

ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC BÁCH KHOA

KHOA ĐIỆN – ĐIỆN TỬ

~~- ---------------------------------------------------------~~

BÁO CÁO BÀI TẬP LỚN:

**THỊ GIÁC MÁY**

Đề Tài:ỨNG DỤNG COMPUTER VISION VÀO ĐO NHỊP TIM

GIẢNG VIÊN:

PHẠM VIỆT CƯỜNG

TP. HỒ CHÍ MINH, THÁNG 12 NĂM 2019



ĐẠI HỌC QUỐC GIA TP.HỒ CHÍ MINH

TRƯỜNG ĐẠI HỌC BÁCH KHOA

KHOA ĐIỆN – ĐIỆN TỬ

|  |  |
| --- | --- |
| **NHÓM 7** | |
| **Họ và tên** | **MSSV** |
| VÕ THỊ CẨM VIÊN | 1733636 |
| PHẠM QUANG KIỆT | 1511662 |
| BÙI MẠNH HUY | 1411409 |
| VŨ ĐĂNG KHOA | 1611645 |
| NGUYỄN HOÀNG KHỞI | 1611663 |

**LỜI CẢM ƠN**

Lời đầu tiên xin gửi lời cảm ơn chân thành đến thầy Phạm Việt Cường đã tận tình giảng dậy bọn em trong suốt học kỳ vừa qua.

Qua những bài giảng của thầy trong suốt quá trình dạy và những bài thiết trình bọn em đã học được một số nền tảng cơ bản về computer vision, lối tư duy hệ thống để tìm hiểu một thuật toán trước khi bắt đầu tìm hiểu sau về nó cũng như phải hệ thống những gì sẽ thuyết trình để các bạn có thể mô hình hoá bài thuyết trình trước khi bắt đầu thuyết trình về một thuật toán cụ thể.

Và qua bày báo cáo cuối khoá này, nhóm em sẽ thể hiện những gì đã học được từ thầy một cách có đầu tư. Sẽ có những sai sốt mong thầy bỏ qua và có những chỉ bảo cần thiết để bọn em có thể tiến xa hơn sau này.

Cuối cùng đại diện cho nhóm xin chúc sức khoẻ đến thầy, mong thầy luôn giữ vững được sự nhiệt huyết và sự khao khát truyền đạt kiến thức đối những thế hệ tiếp theo và sau nữa.

Tp. Hồ Chí Minh, ngày 2 tháng 12 năm 2019

MỤC LỤC

[DANH SÁCH TỪ VIẾT TẮT 4](#_Toc26200460)

[DOANH SÁCH HÌNH MINH HOẠ 5](#_Toc26200461)

[Chương 1. TÓM TẮT VỀ ĐỀ TÀI 7](#_Toc26200462)

[Chương 2. PHƯƠNG PHÁP THỰC HIỆN 9](#_Toc26200463)

[2.1 Nguyên lý nền tản 9](#_Toc26200464)

[2.2 Sơ đồ khối thuật toán 10](#_Toc26200465)

[2.3 Hoạt động của từng frame (khung) trong thuật toán. 12](#_Toc26200466)

[Chương 3. KẾT QUẢ THỰC HIỆN 18](#_Toc26200467)

[Chương 4. NHẬN XÉT VÀ HƯỚNG PHÁT TRIỂN 19](#_Toc26200468)

[TÀI LIỆU THAM KHẢO 20](#_Toc26200469)

# DANH SÁCH TỪ VIẾT TẮT

|  |  |  |
| --- | --- | --- |
| **Doanh sách từ** | **Ký hiệu** | **Ý nghĩa** |
| region of interest | ROI |  |
| Fast Fourier Transform | FFT |  |
| Recurring Point of Interesting | RPoI |  |
|  | PPG |  |
| Infinite Impulse Response | IIR |  |

# DOANH SÁCH HÌNH MINH HOẠ

Hình n: dạng sóng của quá trình tim bơm máu trong một chu kỳ

Hình n: dạng sóng của quá trình tim bơm máu trong một chu kỳ

Hình n: Sơ đồ khối thuật toán

Hình n: (a) ảnh được trích xuất từ webcam. (b) ảnh sau khi được chiếc suất trên 3 khung màu riêng biệt Reb,Green,Blue. (c) biểu diễn tín hiệu bằng đồ thị.

Hình n: (a) ảnh được trích xuất từ webcam. (b) ảnh sau khi được chiếc suất trên 3 khung màu riêng biệt Reb,Green,Blue. (c) biểu diễn tín hiệu bằng đồ thị.

