ĐẠI HỌC BÁCH KHOA HÀ NỘI

VIỆN CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

─────── \* ───────

ĐỒ ÁN 3

**STREAM VIDEO TỪ CAMERA IP**

BỘ MÔN KỸ THUẬT MÁY TÍNH

Sinh viên thực hiện: Vương Chí Sơn

MSSV: 20156407

Lớp CN CNTT 3 - k60

Giảng viên hướng dẫn: TS. Nguyễn Hồng Quang

HÀ NỘI 5-2018

Mục lục

[1. Tìm hiểu chung 4](#_Toc517096281)

[1.1. Các phần mềm thực tế 5](#_Toc517096282)

[1.1.1. Yoosee app 5](#_Toc517096283)

[1.1.2. CMS Client 6](#_Toc517096284)

[1.2. Chuẩn ONVIF 6](#_Toc517096285)

[1.3. Giao thức RTSP 7](#_Toc517096286)

[2. Dùng TCP, UDP tạo gói tin RTSP 9](#_Toc517096287)

[3. Dùng thư viện OpenCV 11](#_Toc517096288)

[3.1. Sơ lược về OpenCV 11](#_Toc517096289)

[3.2. Cài đặt OpenCV 12](#_Toc517096290)

[3.2. Sơ đồ khối 13](#_Toc517096291)

[4. Hướng phát triển 15](#_Toc517096292)

[Tài liệu tham khảo 16](#_Toc517096293)

# 1. Tìm hiểu chung

A close up of a camera

Description generated with high confidenceĐồ án tìm hiểu cách truyền dữ liệu từ 1 chiếc camera IP trực tiếp về server, phục vụ cho quá trình nhận diện bằng AI. Camera dùng trong đồ án do hãng Yoosee sản xuất, tài liệu mô tả của hãng tại file: <http://support.hkvstar.com/file/Yoosee_Instruction_Manual.pdf>.

Hình 1. Camrera Yoosee

Theo tài liệu mô tả, camera dùng chuẩn ONVIF 2.4 Profile S, dùng giao thức RTSP để truyền luồng hình ảnh, âm thanh đi từ camera. Ta cần viết chương trình trêm một chiếc máy tính có thể giao tiếp với camera. Lúc này, máy tính đóng vai trò là client, gửi các yêu cầu đến server camera, camera sẽ gửi luồng hình ảnh, âm thanh lại cho máy tính bằng giao thức RTSP. Nhận được gói tin RTSP, chương trình trên máy tính lấy dữ liệu hình ảnh, sau đó chuyển cho mô-đun AI. Để kết thúc quá trình truyền, client gửi bản tin kết thúc đến camera, cả 2 đóng kết nối.

Trong khuôn khổ đồ án, em tìm hiểu 3 cách để truyền dữ liệu từ camera về. Cách thứ nhất, tạo trực các bản tin RTSP, đóng gói vào TCP, dùng các hàm có sẵn của hệ điều hành, gửi đến cho camera . Cách thứ hai, dùng thư viện OpenCV. Cách thứ ba, dùng phần mềm Red5. Mỗi cách, em sẽ trình bày trong từng chương tiếp theo.

## Các phần mềm thực tế

### Yoosee app

Hãng Yoosee phát triển phần mềm cho các thiết bị di động thông minh giúp người dùng sử dụng, điều khiển, theo dõi camera một cách dễ dàng nhất. Ứng dụng hỗ trợ cho cả iOS và Android.

Hình 2. Giao diện phần mềm Yoosee

A screenshot of a cell phone

Description generated with high confidenceA picture containing electronics

Description generated with high confidence

Để cấu hình cho camera, đầu tiên, cắm nguồn cho camera. Đối với những camera dùng sóng wifi, cần đặt ở nơi có sóng wifi mạnh, ổn định. Đối với những camera có cổng Ethernet, cần nối dây mạng.

