

Gới thiệu giải pháp phần cứng cho bài toán thị giác máy trên nền linux nhúng

nm, lamnt
Trung tâm phát triển công nghệ FPT

1 Gới thiệu về thị giác máy – Computer Vision

Thị giác máy là một lĩnh vực đã và đang rất phát triển. Khái niệm thị giác máy – Computer vision có liên quan tới nhiều ngành học và hướng nghiên cứu khác nhau. Từ những năm 1970 khi mà năng lực tính toán của máy tính ngày càng trở nên mạnh mẽ hơn, các máy tính lúc này có thể xử lý được những tập dữ liệu lớn như các hình ảnh, các đoạn phim thì khái niệm và kỹ thuật về thị giác máy ngày càng được nhắc đến và nghiên cứu nhiều hơn cho tới ngày nay.

Hiện tại lĩnh vực được các chuyên gia đánh giá là vẫn còn “non nớt” và có rất nhiều sự thay đổi trong thời gian tới.

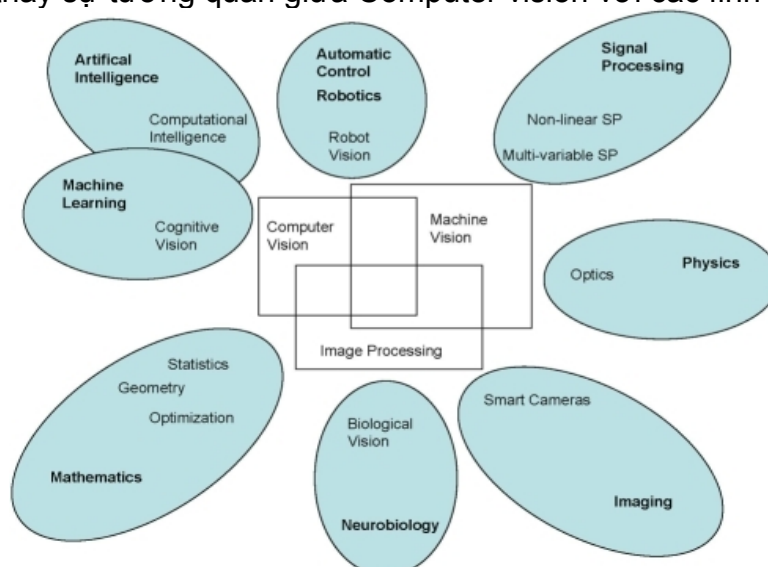
1.1 Thế nào là thị giác máy

Thị giác máy bao gồm lý thuyết và các kỹ thuật liên quan nhằm mục đích tạo ra một hệ thống nhân tạo có thể tiếp nhận thông tin từ các hình ảnh thu được hoặc các tập dữ liệu đa chiều.

Đối với mỗi con người chúng ta, quá trình nhận thức thế giới bên ngoài là một điều dễ dàng. Quá trình nhận thức đó được “học” thông qua quá trình sống của mỗi người. Tuy nhiên với các vật vô tri vô giác như các máy tính, robot v.v thì điều đó quả thực là một bước tiến rất gian nan. Các thiết bị ngày nay không chỉ nhận thông tin ở dạng tín hiệu đơn lẻ mà nay còn có thể có cái “nhìn” thật với thế giới bên ngoài. Cái “nhìn” này qua quá trình phân tích, kết hợp với các mô hình như máy học, mạng nơron v.v sẽ giúp cho thiết bị tiến dần tới một hệ thống nhân tạo có khả năng ra quyết định linh hoạt và đúng đắn hơn rất nhiều.

Lĩnh vực nghiên cứu của thị giác máy rất rộng, và đặc điểm chung là các bài toán về thị giác máy tính đều không có một đề bài chung và cách giải duy nhất. Mỗi giải pháp giải quyết vẫn đều được một kết quả nhất định cho những trường hợp cụ thể.