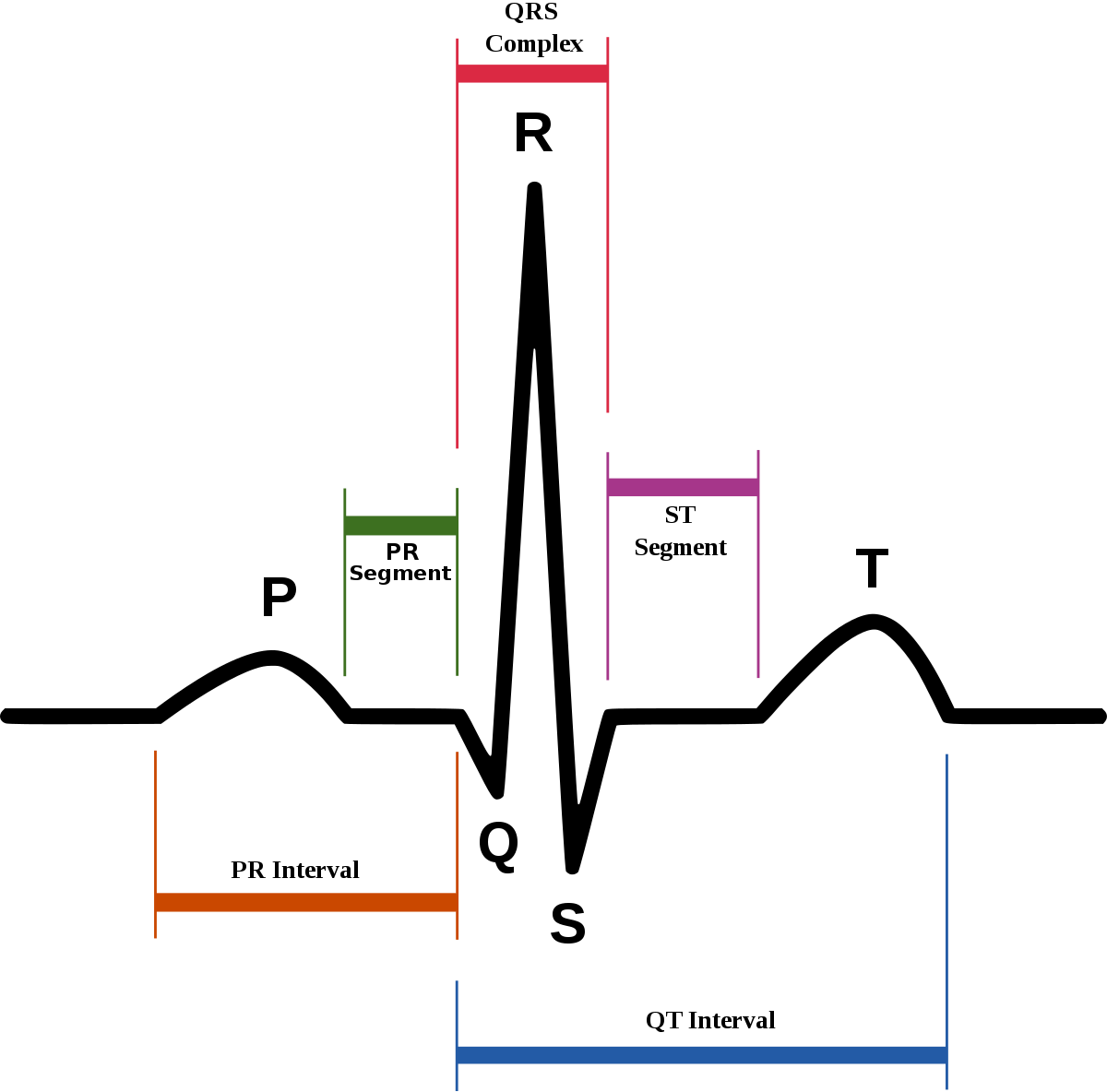
Hình n: ảnh minh hoạ so sánh giá trị thu được của việc sử dụng webcam so với BVP

Hình n: Ảnh minh hoạ cho thuật toán findpeaks

# Chương 1. TÓM TẮT VỀ ĐỀ TÀI

Đề tài “ Ứng dụng computer vision để đo nhịp tim”

Áp dụng những kiến thức đã học về môn điện tử y sinh, thì trong một chu kỳ tuỳ thuộc vào cơ địa, độ tuổi, tình trạng thể lực, hoạt động thể chất mà ứng với mỗi người sẽ có nhịp tim khác nhau. Nhưng xét trong một chu kỳ quá trình bơm máu (đập) của tim gần như không đổi được thể hiện qua dạng sóng:



Hình n: dạng sóng của quá trình tim bơm máu trong một chu kỳ

Do đó nếu ta áp dụng một số công cụ xử lý mạnh mẽ như computer vision, xử lý số tín hiệu thì có thể bắt được sự thây đổi này và hiển thị ra đồ thị nhịp tim.

