Gaze 智能视频云服务

Contents

[Gaze 智能视频云服务 1](#_Toc519502261)

[简介 1](#_Toc519502262)

[设计 1](#_Toc519502263)

[系统架构 2](#_Toc519502264)

[分层架构图 2](#_Toc519502265)

[开发者的角色 2](#_Toc519502266)

[计算图 3](#_Toc519502267)

[节点类别 4](#_Toc519502268)

[执行引擎 5](#_Toc519502269)

[内置组件 5](#_Toc519502270)

[视频源 5](#_Toc519502271)

[中间节点 5](#_Toc519502272)

[汇聚点 6](#_Toc519502273)

[Web界面 6](#_Toc519502274)

[组件录入 6](#_Toc519502275)

[计算图录入 6](#_Toc519502276)

# 简介

Gaze 实时视频分析服务平台整合私有云的计算能力和物联网设备(如摄像头)的数据，在端(Edge)提供可扩展、可伸缩的视频分析能力。

# 设计

摄像头在当今社会中无处不在，无时无刻不在收集大量数据。然而这些数据就像“暗物质”，蕴含的知识，无人知晓；庞大的价值，并未充分发掘。

鉴于此，我们提出 Gaze 来整合私有云的计算能力和物联网设备(如摄像头)的数据。在端(Edge)提供可扩展、可伸缩的视频分析能力。

## 系统架构

### 分层架构图

下图是Gaze的分层架构图：



图1 分层架构图

分布式执行引擎：

1. 基于Kubernetes组织云计算和物联网设备资源；
2. 基于GStreamer 和 OpenCV等开源库处理视频流和图像；
3. 为上层提供计算图执行服务。

## 开发者的角色

基于Gaze平台做视频分析的开发者可以分为：

* 组件开发者(dev 1)：他们有计算机视觉背景，负责开发视频认知与分析组件。可能的组件包括：车辆计数、行人计数、车牌识别、通缉犯识别、特定事件(如斗殴) 识别等等。
* 应用开发者(dev 2)：他们贴近应用场景，理解需求，负责开发计算图，把上述组件组装成完整的应用，来解决特定问题。



图2 系统数据流图

由上面系统数据流图所示，Gaze工作的基本流程为：

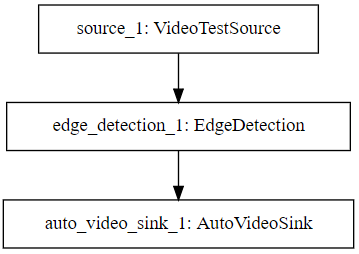
1. 组件开发者创建的视频认知与分析组件，通过Gaze的Web界面录入，并存储；
2. 应用开发者创建计算图，通过Gaze的Web界面录入，并存储；
3. 执行引擎获取计算图，生成容器镜像，创建Pod，并通过Helm charts调度Kubernetes进行编配；
4. 执行引擎获取已分配容器的IP地址/端口；
5. 执行引擎调度视频源，向指定IP地址/端口发送视频流；
6. (公有/私有)云中容器对视频做实时处理；
7. 处理完毕，容器将视频流发送给汇聚点。根据组件定义，汇聚点可以是监视屏幕、大数据存储、Event Hub、Kafka队列等。

## 计算图

应用计算图也叫数据流图，是一种有向无环图。视频分析中每个步骤都是计算图中的一个节点，而节点之间的边描述了计算之间的依赖关系。

下面是最简单的计算图的代码示例和可视化效果：

|  |
| --- |
| x = VideoTestSource()  x = EdgeDetection()(x)  x = AutoVideoSink()(x) |



最简计算图的可视化效果

这段代码基于内置的VideoTestSource类对象生成视频流，用EdgeDetection类对象做基本处理，最后发给AutoVideoSink类对象展示在窗口。

### 节点类别

计算图节点可分为3类：

1. 视频源：这种节点没有输入，只有输出。如“测试视频源“、”摄像头视频源”、”文件视频源“、“网络流视频源”等；
2. 分析节点：既有输入，也有输出。输入是一到多条视频流，输出是视频流或事件流。

如“车辆计数”、 “行人计数”、“车牌识别”、“通缉犯识别”、“特定事件(如斗殴) 识别”等；

1. 汇聚点：只有输入，没有输出。如“Hadoop存储汇聚点“、“Event Hub事件汇聚点”、“Kafka事件汇聚点”、“窗口显示汇聚点”、“网络流汇聚点”等。

