

附件1：视频分析技术挑战赛评测数据集

PKU-SVD-B 数据库由北京大学视频编解码技术国家工程实验室联合北京大学保卫部建立并整理，主要用于全国研究生智慧城市技术与创意设计大赛“视频分析技术挑战赛”。本说明将从命名规则、视频内容、存储方式和共享要求几个方面对PKU-SVD-B 数据库库进行简单的介绍。

一、命名规则

将数据集按照比赛类型、比赛阶段、数据用途和数据难度划分，命名方法为：

T_*

S_*

Dx_*

L_*

- 1) ***比赛类型*** (1位): 1 表示技术擂台赛, 2 表示决赛
- 2) ***数据用途*** (1位): 1 表示训练数据集, 2 表示验证数据集, 3 表示决赛测试数据集
- 3) ***任务类型*** (2位): 01~06 分别对应挑战赛中不同任务。
- 4) ***难易类型*** (1位): 0 为容易, 1 为中等, 2 为困难

对数据库中视频文件的命名方法为：

[name]_

x_

*[wid]*x*[hei]_*

xx_

Px

. [ext]

1) *名称*

视频名称，限定 64 个字符。

2) *场景* (1 位):

表示拍摄视频的摄像机编号。

3) *分辨率* (8 位到 10 位) **

变量 hei 表示视频图像的宽度，用数字表示，位宽 3-4 位。wid 表示图像高度，用数字表示，位宽 3-4 位。wid×hei，格式示例为 720x576, 1600x1200 等。

4) *帧率* (2 位) **

表示视频拍摄帧率。用数字表示，不足部分前面补零，如 10, 15, 30**

5) *公开情况* (2 位) **

公开情况 (两位)，R 表示公开标记，后一位表示该视频是否可以对外界公开，0 表示不可以公开，1 表示可以公开。

6) *后缀名* (2 位到 5 位) **

变量 ext 表示视频的后缀名。包括 avi, yuv, ts, asm, avses, avsts 等。

二、视频内容

PKU-SVD-B 数据库中包含：来自北京大学校园内 16 个摄像头的监控视频，分辨率为 1080p 和 720p。比赛数据集按场景、难易等分为两级，分别用于挑战赛的各任务。根据用途，比赛数据集分为训练数据集、校验数据集、决赛测试数据集三类。

三、存储方式

在 PKU-SVD-B 数据库目录下，我们按照视频的识别号建立目录，同一个视频场景下的序列在同一个目录下。

所有视频按 AVS-S2 标准进行编码存储。组织方将提供该编码标准的视频解码库及程序解析该格式视频的 API。

四、共享要求

在 PKU-SVD-B 数据库的建立过程中，北京大学视频编解码国家工程实验室与联合单位签署了《PKU-SVD-B 数据库使用许可协议》，协议中规定监控视频库中的所有视频未经北京大学许可不得用于任何商业用途和网络共享。如果要获得北京大学视频编解码国家工程实验室共享的北大视频监控比赛数据库，申请者（须是正式工作人员或加盖院系章）必须与北京大学视频编解码国家工程实验室签订《PKU-SVD-B 数据库共享协议》。

附件2：视频分析技术挑战赛评测数据集使用协议

PKU-SVD-B 数据库由北京大学视频编解码技术国家工程实验室联合北京大学保卫部建立并整理，主要用于全国研究生智慧城市技术与创意设计大赛“视频分析技术挑战赛”。因此，PKU-SVD-B 数据库目前仅用于科学研究。所有使用 PKU-SVD-B 数据库的科研人员必须签署本协定并遵守协定所列的条款限制。如有不遵守本协定的行为发生，将被取消获取本数据库及后续版本的权利，并将被追究民事责任。