Để làm được điều đó thuật toán của nhóm em gồm 6 bước:

Bước 1: Lấy ảnh từ webcam để lưu lại xử lý và hiển thị ra màn hình

Bước 2: Sử dụng thuật toán facial landmark để chọn những vùng thích hợp nhất (ROI) của ảnh

Bước 3: Trích xuất ảnh với ba khung màu và chọn khung màu thích hợp

Bước 4: Sau khi đã chọn được khung màu thích hợp chuyển ảnh qua miền tần số với phương pháp FFT

Bước 5: Sử dụng bộ lọc thông giải để lọc bỏ những tần số nằm ngoài giải tần do nhiễu hoặc những phản ứng trừu tượng trong cơ thể.

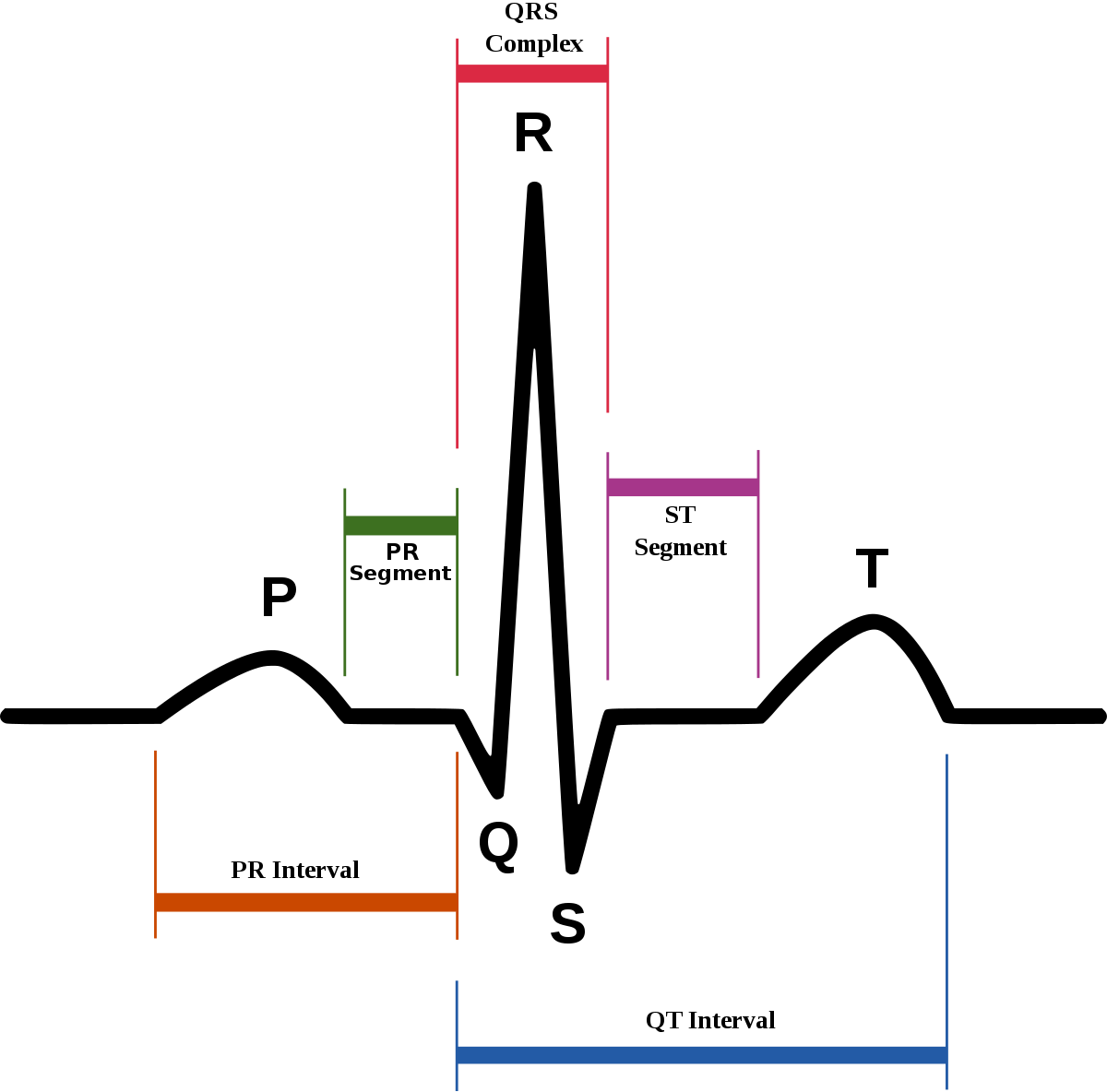
Bước 6: Sử dụng thuật toán tìm đỉnh để chọn ra những điểm có cường độ lớn nhất có khả năng là Recurring Point of Interesting ( những điểm có khả năng chứa thông tin về nhịp tim).