## 执行引擎

执行引擎把计算图编译成

## 内置组件

### 视频源

#### 测试视频源

此组件用于生成各种格式的测试视频数据。默认情况下，此组件将无限期地生成数据。下图是从“测试视频源“截取的静态图片：

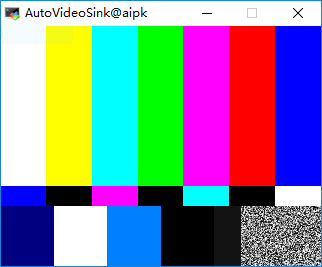


图3 测试视频源

#### 摄像头视频源

此组件用于从网络摄像头设备捕获视频。

#### 文件视频源

从本地文件系统中的视频文件读取数据。

#### 网络流视频源

从网络读取RTP流UDP数据包的网络源。 下图是从“网络流视频源“截取的静态图片：

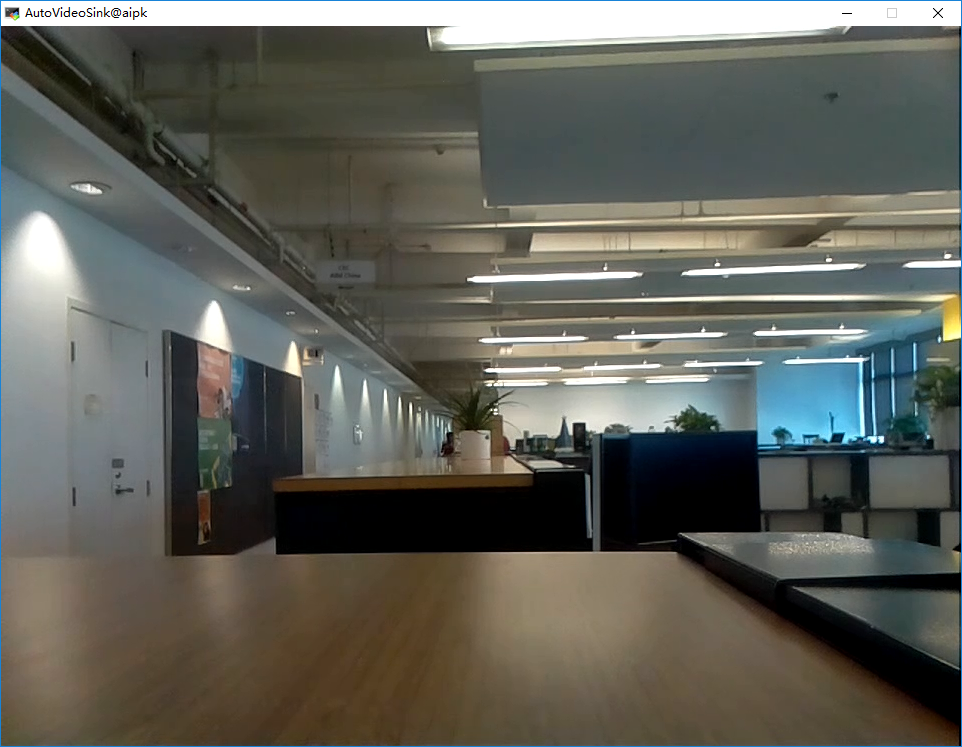


图4 网络流视频源

### 中间节点

中间节点负责视频的处理，下面是一些中间节点示例：

#### 边缘检测

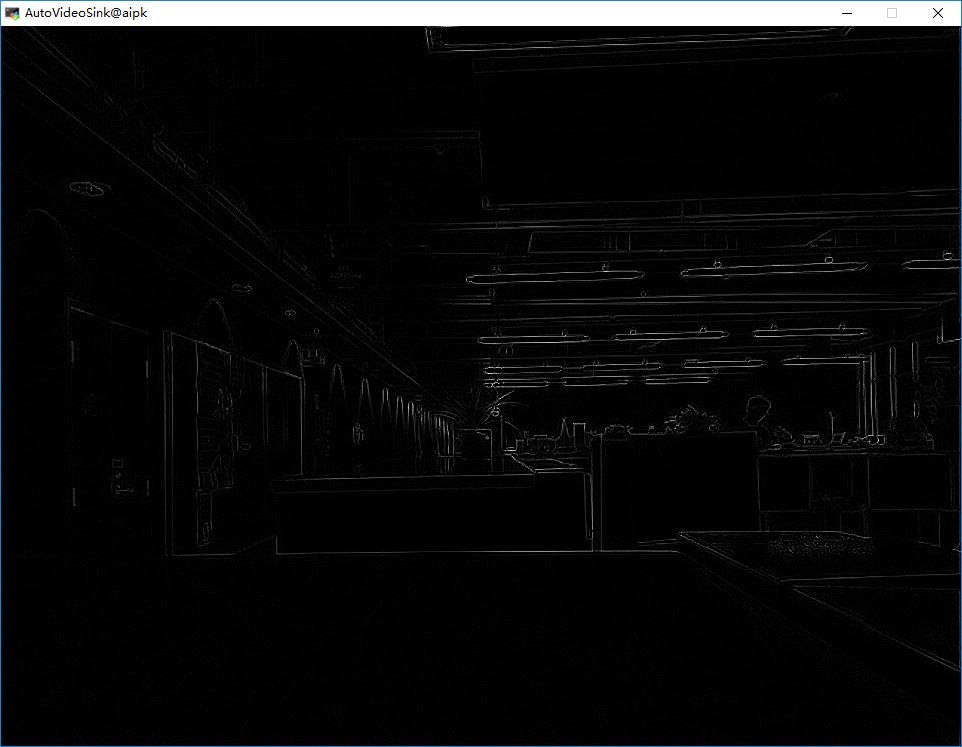


图5 基于网络摄像头的边缘检测效果

#### 行人计数

统计一定时间段内在视频流中出现的人数，核心计数为目标检测和目标跟踪技术。

#### 特定事件识别

指通过分析视频对行人的行为进行识别的技术。分为个体识别和群体识别，主要基于人体运动的图像序列，结合高精度的骨架估计算法，提取人体骨架运动序列，从而实现行为识别。

### 汇聚点

#### Hadoop汇聚点

将传入数据写入分布式文件系统HDFS中的文件。

#### 自动视频汇聚点

将传入数据以实时视频的形式在窗口显示。

## Web界面

管理员在Web界面可以浏览有多少个应用在执行

### 组件录入

### 计算图录入