|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | | |  |
|  | | |
|  | | |
| 方圆ok1 | | | |
|  | 面向5G的VR业务识别技术高效合作需求 | | 附件1-16K |
|  | |
| **文档版本** | **01** |
| **发布日期** | **2018-08-**27 |
| **华为技术有限公司** | |

|  |
| --- |
| 版权所有 © 华为技术有限公司2018。 保留一切权利。  非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。  商标声明  附件3-版权声明页图和其他华为商标均为华为技术有限公司的商标。  本文档提及的其他所有商标或注册商标，由各自的所有人拥有。  注意  您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。  由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。 |

|  |  |
| --- | --- |
| 华为技术有限公司 | |
| 地址： | 深圳市龙岗区坂田华为总部办公楼 邮编：518129 |
| 网址： | http://[enterprise](http://enterprise.huawei.com).huawei.com |

前 言

概述

本文档介绍合作内容，支撑的业务创新

读者对象

本文档主要适用于以下读者：

* 技术工程师
* 维护工程师
* 合作管理部

修改记录

修改记录累积了每次文档更新的说明。最新版本的文档包含以前所有文档版本的更新内容。

文档版本 01 (2018-08-27)

目 录

[前 言 ii](#_Toc516732550)

[目 录 iii](#_Toc516732551)

[1 整体 4](#_Toc516732552)

[1.1 业务需求 4](#_Toc516732553)

[1.2 采集源需求 4](#_Toc516732554)

[1.3 创新需求 5](#_Toc516732555)

[1.4 工具化需求 5](#_Toc516732556)

[2 业务需求描述 5](#_Toc516732557)

[2.1 视频内容解析 5](#_Toc516732558)

[2.1.1 场景描述 5](#_Toc516732559)

[2.1.2 需求描述 6](#_Toc516732560)

[2.2 用户内容推荐 7](#_Toc516732561)

[2.2.1 场景描述 7](#_Toc516732562)

[2.2.2 需求描述 7](#_Toc516732563)

[2.3 内容库融合 8](#_Toc516732564)

[2.3.1 场景描述 8](#_Toc516732565)

[2.3.2 需求描述 9](#_Toc516732566)

# 需求描述

## 背景描述

本次合作保障部门业务识别方面的业界竞争力并在5G相关业务识别方面取得领先。

## 概念漂移的流量识别算法

### 业务描述

为保证概念漂移下的业务识别能力，需要东南大学合作课题组实现：

* **概念漂移算法：**研究适用于网络业务识别的概念漂移算法，在发生概念漂移时，算法的预测准确性不下降或下降在一定阈值范围。

### 需求描述

基于上述的业务需求，完成如下研究：

|  |  |  |
| --- | --- | --- |
| **序号** | **需求名称** | **验证要求** |
| 1 | 概念漂移理论及概念漂移下的业务识别调研 | 输出调研报告内容包含：   * 概念漂移技术的发展，及其在业务识别领域应用调研。 |
| 2 | 概念漂移业务识别研究数据库（招募志愿者采集研究用的数据样本，并对样本进行标记） | 输出样本库，样本数据库应满足：   * 1）覆盖主要的业务场景，各类样本量均衡，为保证样本的多样性，不能为相同场景多次拨测，应覆盖：1）观看视频：YouTube/Netflix/腾讯视频,各100条码流；2）听音频：Spotify/QQ音乐/酷狗音乐,各100条码流；3）网页浏览：Amazon/淘宝/今日头条，各100条码流，浏览3分钟以上；4）发送和接收图片、语音(<2M)：微信/fbmsg/line,各100条码流；5）上传和下载大文件(>2M)：微信//line,各100条码流；6）文本聊天：微信/fbmsg/line,各100条码流，每条聊天3分钟以上；7）Video Call over IP:微信/fbmsg/line视频通话，各100条码流，通话3分钟以上；8）Voice Call over IP:微信/fbmsg/line语音通话，各100条码流，通话3分钟以上； * 2）完备的过程信息记录：一条拨测样本应包含原始数据pcap包、录屏文件、配套的信息记录文档。 |
| 3 | 概念漂移算法原型 | * 概念漂移检测算法具备一定的实时测能力； * 对样本库中的几类业务概念漂移识别率>85%，挑战95%; * 概念漂移时模型自更新后分类精度衰减<5%； * 算法代码，无逻辑错误和影响准确性的bug。 |

## AR/VR业务识别算法

### 业务描述

为了在AR/VR的业务识别能力上取得突破，需要东南大学合作课题组实现：

* **VR流量识别：**在众多具有多其他业务的流量中识别出VR流量；
* **VR场景识别：**1）识别VR流量到具体场景，如游戏、商城、视频等；2）主视野与辅视野识别；3）编码方式识别与分辨率识别；4）VR运动到图形的响应时延计算。

### 需求描述

基于上述的业务需求，完成如下研究：

|  |  |  |
| --- | --- | --- |
| **序号** | **需求名称** | **验证要求** |
| 1 | AR/VR调研 | 调研报告包含如下内容：   * AR/VR技术及应用场景调研和AR/VR业务识别调研。 |
| 2 | VR研究数据库 | * 覆盖VR主要应用场景：游戏/视频/购物各100条，5分钟以上； * 包含各种编码方式和分辨率的VR视频； |
| 3 | VR识别算法原型 | * 从混杂有浏览、消息、视频点播等的流量中识别出VR流量，识别率>85%，挑战95%； * 在VR流量中识别具体场景游戏、商城、视频准确率>85%，挑战95%； * VR主/辅视野识别准确率>85%，挑战95%； * VR编码格式识别率>85%,分辨率识别率>85%，挑战95%; * M2P,M2V,M2S时延预测平均偏差<2s * 算法代码，无逻辑错误和影响准确性的bug |

## 加密视频分辨率识别算法

### 业务描述

为提升分辨率与码率的识别能力，需要东南大学合作课题组实现：

* **加密视频分辨率识别：**能够实现加密视频点播业务的实时分辨率，并能在分辨率发生切换时得以检测。
* **加密视频码率识别：**能够实现加密视频点播业务的分片码率检测。
* **加密视频广告识别：**能够识别视频播放前的广告流量。

### 需求描述

基于上述的业务需求，完成如下研究：

|  |  |  |
| --- | --- | --- |
| **序号** | **需求名称** | **验证要求** |
| 1 | 分辨率和码率识别调研 | 输出识别的调用报告包含：   * 分辨率/码率技术调研，以及业界和学术界的识别方案 |
| 2 | 分辨率和码率识别算法原型 | * 算法覆盖TCP和QUIC传输的点播视频分辨率和码率识别； * 实现单次拨测级别的分辨率预测，准确率>85%，挑战95%； * 实现分片粒度的码率预测，准确率>80%（偏差小于50kbps认为预测准确），挑战90%； * 实现视频播放前广告流的识别，准确率>80%，挑战视频中插播广告的识别； * 算法代码，无逻辑错误和影响准确性的bug。 |