:通过对视频帧进行自适应分块和计算分块之间的结构相似性来获取局部更新率,对背景模型动态更新,得到运动目标。

本文第 1 节引入 SSIM;第 2 节给出本文方法的主要思想;第 3 节给出实验结果,并对结果进行对比分析;最后给出结论。

1 结构相似性

假设输入的图像信号为 X 和 Y ,分别提取出其亮度相似性、对比度相似性和结构相似性函数:

$$l(x, y) = \frac{2\mu_x(x, y)\mu_y(x, y) + C_1}{\mu_x^2(x, y) + \mu_y^2(x, y) + C_1} \quad (1)$$

$$c(x, y) = \frac{2\sigma_x(x, y)\sigma_y(x, y) + C_2}{\sigma_x^2(x, y) + \sigma_y^2(x, y) + C_2} \quad (2)$$

$$s(x, y) = \frac{\sigma_{xy}(x, y) + C_3}{\sigma_x(x, y)\sigma_y(x, y) + C_3} \quad (3)$$

其中,

$$\mu_x = \frac{1}{n} \sum_{i=1}^n x_i \quad (4)$$

$$\sigma_x = \left(\frac{1}{n} \sum_{i=1}^n (x_i - \mu_x)^2 \right)^{1/2} \quad (5)$$

$$\sigma_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) \quad (6)$$

x_i 和 y_i 分别为图像 X 和 Y 第 i 个像素点值; n 为图像像素点个数; μ_x 和 μ_y 分别是 X 和 Y 的平均值,表示图像的亮度; σ_x 和 σ_y 分别是 X 和 Y 的标准差,表示图像的对比度; σ_{xy} 是 X 和 Y 的相关系数,表示失真图像相对于原始图像的改变程度; C_1, C_2 和 C_3 是足够小的正常数,用来保证以上函数的分母趋近于 0 时的计算稳定性。联合式(1)~式(3),得到图像间的结构相似度:

$$S(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [s(x, y)]^\gamma \quad (7)$$

其中, α, β, γ 是调整亮度、对比度和结构 3 个部分的权重系数。一般取 $C_2 = 2C_3, \alpha = \beta = \gamma = 1$,则结构相似度的一种普通形式为:

$$S(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (8)$$

2 运动目标的检测算法

2.1 初始化

本文中,我们仅考虑灰度图像序列 $I_t, t=1, 2, \dots, I_t$,即 t 时刻的图像;如果视频是彩色的,我们只取它的亮度成分。

选取图像序列中初始 L 帧,取其像素灰度均值作为初始背景 B_0 。

2.2 自适应图像分块和相似度计算

为了能够准确地检测出运动目标,图像块大小必须适中。一般情况下,分块越大,包含的元素越多。可能存在下面两种情况:若小运动对象是块的一部分,当整个块被判为背景块时,会导致背景区域中包含小运动对象;若块处于运动对象的边缘位置,当整个块被判为背景块时,运动对象的边缘部分会被当成背景,使得背景区域中包含运动对象的边缘。

另一方面,划分的块过小,容易受单个噪声的干扰,且计算量过大。因此,本文中,我们对视频帧进行自适应大小分块。

首先,将每帧图像 I_t 从左到右、从上到下均匀地分成大小为 16×16 像素大小的子块,得到分块 $I_t^k, k=0, 1, \dots, K-1, K$ 即块的总数。这里,如果图像宽或高不是 16 的整数倍,则将图像依照复制最右边或最下边的像素的方式进行边界延拓。由于视频帧内像素点有高度的结构关系,帧间通常具有很高的相关性,根据式(8),我们计算相邻帧每一对应分块 I_t^k 和 I_{t-1}^k 的相似度 $S_{16}(I_t^k, I_{t-1}^k)$ 。

对于合适选取的阈值 T ,若 $S_{16}(I_t^k, I_{t-1}^k) \geq T$,表明块 I_t^k 和 I_{t-1}^k 间像素差别较小,当前分块中没有明显的运动目标,可归属为背景区域;若 $S_{16}(I_t^k, I_{t-1}^k) < T$,当前分块中出现了较明显的运动目标,可归属为前景区域,此时,可将分块 I_t^k 和 I_{t-1}^k 进一步划分为 8×8 的子块,图像块中前景像素的分布组合有 5 种情况,如图 1 所示,再分别计算对应的子块相似度 $S_8(I_t^i, I_{t-1}^i), i=0, 1, 2, 3$ 。针对 8×8 大小的分块 I_t^k ,可依据上述规则进一步划分为 4×4 的子块,这种多块的划分,更切合图像中实际运动物体的形状,可大大提高运动目标检测的准确度。