Trên điện thoại, tải app Yoosee. Mở phần mềm lên, chọn Smartlink. Phần mềm sẽ dò tìm camera. Khi dò thành công, nhập mật khẩu wifi để giúp camera dùng cho quá trình hoạt động, đặt tài khoản và mật khẩu…. Nếu không thiết lập được, giữ nút reset dưới đáy camera, và thực hiện lại.

### CMS Client

A screenshot of a cell phone

Description generated with very high confidenceĐối với các máy tính Windows, có thể dùng phần mềm CMS Client.

Hình 3. Phần mềm CMS Client

Hướng dẫn sử dụng cả 2 phần mềm đều được hãng Yoosee viết chi tiết tại file: : <http://support.hkvstar.com/file/Yoosee_Instruction_Manual.pdf>.

## 1.2. Chuẩn ONVIF

Với sự bùng nổ mạng Internet, camera IP là một thiết bị tỏ ra hiệu quả, đơn giản , ít tốn kém khi chạy trên nền Internet. Người sử dụng có thể dùng camera để giám sát bất kỳ đâu, bất kỳ lúc nào, miễn là có mạng internet ổn định. Đã có rất nhiều hãng sản xuất phần cứng tham gia sản xuất camera IP. Tuy nhiên các loại camera IP của các hãng khác nhau lại dùng cách truyền dữ liệu khác nhau. Gây khó khăn trong việc thiết lập một hệ thống camera quan sát.   
 Để giải quyết vấn đề này, 27/08/2013 chuẩn ONVIF đã được công bố. Với nghiên cứu này, sẽ tích hợp hệ thống video và kiểm soát truy cập dựa vào IP từ một loạt video khác nhau và kiểm soát luôn vấn đề truy cập. Khả năng tương thích giữa các thiết bị camera có chuẩn ONVIF trên camera quan sát IP sẽ giúp người dùng đơn giản hóa việc cài đặt và quản lý trên các đầu ghi hình IP camera khác nhau. Khi sử dụng chuẩn ONVIF có thể cung cấp thông tin về các điểm truy cập trong hệ thống lắp đặt camera quan sát, ngoài ra nó còn kiểm soát truy cập và báo động. Người dùng có thể truy cập các thiết bị an ninh cũng như kiểm soát video và âm thanh nhờ vào chuẩn ONVIF trên camera quan sát IP này.

ONVIF (Open Network Video Interface Forum) là một tiêu chuẩn để các sản phẩm IP trong việc giám sát hình ảnh và các vùng giám sát an ninh khác có thể giao tiếp với nhau. ONVIF một tổ chức thành lập năm 2008 bởi các thương hiệu lớn như Axis, Bosch, Sony. Là một tổ chức phi lợi nhuận, các thành viên thuộc diễn đàn ONVIF lập ra chuẩn này để các nhà sản xuất, nhà phát triển phần mềm, nhà tư vấn, tích hợp hệ thống, người dùng và các nhóm có quan tâm hướng đến 1 chuẩn giao thức IP nhất định.

Nhiệm vụ của ONVIF là giúp cho việc tích hợp giữa các thiết bị giám sát trở nên dễ dàng bằng cách sử dụng một tiêu chuẩn mở toàn cầu: tiêu chuẩn kết nối thiết bị giám sát trên nền lắp đặt camera quan sát IP giá rẻ.Tiêu chuẩn hóa giao thức giữa các thiết bị thu phát hình ảnh qua mạng. Làm cho tương thích giữa các sản phẩm thu phát hình ảnh qua mạng mà không phụ thuộc vào nhà sản xuất. Được sử dụng cho tất cả các công ty và tổ chức có quy mô đa quốc gia. Như vậy mục tiêu chính của ONVIF là định hướng các thiết bị an ninh hoạt động trên nền giao thức IP có thể giao tiếp với nhau thông qua các tiêu chuẩn có tính chất quốc tế. Hiểu một cách đơn giản, các hãng camera IP khác nhau (Camera IP Avtech, Camera IP Vantech, camera IP Questek...) hoạt động theo tiêu chuẩn ONVIF có thể kết nối với nhau trên cùng phần mềm quản lý giám sát hình ảnh tập trung.