Bước 7: Phân tích các RpoI và hiển thị nhịp tim lên màn hình.

# Chương 2. PHƯƠNG PHÁP THỰC HIỆN

## 2.1 Nguyên lý nền tản

Theo như những gì đã học trong môn điện tử y sinh thì một chu kỳ hoạt động của tim được thể hiện qua dạng sóng



Hình n: dạng sóng trong một chu kỳ bơm máu của tim.

Quá trình bơm máu của tim có thể chia thành 3 giai đoạn ứng với 3 dạng sóng:

Giai đoạn 1: Sóng P(khử phân cực tâm nhỉ)

Giai đoạn 2: sóng phức hợp QRS( khử phân cực tâm thất)

Giai đoạn 3: sóng T(tái phân cực tâm thất)

Mỗi dạng sóng ứng với cường độ máu được tim bơm trên cơ thể, ta có thể thấy sóng phức hợp QRS có cường độ lớn nhất. Đây là nguyên lý cơ bản xuyên suốt đề cương lần này của nhóm.

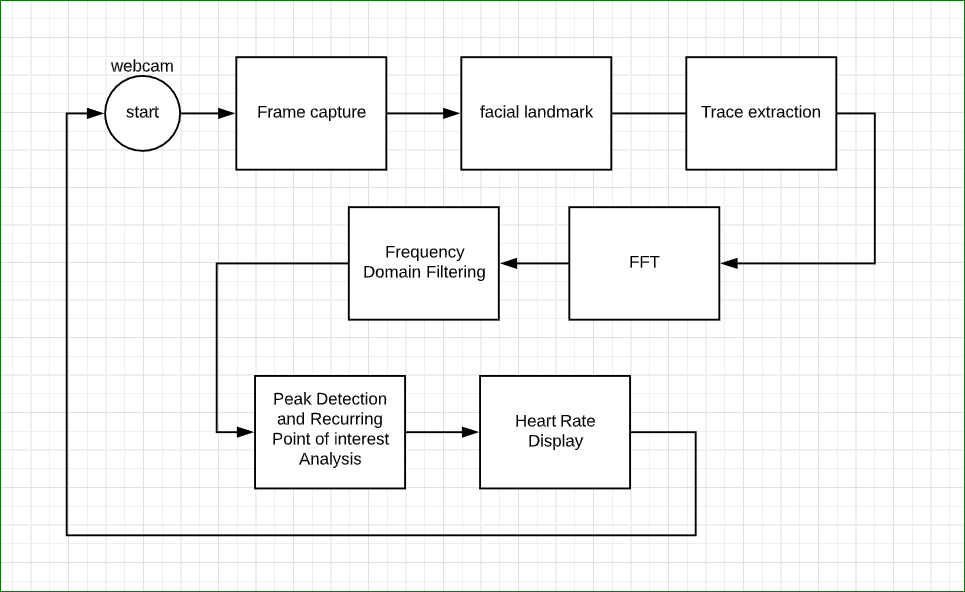
Thêm vào lý thuyết về phản xạ ánh sáng “hiện tượng phản xạ ánh sáng xảy ra khi ta chiếu một nguồn ánh sáng từ môi trường chiếc quang sang môi trường kém chiếc quang hơn (n1>n2)”. Tương ứng với nồng độ máu tại mỗi thời điểm thì chiếc suất n2 sẽ thay đổi tương ứng, dẫn đến ánh sáng phản xạ cũng sẽ thây đổi theo. Việc còn lại là tìm ra một công cụ để thu lại sự thây đổi này.

Theo như nghiên cứu thì tuỳ thuộc vào tần suất hoạt động, vị trí, điều kiện thời tiết, tình trạng sức khoẻ,.. thì mỗi vị trí khác nhau có nồng độ khác nhau. Ví dụ trong khi nói chuyện thì trán, mũi, miệng, tai,má... hoạt đồng nhiều hơn một số bộ phận khác như chân, tay, đùi,cổ tay... hay do vị trí mà sự thây đổi nồng độ máu ở tay và trán có sự khác biệt về thời gian.

Vì vậy với ứng dụng này nhóm em sử dụng một thuật toán nhận dạng cao cấp hơn gọi là facial landmark.