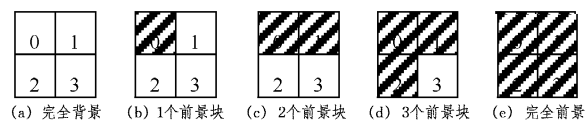


图 1 图像分块的 5 种情况

在很多实际场合,视频中每帧的大部分区域皆属于背景,前景目标只占很小一部分。因此,针对相邻帧的分块,如果两块相等,则对应分块的结构相似度为 1,这样可以避免大量的方差计算,加快检测的速度。

2.3 动态背景更新

背景差分法简单,能提供运动目标最完整的特征数据,但它对外界环境的变化非常敏感,所以,找到高效实用的背景模型并对其进行更新是这种方法的研究重点。

本文中,我们根据相邻帧对应分块的结构相似度获取不同的更新速度即更新率,以实现背景更新。设 T 为给定的阈值,则定义背景更新率为

$$\alpha_i^k = \begin{cases} 0, & S(I_i^k, I_{i-1}^k) \leq T \\ 1 - S(I_i^k, I_{i-1}^k), & S(I_i^k, I_{i-1}^k) > T \end{cases} \quad (9)$$

其中, $k=0, 1, \dots, K-1$ 。当 $S(I_i^k, I_{i-1}^k) \geq T$ 时, 当前分块中没有明显的运动目标, 可归属为背景区域。但为了避免局部噪声的影响, 需对背景进行更新, 更新率设置为 $1 - S(I_i^k, I_{i-1}^k)$; 当 $S(I_i^k, I_{i-1}^k) < T$ 时, 当前分块中出现了较明显的运动目标, 可归属为前景区域, 不对背景进行更新, 直接选取前一帧对应分块的背景作为当前分块的背景。

进而, 采取与 I_i 同样的方式对其对应的背景图像 B_i 进行分块, 得到 $B_i^k, k=0, 1, \dots, K-1$ 。这样, 对背景图像实施逐块的实时更新策略:

$$B_i^k = \alpha_k I_i^k + (1 - \alpha_k) B_{i-1}^k \quad (10)$$

联立式(9)与式(10), 求得

$$B_i^k = \begin{cases} B_{i-1}^k, & S(I_i^k, I_{i-1}^k) < T \\ (1 - R_i^k) I_i^k + R_i^k B_{i-1}^k, & S(I_i^k, I_{i-1}^k) > T \end{cases} \quad (11)$$

合并诸 B_i^k 就可得到完整的当前帧背景图像 B_i 。

2.4 运动目标的获取及后处理

承上, 将 I_i 与 B_i 差分:

$$M_i = I_i - B_i \quad (12)$$

得到运动目标的近似图像, 再对差分结果 M_i 二值化, 获取运动目标的二值近似图像, 对其进行数学形态学开、闭操作, 去除可能存在的细小残留噪声, 得到更为精确的运动目标。

2.5 阈值参数的设定

为了获取阈值 T , 我们随机选取测试视频中有运动目标出现的 3 帧, 对不同分块之间的相似度进行直方图统计分析。

图 2 为选取的 3 帧分别与其前一帧背景图像对应分块相似度的直方图, 横坐标表示各个分块间相似度大小(为便于分析, 将相似度结果乘以 255), 纵坐标表示同一相似度下分块的个数。从图中看出, 在标注黑点右边的分块数, 即相似度较大时, 图像块数较为集中且总数多, 而一幅图像中大部分分块往往集中在背景区域, 少部分分块集中在前景区域, 则这些分块可判为背景区域; 在黑点左边, 即相似度较小时, 块数较少, 可认为有较明显的运动目标, 判为前景区域。经过实验统计分析, 得到更新阈值参数 $T, T \in (0.8, 0.9)$, 本文中取 0.84 作为最终阈值。