## 1.3. Giao thức RTSP

Real Time Streaming Protocol (RTSP) – Giao thức truyền tin thời gian thực là một giao thức điều khiển truyền thông mạng ở tầng ứng dụng được thiết kế để sử dụng trong các hệ thống giải trí và truyền thông để điều khiển máy chủ chứa các dữ liệu truyền tin đa phương tiện (streaming media). Giao thức này được sử dụng để thiết lập và điều khiển các phiên truyền thông giữa các trạm cuối. Các máy khách của các máy chủ truyền thông ban ra các lệnh kiểu VCR, chẳng hạn như chơi, thâu và tạm dừng, để điều khiển thời gian thực của các phương tiện truyền tin trực tuyến từ máy chủ tới máy khách (Video On Demand) hoặc từ máy khách đến máy chủ (Voice Recording) .

Việc truyền tải dữ liệu trực tuyến không phải là một nhiệm vụ của RTSP. Hầu hết các máy chủ RTSP sử dụng giao thức truyền tải thời gian thực (RTP) kết hợp với giao thức điều khiển thời gian thực (RTCP) để phân phối luồng truyền thông. Tuy nhiên, một số nhà cung cấp thực hiện các giao thức vận tải độc quyền. Phần mềm máy chủ RTSP của RealNetworks, thí dụ, cũng sử dụng Real Data Transport (RDT) độc quyền của RealNetworks.

RTSP hoạt động theo mô hình Client-Server. Camera đóng vai trò là Server, luôn chờ các request yêu cầu Streaming, khi có thì gửi lại các response trả lời chứa hình ảnh, âm thanh. Chương trình muốn lấy dữ liệu đóng vai trò là client. Các request của RTSP khá giống HTTP, gồm các phương thức (OPTIONS, DESCRIBE, SETUP, PLAY, PAUSE, RECORD,…), các mã trả về (200 OK,…), các trường header….

RSTP có các Control Sequence (Cseq), mỗi request sẽ được đánh số, response tương ứng có Cseq giống request. RTSP là giao thức statefull. Một phiên Streaming sẽ cần nhiều request-response thiết lập , điều khiển. RTSP dùng TCP để điều khiển, dùng UDP để truyền dữ liệu. Nó dùng cổng 554 cho cả TCP và UDP.

Ví dụ một vài RTSP Request:

**OPTIONS**

C->S: OPTIONS rtsp://example.com/media.mp4 RTSP/1.0

CSeq: 1

Require: implicit-play

Proxy-Require: gzipped-messages

S->C: RTSP/1.0 200 OK

CSeq: 1

Public: DESCRIBE, SETUP, TEARDOWN, PLAY, PAUSE

OPTIONS request trả về các loại request mà server hỗ trợ. Ở đây, server hỗ trợ DESCRIBE, SETUP, TEARDOWN, PLAY và PAUSE.

Dòng đầu tiên của request gồm: tên lệnh (OPTIONS), tiếp theo là đường dẫn (rtsp:// example.com/media.mp4), cuối cùng là phiên bản (RTSP/1.0). Dòng thứ hai, Control Squence. Các dòng tiếp theo, mỗi dòng là một cặp trường giá trị khác, bắt đầu là tên trường, tiếp theo là dấu 2 chấm, theo sau là giá trị của trường.

Trong response, dòng đầu tiên gồm phiên bản (RTSP/1.0), tiếp theo là mã và tên trạng thái (200 OK, nghĩa là bình thường, không có lỗi). Dòng tiếp theo là Control Sequence, Cseq có giá trị 1, nghĩa là nó ứng với request có Cseq cũng là 1.