Ta có thể thấy sự tương quan giữa Computer vision với các lĩnh vực khác như sau:



Ảnh 1. Quan hệ giữa thị giác máy tính với các lĩnh vực liên quan

1.2 Ứng dụng

Một vài lĩnh vực mà Computer Vision được ứng dụng có thể kể tới như sau:

- Điều khiển tiến trình (ví dụ: trong các robot công nghiệp, hay các thiết bị, xe tự hành)
- Phát hiện sự thay đổi (ví dụ: các thiết bị giám sát)
- Tổ chức thông tin (ví dụ: chỉ số kho dữ liệu các ảnh hoặc chuỗi ảnh liên tục)
- Mô hình hoá đối tượng (ví dụ: quá trình kiểm tra trong môi trường công nghiệp, xử lý ảnh trong y học)
- Tương tác (đóng vai trò làm đầu vào cho thiết bị trong quá trình tương tác giữa người và máy)

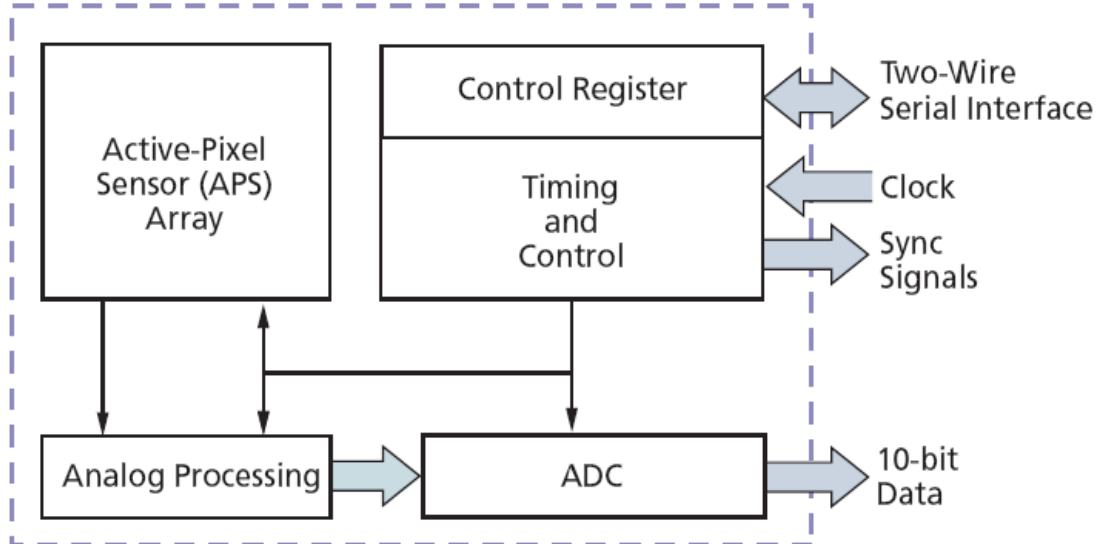
2 Các giải pháp phần cứng khả dụng hiện nay

Để giải bài toán xử lý bằng hình ảnh, ta có rất nhiều giải pháp khác nhau. Từ phần cứng cho tới các module phần mềm. Ta có thể điểm qua một vài phương pháp hay được sử dụng và có hiệu quả cao như sau.

2.1 FPGA

Đây là giải pháp về phần cứng mà hay được sử dụng hiện nay. Tận dụng đặc tính của FPGA là linh hoạt và tốc độ xử lý nhanh (cứng hoá các giải thuật). Ta có thể kết nối FPGA với các CCD sensor để điều khiển và nhận các frame hình thông qua bus dữ liệu tốc độ cao.