Một số ứng dụng khác khá phổ biến trong đo nhịp tim sử dụng nguyên lý này như photoplethysmography (PPG), đo nhịp tim bằng ngón tay,…

## 2.2 Sơ đồ khối thuật toán

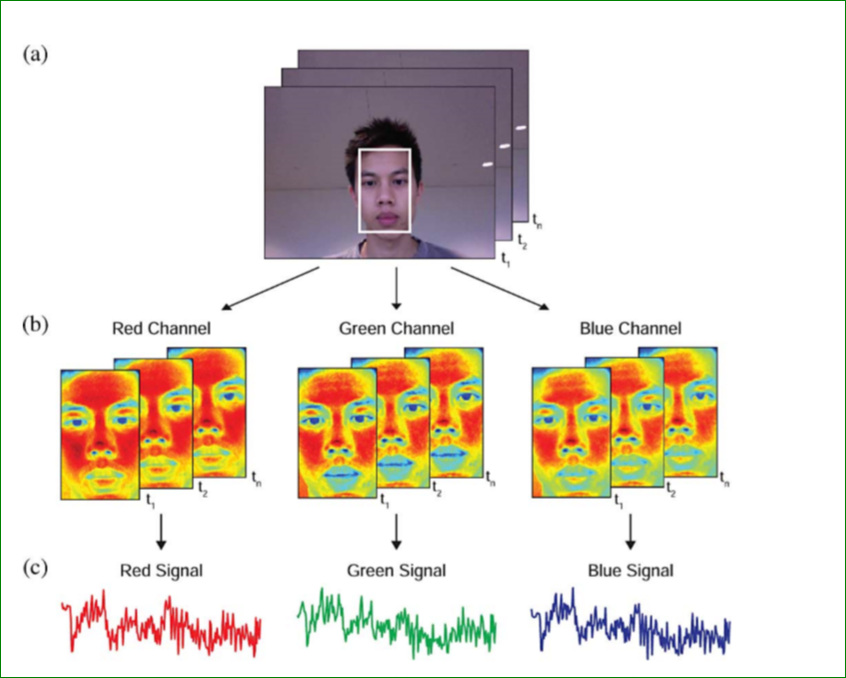


Hình n: Sơ đồ khối thuật toán

Nhịp tim con người 40-200bpm tuỳ thuộc vào độ tuổi, tình trạng thể lực, hoạt động thể chất. Để đảm bảo bắt được sự thây đổi tốt nhất nhóm em chọn băng thông từ 40-230bpm, tương ứng với tần số 3.833hz. Theo Nyquist thì tần số lấy mẫu phải lớn 2 lần tần số lớn nhất tức 8hz, với tốc độ chụp ảnh của webcam laptop (khoảng10 fps) thì tần số lấy mẫu này là hợp lý.

Ứng với mỗi vị trí khác nhau thì nồng độ máu là khác nhau phụ thuộc vào tần suất hoạt động, vị trí,.. Ở đây nhóm bọn em sử dụng thuật toán facial landmark để chọn 2 vùng có tần suất hoạt động lớn và vị trí gần nhau để phát hiện sự thây đổi đột ngột trong một chu kỳ bơm máu của tim là vùng trán và mũi.

Tiếp theo ta Trace extraction tức là chiếc xuất ảnh thành 3 khung màu red, green and blue. Ở bước này ta đã có 3 ảnh riêng biệt trên 3 khung màu khác nhau, công việc còn lại là kiểm tra sự thay đổi trên mỗi khung màu.



Hình n: (a) ảnh được trích xuất từ webcam. (b) ảnh sau khi được chiếc suất trên 3 khung màu riêng biệt Reb,Green,Blue. (c) biểu diễn tín hiệu bằng đồ thị.

Ta chọn ảnh với khung màu xanh để xử lý vì khi phân tích thì ảnh với khung màu này có đồ thị mượt nhất.

Sau khi đã chọn ra khung màu phù hợp ta chuyển chúng qua miền tần số với phương pháp Fast Fourier Transform (FFT) và cho ảnh qua bộ Frequency Domain Filtering( bộ lọc trên miền tần số), cụ thể hơn ở bài này em sử dụng bộ lọc thông giải Butterworth để giới hạn băng thông trong vùng 40-230 bpm nhằm lọc bớt nhiễu và loại bỏ những thành phần nhiễu.

Dữ liệu thu được sẽ tiếp tục được làm mượt để dễ dàng trong việc phát hiện đỉnh và tìm ra những điểm RPoI.

Ở bước cuối cùng này ta sẽ phân tích các điểm RPoI tìm được để tính toán giá trị nhịp tìm và vẽ đồ thị lên màn hình.