**DESCRIBE**

C->S: DESCRIBE rtsp://example.com/media.mp4 RTSP/1.0

CSeq: 2

S->C: RTSP/1.0 200 OK

CSeq: 2

Content-Base: rtsp://example.com/media.mp4

Content-Type: application/sdp

Content-Length: 460

m=video 0 RTP/AVP 96

a=control:streamid=0

a=range:npt=0-7.741000

a=length:npt=7.741000

a=rtpmap:96 MP4V-ES/5544

a=mimetype:string;"video/MP4V-ES"

a=AvgBitRate:integer;304018

a=StreamName:string;"hinted video track"

m=audio 0 RTP/AVP 97

a=control:streamid=1

a=range:npt=0-7.712000

a=length:npt=7.712000

a=rtpmap:97 mpeg4-generic/32000/2

a=mimetype:string;"audio/mpeg4-generic"

a=AvgBitRate:integer;65790

a=StreamName:string;"hinted audio track"

DESCRIBE trả về mô tả chi tiết về server, theo định dàng SDP (Session Description Protocol).

**SETUP**

C->S: SETUP rtsp://example.com/media.mp4/streamid=0 RTSP/1.0

CSeq: 3

Transport: RTP/AVP;unicast;client\_port=8000-8001

S->C: RTSP/1.0 200 OK

CSeq: 3

Transport: RTP/AVP;unicast;client\_port=8000-8001;server\_port=9000-9001;ssrc=1234ABCD

Session: 12345678

SETUP request chỉ định cách truyền một luồng phương tiện. Điều này phải được thực hiện trước khi gửi PLAY request. Request chứa URL luồng và cách thức truyền tải. Thông số này thường bao gồm một cổng để nhận dữ liệu RTP (âm thanh hoặc video) và một cổng khác cho dữ liệu RTCP (thông tin meta). Response thường xác nhận các tham số đã chọn và điền vào các phần bị thiếu, chẳng hạn như các cổng đã chọn của máy chủ. Mỗi luồng phải được định cấu hình bằng SETUP trước khi có thể gửi yêu cầu PLAY.

**PLAY**

C->S: PLAY rtsp://example.com/media.mp4 RTSP/1.0

CSeq: 4

Range: npt=5-20

Session: 12345678

S->C: RTSP/1.0 200 OK

CSeq: 4

Session: 12345678

RTP-Info: url=rtsp://example.com/media.mp4/streamid=0;seq=9810092;rtptime=3450012

PLAY request giúp bắt đầu quá trình truyền. Server sẽ truyền luồng cho client như đã quy định khi SETUP.