Ví dụ ta có thể tham khảo mô hình của CCD Sensor MT9T001 của Micron như sau:



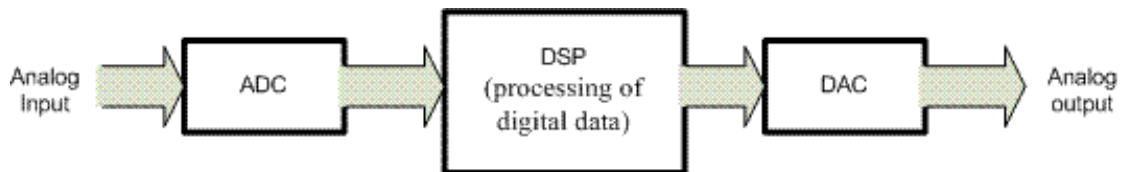
Ảnh 2 : CCD Sensor MT9T001 của hãng Micron

Ta có thể thao tác trực tiếp vào các thanh ghi điều khiển của MT9T001 thông qua bus TWI, và lấy các frame hình qua bus data 10bit. FPGA có tốc độ nhanh sẽ làm nhiệm vụ điều khiển MT9T001 và lấy dữ liệu liên tục từ data bus sau đó đưa vào bộ nhớ tạm để xử lý.

2.2 DSP Processor

DSP Processor được giới thiệu đầu tiên vào những năm 1978, 1979 bởi Intel, Bell Labs. Các bộ xử lý DSP có những đặc tính nổi bật như sau:

- Thích hợp cho các quá trình cần xử lý theo thời gian thực
- Hiệu năng được tối ưu với dữ liệu dạng luồng
- Chương trình và dữ liệu được bố trí riêng biệt (kiến trúc Harvard)
- Tích hợp các chỉ thị lệnh đặc biệt SIMD (Single Instruction, Multiple Data)
- Không hỗ trợ đa nhiệm
- Tương tác trực tiếp với bộ nhớ của thiết bị
- Tích hợp sẵn ADC và DAC



Ảnh 3. Mô hình của một DSP Processor

DSP Processor ngày nay đã được tích hợp nhiều thành phần khác nhau, làm tăng khả năng linh hoạt và tốc độ xử lý. Đặc biệt các DSP Processor rất thích hợp cho những nhu cầu cần tính toán nhanh, xử lý số thực. Đặc biệt một số còn có sẵn những chỉ thị lệnh giúp cho việc tính toán ma trận, tích chập hay thậm chí các phép biến đổi DCT trong quá trình nén ảnh. Với những ưu điểm đó DSP Processor được dùng trong nhiều thiết bị xử lý ảnh chuyên nghiệp.

2.3 Mainboard, laptop

Đây là một trong những phương pháp đơn giản nhất. Có thể tận dụng các mainboard máy tính hay thậm chí các máy tính xách tay với chức năng là một đơn vị xử lý ảnh, và đưa ra quyết định. Với việc kết nối một camera hay webcam ta hoàn toàn chủ động trong quá trình nhận/xử lý ảnh. Các giao tiếp ngoại vi phổ biến như UART, Parallel, USB hay Keyboard.

Việc sử dụng mainboard, laptop sẽ có những ưu/ nhược điểm sau:

- Ưu điểm
- Có tốc độ xử lý cao
- Dễ dàng lập trình, kiểm lỗi
- Hệ điều hành quen thuộc (windows/linux)
- Các công cụ lập trình/biên dịch phổ biến (MSVC, gcc)
- Nhược điểm
- Kích thước, khối lượng lớn
- Dễ hư hỏng do va đập hay các tác nhân khác
- Giá thành rất đắt (>600\$)
- Có quá nhiều thành phần không sử dụng đến
- Chỉ có thể giao tiếp với ngoại vi thông qua các chuẩn phổ biến như UART, USB

2.4 Main công nghiệp, PC 104



Ảnh 4. Mainboard pc-104, kích thước 90.17mm × 95.89mm

Ta có thể sử dụng mainboard PC104 với các chức năng như một máy tính thông thường. Việc sử dụng PC104 sẽ có những ưu/ nhược điểm sau:

- Ưu điểm
- Có tốc độ xử lý cao
- Dễ dàng lập trình, kiểm lỗi
- Hệ điều hành quen thuộc (windows/linux)
- Các công cụ lập trình/biên dịch phổ biến (MSVC, gcc)

- Nhược điểm
- Kích thước lớn
- Có quá nhiều thành phần không sử dụng đến
- Giá thành đắt (>300\$)
- Chỉ có thể giao tiếp với ngoại vi thông qua các chuẩn phổ biến như UART, USB

2.5 Single Board Computer(SBC) sử dụng Single on Chip (SoC)



Ảnh 5. SBC uDimm, Kích thước 43.18x68.58 mm.