## 2.3 Hoạt động của từng frame (khung) trong thuật toán.

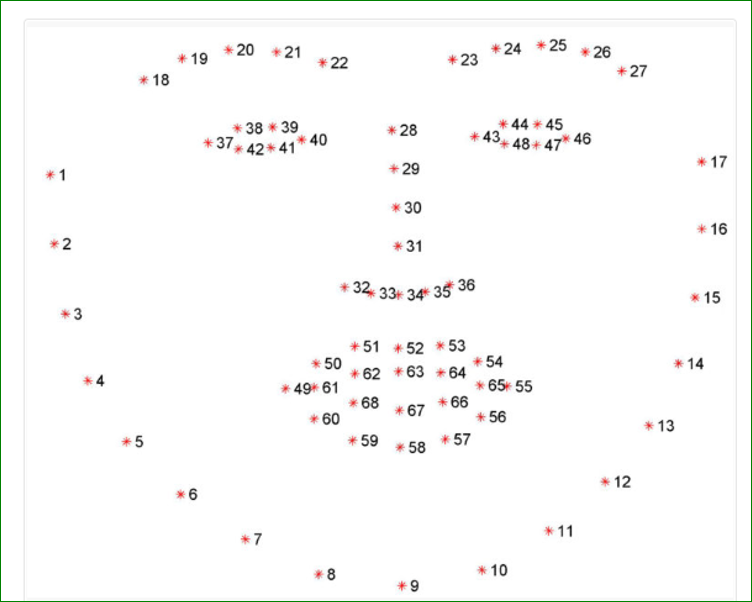
a. Frame capture

Ảnh từ webcam sẽ được lưu lại để xử lý và hiển thị ra màn hình.

Trong python-OpenCv để trích xuất ảnh từ webcam ta gọi hàm:

*webcam = cv2.VideoCapture(0;*

b. Facical lankmark



Hình n: ảnh minh hoạ cho thuật toán facial landmark

\* Định nghĩa

Facial landmarks là tên gọi chung dùng để chỉ những vùng quan trọng trong khuôn mặt như: mắt, lông mày, mũi, miệng,hàm. Vì vậy thuật toán này thì dùng để tìm các thành phần trong khuôn mặt như mắt, lông mày,mũi,miệng,hàm và chúng được đánh nhãn để dễ dạng trong việc trích xuất.

Đây là thuật toán cao cấp hơn so với nhận diện khuôn mặt bởi vì chính bản thân nó đã bao gồm nhận diện khuôn mặt.

\* Cách thức hoạt động của facial landmark.

Detecting facial lankmarks là một bài toán con của shape predictor(dự đoán hình dạng vật thể). Shape predictor là bài toán đưa ảnh đầu vào sau đó tìm ROI của đổi tượng cần dự đoán trong ảnh, thuật toán sẽ thực thi để tìm vị trí của những điểm quang trọng (points of interest) dọc theo hình dạng của vật thể.

Trong bài này facial lankmarks được sử dụng với mục đích chính là detect important facial structures (các chấm tròn được đánh với tronh hình n) trên khuôn mặt sử dụng kỹ thuật shape prediction.

Detecting facial landmarks thì gồm 2 bước:

Step #1: thu thập khuôn mặt trên ảnh.

Step #2: detect the key facial structures on the face ROI.

Face detection (step #1) có thể hoàn thành bằng nhiều cách:

sử dụng haar-casecade

Chúng ta có thể ứng dụng HOG

Hoặc chúng ta cũng có thể sử dụng thuật toán deep learning-based như CNN,...

Trong bài này nhóm em sử dụng HOG lấy ra từ thu viện dlib của python-Opencv:

*Detector= dlib.get\_frontal\_face\_detector()*

Sau khi tìm ra khuôn mặt chúng ta có thể tiến đến step #2: detect important facial structures trong khuôn mặt như:

+Trán

+ Mũi

+ Miệng

+ Mắt trái

+ Mắt phải

+ Mi trái

+ Mi phải

The facial landmark detector được Kazemi và Sullivan đưa ra trong bài báo One millisecond Face Alignment with an Ensemble of Regression Trees vào năm 2014

Phương pháp này bắt đầu bằng cách:

1. Một tập dữ liệu huấn luyện các nhãn facial landmarks trên ảnh. Những ảnh này được dán nhãn thủ công, chỉ định cụ thể trục (x,y) của các vùng lần cận trên mỗi facial structure.