**PAUSE**

C->S: PAUSE rtsp://example.com/media.mp4 RTSP/1.0

CSeq: 5

Session: 12345678

S->C: RTSP/1.0 200 OK

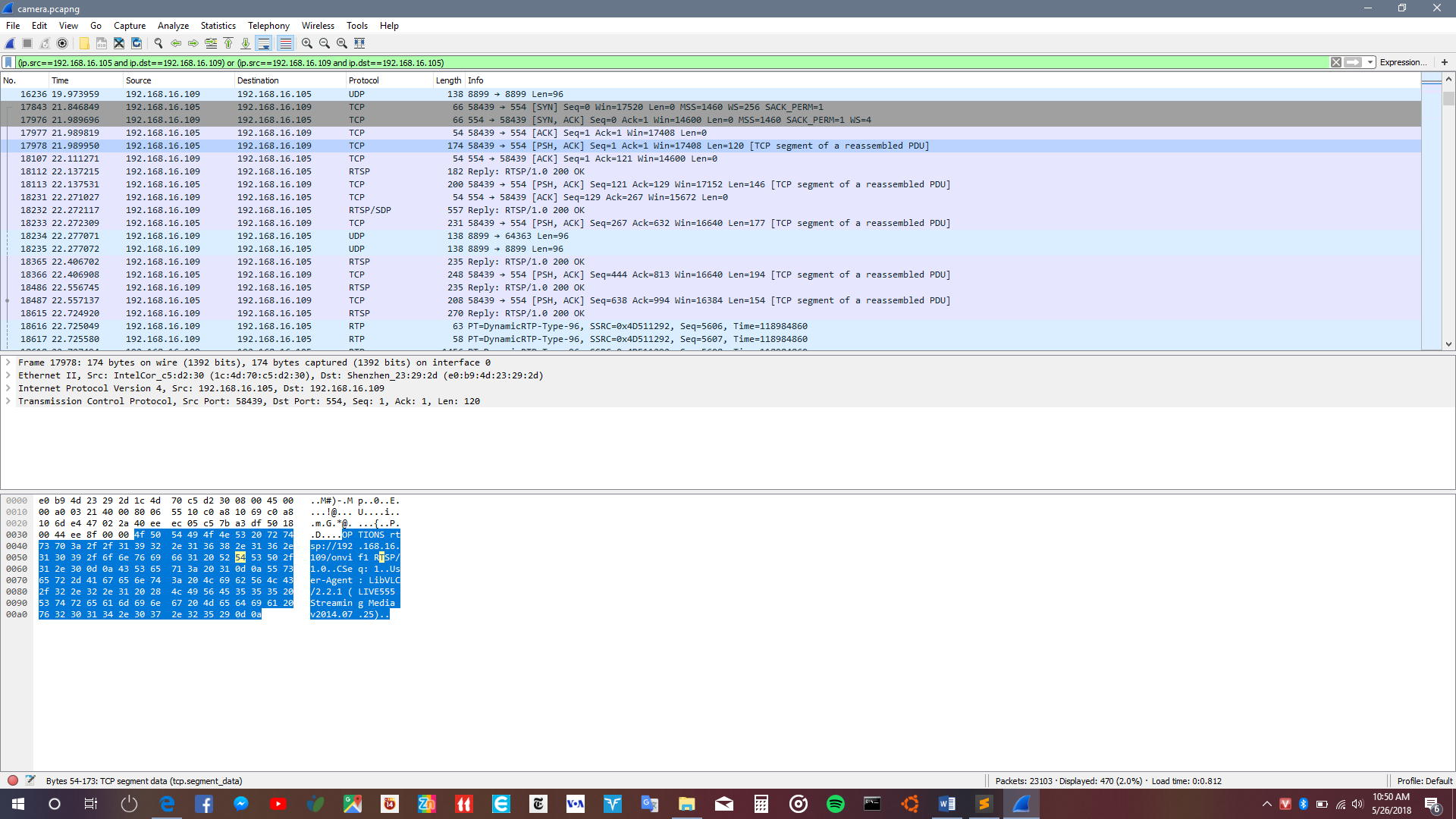
CSeq: 5

Session: 12345678

PAUSE giúp tạm thời dừng một hoặc tất cả các luồng đang phát. Có thể tiếp tục lặi bằng PLAY.

# 2. Dùng TCP, UDP tạo gói tin RTSP

Ý tưởng khá đơn giản, dùng cú pháp request thiết lập của RTSP, đóng gói vào bản tin TCP. Gửi cho camera. Sau đó, mở kết nối UDP, nhận gói tin RTSP chứa dữ liệu từ camera. Bóc tách, lấy được dữ liệu hình ảnh.  
 Để cho đơn giản, em dùng phần mềm CMS Client (đã cài ở phần 1), mở kết nối đến camera. Sau đó dùng WireShark, bắt và phân tích các gói tin trao đổi giữa CMS và camera IP. Tiếp theo, viết chương trình “nhái” lại các gói tin này và trao đổi với camera.



Hình 4. Dùng WireShark bắt các gói tin khi chạy CMS client

Hình trên là gói tin OPTIONS của CMS Client gửi đến Camera. Khi lập trình, chỉ cần copy nội dung gói tin, đưa vào xâu buffer, dùng TCP gửi buffer đó đến camera bằng thư viện sys/socket. Bằng cách này, chỉ cần áp dụng đúng thứ tự, nghe đúng cổng, thì chương trình hoàn toàn có thể trao đổi dữ liệu bình thường với Camera. Cách này đơn giản, thực hiện dễ dàng, nhanh chóng. Tuy nhiên, cách này khá cứng nhắc, khi thay đổi camera, lại phải làm lại.