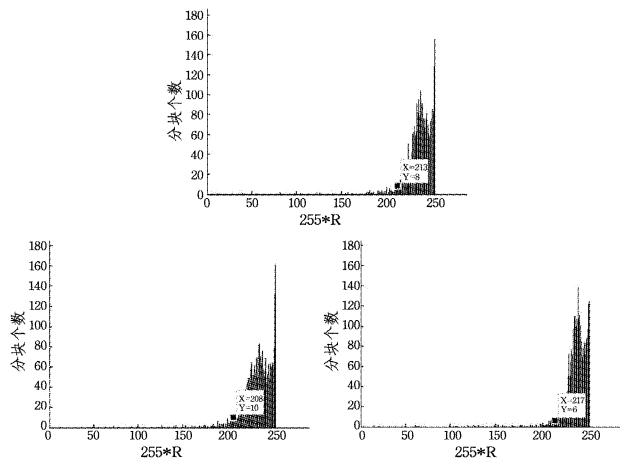


图 2 随机选取的 3 帧与前一帧背景对应分块相似度的直方图

3 实验结果及分析

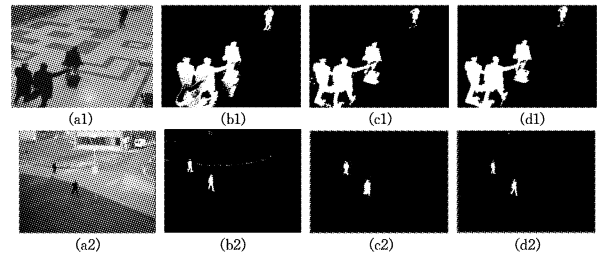
本文从 PETS^[15] 数据库中取两段视频进行实验对比, 每帧大小为 480×360 , 帧率为 14 fps。实验计算机 CPU 为 Intel

Pentium 1.86 GHz, 操作系统为 Windows XP, 视频处理平台为 VS2008 和 OpenCV 2.0。

将本文的算法与高斯混合背景建模算法和文献[9]的算法进行对比分析。高斯混合背景建模算法对每个像素点进行建模, 检测结果出现噪声点, 且由于采用固定的更新系数来更新各个高斯分布的权重, 若更新得过快, 会降低对运动目标检测的灵敏度, 若更新过慢, 则易产生阴影, 此算法得到的运动目标轮廓也不够清晰。另外, 由表 1 知, 混合高斯背景建模算法对每个像素建立多个模型; 文献[9]提出的算法采用 Hausdorff 距离来计算块间的不匹配性, 得到更新率, 这些都增加了计算机处理的时间。而本文的方法是基于自适应分块处理的, 算法时间消耗较少, 且针对不同的分块采取不同的更新系数, 所以在很大程度上能满足实时性要求。实验结果如图 3 所示。

表 1 算法时间消耗对比

	高斯混合背景建模	文献[9]算法	本文算法
Tree 视频	640.5	150.4	95.3
PETs 视频	1709.6	399.3	228.7



(a1)和(a2)为原始视频帧;(b1)和(b2)为 GMM 算法[5]的结果;(c1)和(c2)为文献[10]算法的结果;(d1)和(d2)为本文算法的结果

图 3 实验结果

结束语 在实际的视频序列中, 运动目标的尺寸往往不一致, 且感兴趣前景目标区域与背景区域灰度值变化大小往往不一致, 采用统一的分块会导致背景模型不够精确, 提取的目标边缘模糊。因此本文利用帧间对应分块的结构相似度来获取更新率, 并以此对图像自适应分块, 最终建立一个实时更新的背景参考模型, 该模型能够满足实时性的运动目标检测, 达到了较好的效果。

本文算法的局限在于只适用于固定场景下的目标检测, 针对动态场景下的运动目标检测, 尚有很多问题亟待解决, 这是我们下一步的工作。

参考文献

- [1] Elhabian S Y, El-Sayed K M, Ahmed S H. Moving Object Detection in Spatial Domain using Background Removal Techniques—State-of-Art [C]//Recent Patents on Computer Science, 2008, 1:32-54
- [2] Stauffer C, Grimson E. Learning Patterns of Activity Using Real Time Tracking [J]. IEEE Transactions on Pattern Recognition and Machine Intelligence, 2000, 22(8):747-757
- [3] Stauffer C, Grimson W E L. Adaptive Background Mixture Models for Real-Time Tracking [C]//Computer Vision and Pattern Recognition, 1999
- [4] 梁华, 刘云辉. 自适应多模快速背景差方法[J]. 中国图象图形学报, 2008, 13(2):345-350

[5] Li Li-yuan, Huang Wei-min, Irene Y H, et al. Foreground Object Detection from Videos Containing Complex Background [C]// Proceedings of the eleventh ACM international conference on Multimedia, 2003; 2-10

[6] 张文涛, 李晓峰, 李在铭. 高速密集视频目标场景下的运动分析 [J]. 电子学报, 2000, 28(10): 114-117

[7] 杨广林, 孔令富. 基于图像分块的背景模型构建方法 [J]. 机器人, 2007, 29(1): 29-34

[8] 姚春莲, 周兵. 运动对象检测及其在视频压缩与处理中的应用 [M]. 北京: 冶金工业出版社, 2010

[9] 李庆武, 王敏, 陈琦, 等. 基于图像分块和 Hausdorff 距离的背景更新方法 [J]. 测试技术学报, 2011, 25(6): 544-547

[10] Jodoin J P, Bilodeau G A, Saunier N. Background subtraction based on local shape [J]. arXiv preprint arXiv:1204.6326, 2012

[11] Wang Zhou, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612

[12] Wang Zhou, Bovik A C, Simoncelli. Structural approaches to image quality assessment [J]. Handbook of Image and Video Processing, 2005, 7: 18