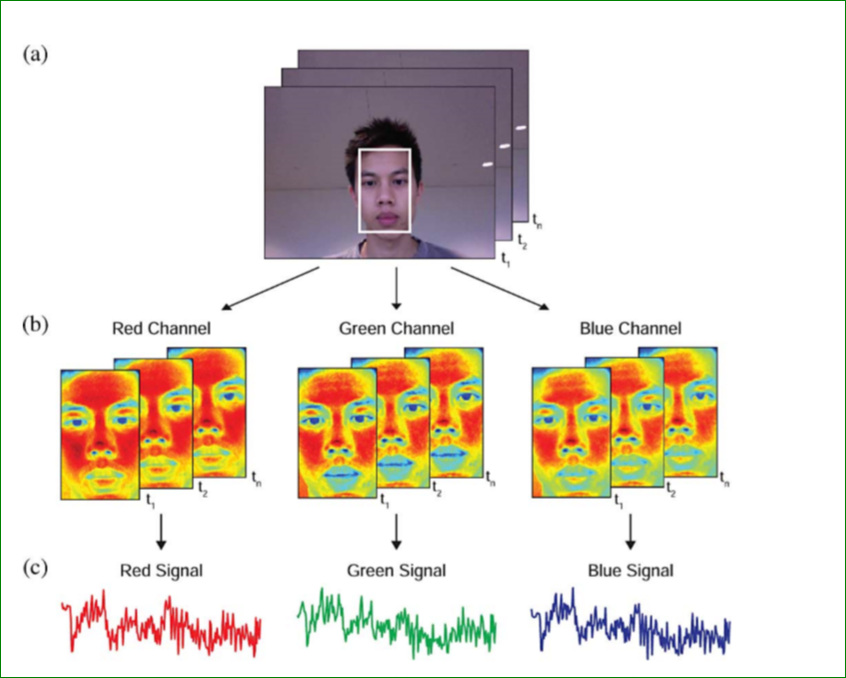
2. Cụ thể hơn, xác suất trên khoản cách giữa các cặp pixel đầu vào.

với dữ liệu huấn luyện này, một nhóm các cây hổi quy được đào tạo để estimate các vị trí mốc trên khuôn mặt trực tiếp từ chính cường độ điểm ảnh ( no “ feature extraction” đang điễn ra)

Kết quả cuối cùng của facial landmark detector là nó có thể được sử dụng để detect facial landmarks trong thời gian thực với các dự đoán chất lượng cao.

c. Trace extraction

Ảnh trong ROI sau khi đã được cắt sẽ đi qua bước tiếp theo là Trace extraction tức chia một ảnh màu RGB thành ba khung màu riêng biệt là red, green, blue để tìm sự thay đổi trên từng khung màu.



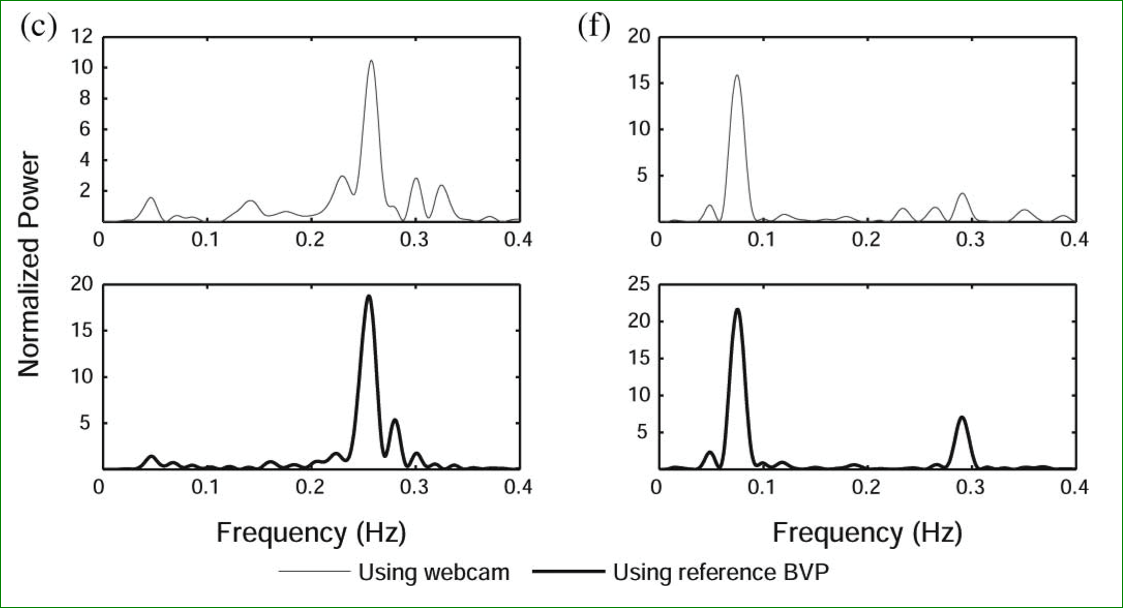
Hình n: (a) ảnh được trích xuất từ webcam. (b) ảnh sau khi được chiếc suất trên 3 khung màu riêng biệt Reb,Green,Blue. (c) biểu diễn tín hiệu bằng đồ thị.