A screenshot of a cell phone

Description generated with high confidence

Đầu tiên, mở kết nối TCP, kết nối này dành cho quá trình thiết lập. Sau đó, gửi lần lượt các phương thức OPTIONS, DESCRIBE, SETUP. Tiếp theo đó, mở kết nối UDP, tất cả luồng hình ảnh, âm thanh từ camera sẽ truyền qua kết nối này. Gửi lệnh PLAY, camera sẽ liên tục gửi dữ liệu qua kết nối UDP vừa tạo. Khi muốn tạm dừng / kết thúc quá trình nhận, gửi thông điệp PAUSE cho camera.

# 3. Dùng thư viện OpenCV

## 3.1. Sơ lược về OpenCV

OpenCV là một thư viện mã nguồn mở hàng đầu cho thị giác máy tính (computer vision), xử lý ảnh và máy học, và các tính năng tăng tốc GPU trong hoạt động thời gian thực.

OpenCV được phát hành theo giấy phép BSD, do đó nó hoàn toàn miễn phí cho cả học thuật và thương mại. Nó có các interface C++, C, Python, Java và hỗ trợ Windows, Linux, Mac OS, iOS và Android. OpenCV được thiết kế để tính toán hiệu quả và với sự tập trung nhiều vào các ứng dụng thời gian thực. Được viết bằng tối ưu hóa C/C++, thư viện có thể tận dụng lợi thế của xử lý đa lõi. Được sử dụng trên khắp thế giới, OpenCV có cộng đồng hơn 47 nghìn người dùng và số lượng download vượt quá 6 triệu lần. Phạm vi sử dụng từ nghệ thuật tương tác, cho đến lĩnh vực khai thác mỏ, bản đồ trên web hoặc công nghệ robot.

OpenCV có class VideoCapture, class này giúp bắt luồng video theo từng frame một, hỗ trợ video lưu trên máy, lẫn cả video streaming.

## 3.2. Cài đặt OpenCV

Dưới đây, em xin trình bày cách cài OpenCV trên Ubuntu 16.10. Trên một vài phiên bản Ubuntu mới hơn, do Ubuntu thay đổi chế độ đồ họa, nên một số gói có thể bị lỗi.

Bước 1: Updating Ubuntu

$ sudo apt-get update

$ sudo apt-get upgrade

Bước 2: Cài đặt các gói

$ sudo apt-get install build-essential cmake git libgtk2.0-dev pkg-config libavcodec-dev libavformat-dev libswscale-dev

$ sudo apt-get install python3.5-dev python3-numpy libtbb2 libtbb-dev

$ sudo apt-get install libjpeg-dev libpng-dev libtiff5-dev libjasper-dev libdc1394-22-dev libeigen3-dev libtheora-dev libvorbis-dev libxvidcore-dev libx264-dev sphinx-common libtbb-dev yasm libfaac-dev libopencore-amrnb-dev libopencore-amrwb-dev libopenexr-dev libgstreamer-plugins-base1.0-dev libavutil-dev libavfilter-dev libavresample-dev

Bước 3: Clone OpenCV từ github

$ sudo -s

$ cd /opt

/opt$ git clone https://github.com/Itseez/opencv.git

/opt$ git clone <https://github.com/Itseez/opencv_contrib.git>

Bước 4: Build và cài đặt OpenCV

/opt$ cd opencv

/opt/opencv$ mkdir release

/opt/opencv$ cd release

/opt/opencv/release$ cmake -D BUILD\_TIFF=ON -D WITH\_CUDA=OFF -D ENABLE\_AVX=OFF -D WITH\_OPENGL=OFF -D WITH\_OPENCL=OFF -D WITH\_IPP=OFF -D WITH\_TBB=ON -D BUILD\_TBB=ON -D WITH\_EIGEN=OFF -D WITH\_V4L=OFF -D WITH\_VTK=OFF -D BUILD\_TESTS=OFF -D BUILD\_PERF\_TESTS=OFF -D CMAKE\_BUILD\_TYPE=RELEASE -D CMAKE\_INSTALL\_PREFIX=/usr/local -D OPENCV\_EXTRA\_MODULES\_PATH=/opt/opencv\_contrib/modules /opt/opencv/