[13] Loza A, Mihaylova L, Canagarajah N, et al. Structural Similarity-Based Object Tracking in Video Sequences [C]// 2006 9th International Conference on Information Fusion, IEEE, 2006; 1-6

[14] Loza A, Mihaylova L, Bull D, et al. Structural Similarity-Based Object Tracking in Multimodality Surveillance Videos [J]. Machine Vision and Applications, 2009, 20(2): 71-83

[15] PETS 2006 benchmark data [OL]. <http://www.pets2006.net>, 2006

(上接第 113 页)

表 3 中可以看出, CRAC-lift 在少数类 C_2 、 C_4 、 C_5 中产生更多的规则数。

表 3 产生各类别的规则数对比

类	C_1	C_2	C_3	C_4	C_5
支持度	76.1%	11.0%	7.5%	4.5%	0.9%
CRAC-conf	21	7	1	1	0
CRAC-lift	18	11	1	2	3

表 4 和表 5 显示了 CRAC-lift 和 CRAC-conf 的混淆矩阵。在 89 个测试实例中, CRAC-lift 将少数类 C_5 的一个实例做出了正确预测, 而 CRAC-conf 没有匹配该实例的规则。CRAC-lift 正确分类了 C_2 中的 10 个实例, 而 CRAC-conf 只正确分类了 7 个实例, 另外 3 个没有匹配规则。因此, CRAC-lift 主要提高了少数类的分类准确率。

表 4 CRAC-lift 在 Anneal 上的混淆矩阵

	C_1	C_2	C_3	C_4	C_5	NoRules
C_1	68			1		
C_2		10				
C_3			6			
C_4				3		
C_5					1	

表 5 CRAC-conf 在 Anneal 上的混淆矩阵

	C_1	C_2	C_3	C_4	C_5	NoRules
C_1	68					1
C_2		7				3
C_3			6			
C_4				3		
C_5					0	1

结束语 在不平衡数据中, 正确分类少数类的实例具有更重要的意义。而置信度、支持度等度量都偏向产生更多的多数类的规则, 使得少数类的规则较少, 甚至没有。本文提出了一种基于相关规则的不平衡数据的关联分类。以提升度来挑选规则, 将产生更多少数类的规则。以结合提升度、支持度和 CCS 的强度排序规则, 使得少数类的规则的优先级较高, 因而在分类少数类实例时, 有可能优先选中少数类规则做出正确分类。实验表明, 该算法取得了较高的平均分类准确率且在分类少数类的实例时具有更高的准确率。

参考文献

[1] Liu B, Hsu W, Ma Y. Integrating classification and association rule mining [C]// Proc of the 4th International Conference on Knowledge Discovery and Data Mining (KDD'98), 1998; 80-86

[2] Li W, Han J, Pei J. CMAR: Accurate and efficient classification based on multiple class-association rules [C]// Proc of the 1st International Conference on Data Mining, 2001; 369-376

[3] Yin X, Han J. CPAR: classification based on predictive association rules [C]// Proc of the SIAM International Conference on Data Mining (SDM'03), 2003; 331-335

[4] Dong G, Zhang X, Wong L, et al. CAEP: Classification by aggregating emerging patterns [C]// Discovery Science, Springer Berlin Heidelberg, 1999; 30-42

[5] Wang J, Karypis G. HARMONY: Efficiently mining the best rules for classification [C]// Proc. of SDM, 2005; 205-216

[6] Quinlan J R. C4. 5; programs for machine learning [M]. Morgan kaufmann, 1993

[7] Verhein F, Chawla S. Using significant, positively associated and relatively class correlated rules for associative classification of imbalanced datasets [C]// Seventh IEEE International Conference on Data Mining, 2007, ICDM 2007. IEEE, 2007; 679-684

[8] Arunasalam B, Chawla S. CCCS: a top-down associative classifier for imbalanced class distribution [C]// Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 2006; 517-522

[9] Omiecinski E R. Alternative interest measures for mining associations in databases [J]. IEEE Transactions on Knowledge and Data Engineering, 2003, 15(1): 57-69

[10] Zhao Y, Karypis G. Criterion functions for document clustering: Experiments and analysis [Z]. Machine Learning, 2001

[11] Agrawal R, Srikant R. Fast algorithms for mining association rules [C]// Proc of the 20th International Conference on Very Large Data Bases (VLDB'94), 1994; 487-499

[12] Thabtah F A, Cowling P, Peng Y. MMAC: A New Multi-class, Multi-label Associative Classification Approach [C]// Proc of the 4th International Conference on Data Mining (ICDM'04), 2004; 217-224

[13] CBA; <http://www.comp.nus.edu.sg/dm2/p-download.html>