所有科研人员需要遵守 PKU-SVD-B 数据库的下列约束条款：

1. 仅本数据库中公开数据可以并仅可以发表在监控视频分析领域的论文或者技术报告中，对于未公开部分不准以任何形式出现在任何的文档中。数据库内所有数据不准发布在商业资料、新闻报刊或者其他公共媒体上。
2. 所有使用了 PKU-SVD-B 数据库的科技论文、文档和报告必须在致谢中表明使用了本数据库。所有使用了 PKU-SVD-B 数据库的科技论文、文档和报告必须立即转发给北京大学视频编解码技术国家工程实验室。
3. 不论盈利与否，均不得采用任何方式对本数据库进一步发布、发表、复制或者进一步传播。本条款之限定包括应其他大学、单位或者公司的要求，进一步分发、复制或者发布。
4. 本数据库内所有数据仅可用于科学研究。本数据库整体或者部分均不得以任何方式用于商业用途。
5. 本协定最终解释权归北京大学视频编解码技术国家工程实验室所有。

组织名称（盖章）: _____

姓名: _____ 职务/职称: _____

签名: _____

日期: _____

通讯地址: _____

电话: _____

传真: _____

电子邮箱: _____

附件3、视频分析技术挑战赛数据管理的保密制度

- 1、【目的】为保证比赛的公平性和严肃性，特制定技术挑战赛过程中应遵循的数据管理保密制度（以下简称“本制度”）。
- 2、【适用范围】根据用途，比赛数据集分为训练数据集、校验数据集、决赛测试数据集三类。本制度的适用范围，包括技术挑战赛的验证数据集和决赛测试数据集；公开的训练数据集不在本制度的适用范围内。
- 3、【数据产生方式】整个技术挑战赛的数据采集自来自北京大学校园内 20 个摄像头的监控视频，对每类任务采集的数据量不少于 3 小时，其中以 15 分钟作为一个基本数据单元。上述三类数据采用统一的标注标准和方法。
- 4、【数据划分方式】在所有数据采集标注完成后，在专家委员会的监督下，采用数据集划分程序以随机的方式产生验证数据集和决赛测试数据集，所得的划分结果封存在教育部学位与研究生教育发展中心。
- 5、【数据分发】在技术挑战赛的擂台赛或决赛阶段开始之前，在专家委员会的监督下，从教育部学位与研究生教育发展中心获取相应的验证数据集或决赛测试数据集，并进入比赛评测系统。

附件4：视频分析技术挑战赛评测指标

视频分析技术挑战赛采用国际通用的评测指标，根据参赛队伍提交的结果或现场运行产生的结果进行评测，比赛排名分任务按指标进行排序。不同任务的具体评测指标如下：

一、单摄像头行人评测

参照 PASCAL VOC 比赛评测方法，对于任意一个检测结果，如果与 Ground Truth 的唯一标注的矩形框满足公式（1），则被认为是正检，否则为误检，多个检测框对应一个标注框，则只有其中一个检测框被认为正检，其余均为误检。

$$\alpha = \frac{\text{area}(B_p \cap B_{gt})}{\text{area}(B_p \cup B_{gt})} \geq 50\% \quad (1)$$

其中 B_p 为所提交的任一检测结果框， B_{gt} 为 Ground Truth 标注框，函数 area • 用于计算矩形框的面积。

根据上述正检判断规则，计算检测的召回率 Recall 和准确率 Precision 以及 Fscore 作为评测指标，其中 Fscore 为大赛第一排序指标，该值越大说明算法性能越好。

$$\text{Recall} = \frac{t_p}{t_p + t_n}, \text{Precision} = \frac{t_p}{t_p + f_p}, \text{Fscore} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (2)$$

其中， t_p 为正检数目， f_p 为误检数目， t_n 为漏检数目。

二、单摄像头多类对象检测评测

参照任务一评测方案，但是输出需要带有类别标签，必须在类别标签一致前提下同时满足公式（1）约束条件才能判别为正检。例如，车

的检测结果只能和车的 **Ground Truth** 比对，而不能和人的 **Ground Truth** 比对。

三、单摄像头指定对象跟踪

参照 **PETS** 比赛中对象跟踪评测方法，选取 *MOTA* 和 *MOTP* 作为对象跟踪性能评测指标，该值越大说明跟踪算法性能越好。

$$MOTA = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t}, MOTP = \frac{\sum_{i,t} d_t^i}{\sum_t c_t} \quad (3)$$