/opt/opencv/release$ make -j4

/opt/opencv/release$ make install

/opt/opencv/release$ ldconfig

/opt/opencv/release$ exit

/opt/opencv/release$ cd ~

## 3.2. Sơ đồ khối

Mô tả:

* Đưa ra 3 lựa chọn định dạng đầu ra: video, image, pipe. Lựa chọn đầu tiên, lưu lại thành 1 đoạn video trong máy. Lựa chọn thứ 2, lưu thành 1 ảnh, ảnh này liên tục thay đổi theo các frame video truyền về. Lựa chọn thứ 3, lưu thành FIFO file, tạo thành pipeline trao đổi thông tin giữa 2 tiến trình.
* Dùng VideoCapture, mở kết nối RTSP, kiểm tra kết nối, tìm định dạng mã hóa, kích thước frame. Liên tục nhận luồng stream từ camera về, cắt thành từng frame một, lưu lại.

Chương trình nhận tham số từ dòng lệnh, dạng:

./Capture -v video.mp4 -i image.jpg -p pipe rtsp://192.168.1.15:554/onvif1

Trong đó, Capture là tên chương trình. -v -i -p là các tùy chọn định dạng đầu ra (video, image, pipeline), sau mỗi tùy chọn là tên file cần lưu. Tham số cuối cùng là địa chỉ stream của camera.

Chương trình kiểm tra các tham số này, đánh dấu bởi 3 biến flag tương ứng: videoFlag, imageFlag, pipeFlag. Biến flag có giá trị true khi tùy chọn tương ứng được bật, false trong trường hợp ngược lại.

Sơ đồ khối:

A close up of a logo

Description generated with very high confidence

# 4. Hướng phát triển

Tiếp tục cải tiến, nâng cấp thêm cho phiên bản dùng OpenCV. Bổ sung giao diện đồ họa để người quản trị dễ dàng truy cập, điều khiển.

Phát triển thêm phiên bản dùng Red5. Red5 là phần mềm stream mã nguồn mở nổi tiếng. Nó có thể stream cùng lúc đến nhiều thiết bị với hiệu năng cao. Do tài liệu về red5 không nhiều, việc cài đặt không thuận lợi, kiến thức có hạn nên em chưa thể triển khai được. Trong tương lai, em sẽ cố gắng cài đặt, tùy chỉnh red5 thành 1 server stream.

Chuẩn ONVIF có hỗ trợ tìm thiết bị trong cùng mạng LAN, không cần phải nhập địa chỉ IP thủ công nữa. Chức năng này là rất cần thiết, bởi không phải lúc nào cũng dễ dàng biết địa chỉ IP của camera. Hơn nữa khi gặp sự cố, địa chỉ IP thay đổi, phần mềm có thể tự động dò tìm lại camera thì tiết kiệm thời gian hơn rất nhiều so với người quản trị phải dò lại IP một cách thủ công.

# Tài liệu tham khảo

[1] Chuẩn ONVIF là gì? Tại sao phải sử dụng chuẩn ONVIF trong camera IP?

<http://www.cameravinhphat.com/tin-tuc/chuan-onvif-la-gi-tai-sao-phai-su-dung-chuan-onvif-trong-camera-ip-337.html>

[2] Real Time Streaming Protocol

<https://wikipedia.org/wiki/Real_Time_Streaming_Protocol>

[3] Hồ Sỹ Hùng, OpenCV và các ứng dụng của nó hiện nay

<https://techmaster.vn/posts/33943/opencv-va-cac-ung-dung-cua-no-hien-nay>

[4] How to Install OpenCV in Ubuntu 16.04 LTS for C / C++

<http://www.codebind.com/cpp-tutorial/install-opencv-ubuntu-cpp/>