SBC (Single board computer) có sử dụng các SoC là bước tiến lớn trong việc xây dựng thiết bị với khả năng tính toán và xử lý như một máy tính thông thường. Với nhiều ưu điểm hơn hẳn so với các thiết bị khác như :

- Kích cỡ
- Giá thành.

- Sử dụng các SoC làm bộ xử lý trung tâm có khả năng chuyên biệt về xử lý hình ảnh hay xử lý trên môi trường mạng rất hữu hiệu.

Do vậy tùy và bài toán đặt ra mà ta có thể chọn các SBC cho các SoC phù hợp với yêu cầu cụ thể. Hiện nay đã có nhiều SoC có khả năng tích hợp các DSP Processor vào trong nhân nhằm tăng khả năng xử lý (sign processing).

Việc sử dụng mạch SBC sẽ có những ưu/ nhược điểm sau:

- Ưu điểm
- Kích thước rất nhỏ
- Giá thành rẻ (~<100\$)
- Hỗ trợ SPI, I2C, I2S, MMC, SDCard, UART, USB2.0/1.1 v..v
- Có tốc độ xử lý cao
- Sử dụng hệ điều hành Linux
- Được sự hỗ trợ rất lớn của cộng đồng mã nguồn mở. Từ HĐH, kernel hay rất nhiều các ứng dụng
- Các công cụ biên dịch phổ biến gcc
- Các công cụ hỗ trợ lập trình rất nhiều. Eclipse, Vim, Emacs v..v
- Nhược điểm
- Việc chạy/kiểm thử phải thực hiện giả lập trên máy tính trước khi đưa vào mạch
- Am hiểu kiến thức về các giao tiếp ngoại vi, kiến trúc về SBC

3 Đâu là cách tiếp cận dễ dàng và nhanh chóng nhất ?

Theo quan điểm của chúng tôi, cách tiếp cận nhanh chóng nhất với các bạn sinh viên đó là tận dụng môi trường mã nguồn mở hiện nay. Với mã nguồn mở các bạn có thể sử dụng lại những thành quả của cộng đồng và phát triển, sáng tạo những cái riêng cho mình một cách dễ dàng và nhanh chóng nhất.

Chúng ta có thể sử dụng HĐH Linux làm môi trường chính cho thiết bị, với mã nguồn mở ta có thể chủ động được cấu hình của hệ thống, tăng hiệu năng hay xử lý thô với các giao tiếp ngoại vi bên ngoài một cách nhanh chóng và thuận tiện nhất. Điều đáng quan tâm nữa là ta có thể có được sự trợ giúp (miễn phí) rất lớn cộng đồng mã nguồn mở. Đây là điều mà môi trường mã nguồn đóng như Microsoft chưa có thể có được.

3.1 Phần cứng

Ta có thể tận dụng lợi thế của các SBC có sử dụng SoC để làm phần cứng cho các thiết bị trong bài toán thị giác máy. Lợi thế của sự chọn lựa này là các mạch SBC giá rẻ và rất linh hoạt giúp cho người thiết kế tập trung hơn vào các module phần mềm.

Ta có thể điem qua một vài mạch SBC như sau:

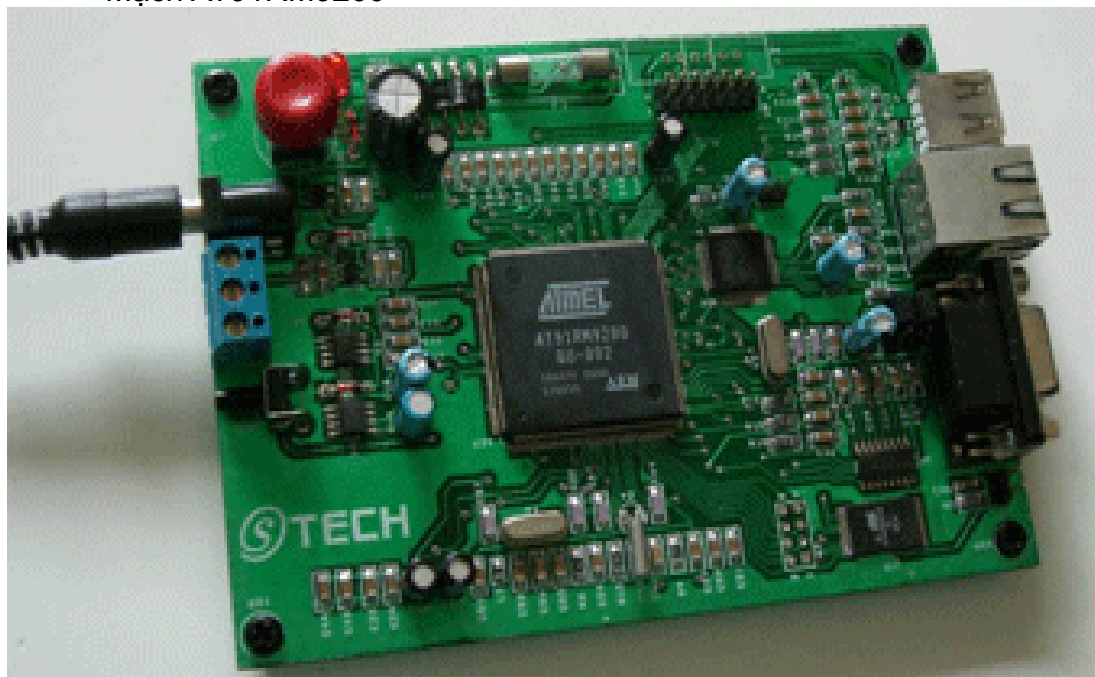
- Mạch NSLU2



Các thông số kỹ thuật

- Hãng sản xuất **Linksys**
- Kiến trúc ARM
- Sử dụng bộ VĐK SoC, Intel IXP420. Tốc độ 533MIPs
- 32MB Bộ nhớ SDRAM , 8MB Flash
- Hai cổng USB 2.0 Host
- Hai cổng Ethernet
- Giao tiếp UART với tốc độ 900Kb/s
- 16 cổng lập trình được (GPIO)
- Giá ~<100\$

- Mạch AT91RM9200



Các thông số kỹ thuật

- Hãng sản xuất : Công ty phát triển công nghệ FPT
- Kiến trúc ARM
- Sử dụng bộ VĐK SoC, Atmel AT91RM9200. Tốc độ 200MIPs
- 32MB Bộ nhớ SDRAM , 4MB DataFlash
- USB 2.0 Host
- Ethernet controller
- 4 cổng UART với tốc độ 900Kb/s
- 32 cổng lập trình được (GPIO)

3.2 Phần mềm

Như đã nói ở trên, có rất nhiều ứng dụng sử dụng thị giác máy. Mỗi ứng dụng sử dụng một dãy các chức năng thuộc về thị giác máy. Sau đây là danh sách các chức năng đặc trưng của thị giác máy

Nhận dạng đối tượng

Nhận dạng ảnh là một trong những vấn đề kinh điển trong lĩnh vực thị giác máy và xử lý ảnh. Mục đích của nhận dạng ảnh tức là xác định xem liệu một bức ảnh có tồn tại những đối tượng, đặc điểm đặc biệt hay không. Chức năng này có thể thực hiện tự động, không cần sự tác động của con người; tuy nhiên nó không đảm bảo trong các trường hợp chung, với những đối tượng tùy ý và trong tình huống tùy ý. Các phương pháp hiện nay để giải quyết vấn đề này chỉ được ứng dụng với những đối tượng đặc biệt(những đối tượng hình học cơ bản, nhận dạng mặt người, nhận dạng chữ in và chữ viết tay, ...) và trong những tình huống đặc biệt (trong điều kiện chiếu sáng được xác định trước, ...)

Nhận dạng chuyển động

Nhận dạng chuyển động được áp dụng để xử lý một chuỗi các ảnh liên tiếp để ước lượng tốc độ chuyển động của từng điểm ảnh.