其中， m_t 为漏跟目标个数， fp_t 为误跟目标个数， mme_t 则为跟踪轨迹 id switch 次数； d_t^i 为第 t 帧的第 i 个正确的跟踪框， c_t 为第 t 帧的 **Ground Truth** 目标个数。

四、跨摄像头指定行人跟踪

对所有 N 个摄像头中的每个摄像头，在时间轴上设置 $T_n (n = 1, \dots, N)$ 个匹配节点，每个匹配节点在时间轴上具有一定的范围（例如前后 10 帧）。如果某个对象的运动轨迹在这个时间范围内经过一个摄像头，则认为这条轨迹包含了对应的匹配节点。

对于某一个指定跟踪对象，在某个包含的匹配节点范围内，当提交的跟踪轨迹和 **Ground Truth** 跟踪轨迹有超过 50% 的帧都能匹配上（参照对象检测匹配方式）时，则认定该跟踪对象在跨头跟踪时匹配到了该匹配节点。

对于第 i 个跟踪对象，评测指标为：

$$Recall(i) = \frac{\sum_{n=1}^N C_n^i}{\sum_{n=1}^N G_n^i}; \quad Precision(i) = \frac{\sum_{n=1}^N C_n^i}{\sum_{n=1}^N D_n^i}; \quad (4)$$

其中 N 是总的摄像头数， $G_n^i (G_n^i < T_n)$ 是第 i 个指定跟踪对象的跟踪

Ground Truth 轨迹在摄像头 n 里包含的匹配节点的数目。 C_n^i 是第 i 个指定跟踪对象在摄像头 n 里匹配到的匹配节点的数目。 D_n^i 是第 i 个指定跟踪对象在摄像头 n 里包含的匹配节点的数目。



假设上图为某个指定跟踪对象的评测过程：时间轴上标识的是不同的匹配节点；提交结果为提交的跟踪轨迹所包含的匹配节点（1号摄像头下1个，3号摄像头下1个，7号摄像头下2个，8号摄像头下1个，共5个）；标准结果为当前对象的**GroundTruth**轨迹包含的匹配节点（1号摄像头下2个，3号摄像头下1个，6号摄像头下2个，8号摄像头下2个，共7个）。由图可知，匹配到的节点为：1号摄像头下第一个节点，3号摄像头的节点，8号摄像头的第二个节点，共3个。

因此，这条轨迹的评测结果为： $Recall = 3 / 7$, $Precision = 3 / 5$ 。

对于所有的指定跟踪目标，评测指标为：

$$Recall = \frac{\sum_{i=1}^M \sum_{n=1}^N C_n^i}{\sum_{i=1}^M \sum_{n=1}^N G_n^i}; \quad Precision = \frac{\sum_{i=1}^M \sum_{n=1}^N C_n^i}{\sum_{i=1}^M \sum_{n=1}^N D_n^i}; \quad Fscore = \frac{2 * Recall * Precision}{Recall + Precision} \quad (5)$$

其中 M 是所有被指定跟踪目标对象数目。

五、人脸检测与识别

人脸检测的评测参照比赛任务一的评测方法，阈值 $\alpha=40\%$ 。

对于人脸识别任务（给定待识别的人脸库（含证件照），要求检测

并识别出摄像头视域内出现的目标人脸), Watchlist 由采集对象的证件照构成。每个采集对象对应一张或多张证件照, 有唯一标识的 id。

测试集 Query Set 由采集对象进出校门的监控视频帧构成, 以 ROC (Receiving Operator Characteristic) 曲线来评判识别性能, 曲线的横纵坐标分别为 *Recall* 和 *1-Precision*。识别结果评测类似于任务二评测, 即识别出的目标 id 必须与 Watchlist 中的目标 id 一致。