// cần hình ảnh để thây thế (phải lấy thực tế từ trong code)

d. Fast Fourier Transform (FFT)

Với dữ liệu thu được ở hình n (c) ta sẽ chuyển qua miền tần số với phương pháp FFT để dễ trong việc tính toán.

// cần tìm một công thức hay hình ảnh thêm vào



Hình n: ảnh minh hoạ so sánh giá trị thu được của việc sử dụng webcam so với BVP

e. Frequency Domain Filtering.

Một bộ là cần thiết để giảm những thành phần không cần thiết do nhiễu hay các hoạt động trừu tượng trong cơ thể.

Nhóm em chọn bộ lọc thông dải lọc bỏ những tần số bên ngoài dải tần, Điều này giúp giảm các thành phần không cần thiết bên ngoài dải tần.

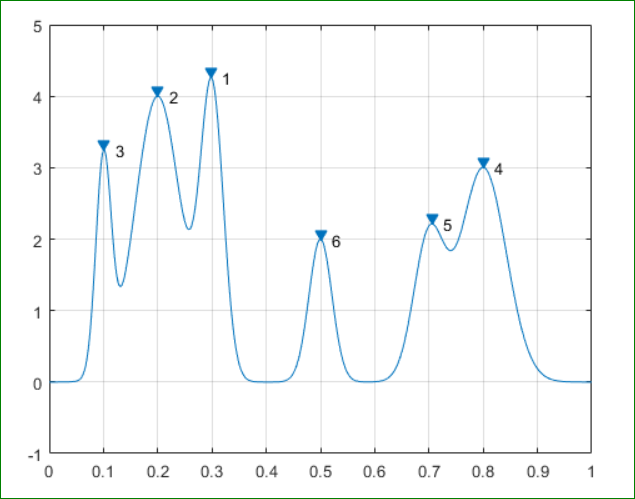
Đối với trường hợp này nhóm sử dụng bộ lọc butterworth bậc 2 được thiết kế trong giải tần 40-230hz. Nhóm chọn bộ lọc này thì theo như những bài báo bọn em đọc được thì đây là bộ lọc FIR đơn giản và tốt so với nhiều bộ lọc thông giải khác.

f. Peak Detetion and Recurring Point of interest Analysis.

Có nhiều giải thuật tìm đỉnh khác nhau, một giải thuật đơn giản là:

* Chia đồ thị thành những khoản nhỏ ví dụ n=linspace(40,230)
* Sau đó dùng một vòng lặp để kiểm tra sự thây đổi của từng giá trị và chọn những giá trị lớn nhất ở giữa một lần tăng và một lần giảm.

Lưu ý: trước khi sử dụng phương pháp nào thì đồ thị nên được smoothing (làm mượt)



Hình n: Ảnh minh hoạ cho thuật toán findpeaks

Sau bước này ta đã tìm được một số điểm RPoT tìm năng

Lưu ý:

*Trước khi sử dụng phương pháp này thì đồ thị nên được smoothing (làm mượt)*

*Ta có thể tìm hiểu thêm về một số thuật toán tìm đỉnh bằng cách gõ findpeaks in matlab.*

l. Heart Rate Display

# Chương 3. KẾT QUẢ THỰC HIỆN

# Chương 4. NHẬN XÉT VÀ HƯỚNG PHÁT TRIỂN

# TÀI LIỆU THAM KHẢO

[1] F. Xavier & A.Dias & Eduardo Peixoto. SOFTWARE FOR REAL TIME HEART RATE DETECTION USING A STANDARD WEBCAM, inc.

[2] Ming-Zher Pob & Daniel J. McDuff & Rosalind W. Picard. ADVANCEMENTS IN NONCONTACT, MUTILPARAMETER PHYSIOLOGICAL MEASUREMENTS USING A WEBCAM, inc 2011.

[3] H. Radman & M.U. Ahmed & S. Pegum & P. Funk. READ TIME HEART RATE MONOTORING FROM FACIAL RGB COLOR VIDEO USING WEBCAM, inc 2016

[4] Isayiyas Nigatu Tiba & Li Li. IMAGE-BASED AUTOMATIC PULSE RATE MONITORING SYSTEM USING PC WEBCAM, inc 2013.

[5] Larissa Carvalho & Mr. H.G.Virani & Mr.Shajahan Kutty. ANALYSIS OF HEART RATE MONITORING USING A WEBCAM, inc 2014.

[6] <https://www.mathworks.com/help/signal/ref/findpeaks.html>

[7] [https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/#](https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/)

[8] kazemi & Sullivan. ONE MILLISECOND FACE ALIGNMENT WITH AN ENSEMPLE OF REGRESSTION TREES, inc 2014.