- Egomotion: Xác định chuyển động trong khung cảnh 3D từ camera. Áp dụng cho các camera theo dõi tự động.
- Tracking: Theo dõi chuyển động của một đối tượng nào đó, ví dụ như người hay xe cộ

Xây dựng lại khung cảnh

Chức năng này dùng để tính toán khung cảnh 3D từ một vài bức ảnh, đoạn video.

Khôi phục ảnh

Mục đích của chức năng này là loại bỏ nhiễu có trong ảnh.

4 Ứng dụng cho bài toán robot di chuyển

Việc áp dụng thị giác máy vào bài toán robo di chuyển là một việc phức tạp. Tuy nhiên cũng có nhiều thuận lợi như chúng ta đã biết trước được mô hình sân đấu, phần lớn đối tượng có thể có trên sân. Tuy nhiên chúng ta cũng sẽ gặp nhiều khó khăn như không biết trước được điều kiện chiếu sáng của sân và robot của đối thủ. Sau đây tôi sẽ đưa ra từng bước các chức năng có thể ứng dụng vào bài toán này:

Thu thập ảnh:

Chụp ảnh từ webcam. Chất lượng của ảnh thu được từ bước này phụ thuộc nhiều vào

chất lượng webcam, tốc độ di chuyển khi chụp, góc nhìn và độ sáng.

Các bước tiền xử lý:

Trước khi áp dụng các chức năng của thị giác ảnh, chúng ta cần xử lý ảnh để đảm bảo rằng ảnh sẽ thỏa mãn một vài giả thiết nào đó. Ví dụ như sau

- Lấy mẫu lại để đảm bảo hệ thống tọa độ là đúng. Nếu 2 ảnh lấy liên tiếp nhau có sự sai khác quá nhiều thì chúng ta có thể loại bỏ ảnh đó vì tư thế của robot lúc chụp không được ổn định.
- Loại bỏ nhiễu để giảm thiểu khả năng xuất hiện các thông tin sai
- Tăng độ tương phản để đảm bảo sẽ xác định được các thông tin cần thiết.

Trích rút các điểm đặc trưng:

Đặc trưng của ảnh ở những mức phức tạp khác nhau sẽ được trích rút từ ảnh, như là các đường, cạnh, góc.

Phân đoạn :

Tại một vài điểm ảnh, chúng ta chọn ra một tập các điểm ảnh hoặc các vùng ảnh phù hợp với các thao tác sau này. Ví dụ

- Chọn ra một tập các điểm ảnh có màu gần với màu của “ụ” trong sân đấu
- Phân ra làm một hay nhiều vùng ảnh có khả năng chứa đối tượng “ụ”

Xử lý cấp cao

Đến bước này thì dữ liệu đầu vào chỉ là một phần của dữ liệu ban đầu, đó có thể là tập các điểm ảnh hoặc một vùng ảnh có khả năng chứa một đối tượng đặc biệt nào đó. Quá trình xử lý sẽ qua các bước sau:

- Thẩm định lại dữ liệu có phù hợp với các yêu cầu cơ bản và đặc biệt
- Ước lượng các tham số đặc biệt
- Phân lớp đối tượng xác định được.

5 Tham khảo

1. [Computer vision](http://en.wikipedia.org/wiki/Computer_vision) http://en.wikipedia.org/wiki/Computer_vision
2. [Image Processing](http://en.wikipedia.org/wiki/Image_processing) http://en.wikipedia.org/wiki/Image_processing
3. [Intel Open Computer Vision](http://www.intel.com/technology/computing/opencv/) <http://www.intel.com/technology/computing/opencv/>
4. [CMVision library](http://www.cs.cmu.edu/~jbruce/cmvision/) <http://www.cs.cmu.edu/~jbruce/cmvision/>
5. [Embedded computer vision](http://www.movesinstitute.org/~kolsch/pubs/EmbeddedCV.html)
<http://www.movesinstitute.org/~kolsch/pubs/EmbeddedCV.html>