

**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ
KHOA CƠ HỌC KỸ THUẬT VÀ TỰ ĐỘNG HÓA**



BÁO CÁO TÌM HIỂU CẢM BIẾN KINECT

GV : ThS.ĐOÀN THỊ HƯƠNG GIANG

NHÓM 11

LỚP K55 M

Thành viên nhóm 11

Nguyễn Tiến San

Nguyễn Thái Bảo Hùng

Nguyễn Văn Linh

Phạm Vũ Cường

Lê Sơn Hải

Trần Văn Kiên

Bùi Hồng Hà

Mục lục :

- I. Giới thiệu về cảm biến Kinect
- II. Cấu tạo
 - 1. Hệ thống cảm biến chiều sâu
 - 2. Camera RGB-D
 - 3. Động cơ, máy đo gia tốc, micro
- III. Cảm biến đo chiều sâu (depth sensor)
- IV. Cảm biến hình ảnh RGB-D
- V. Một vài ứng dụng quan trọng

I. Giới thiệu chung

Sử dụng video :

1_Microsoft Kinect for Xbox 360

Kinect là gì?

Kinect là một thiết bị đầu vào, là cảm biến chuyển động do hãng Microsoft sản xuất dành cho Xbox 360 và máy tính Windows. Dựa trên một webcam kiểu add-on ngoại vi cho Xbox 360, nó cho phép người dùng điều khiển và tương tác với Xbox 360 mà không cần phải dùng đến một bộ điều khiển tay cầm, thông qua một giao diện người dùng tự nhiên bằng cử chỉ và lệnh nói.

Thiết bị được giới thiệu vào tháng 11 năm 2010 như một phụ kiện của Xbox 360. Cảm biến chiều sâu (depth sensor) được sử dụng trong Kinect được phát triển bởi Zeev Zalevsky, Alexander Shpunt, Aviad Maizels và Javier Garcia, năm 2005. Kinect đã chính thức được công bố vào ngày 01 Tháng Sáu Năm 2009, dưới cái tên "Project Natal" (Natal là tên một thành phố ở Brazil, nơi sinh của 1 giám đốc dự án này). Chức năng chính của Kinect là một công cụ để người dùng tương tác với Xbox 360 bằng cử chỉ và lệnh nói. Vì lý do này, các bộ cảm biến có khả năng thu thập dữ liệu ở độ phân giải 640x480 điểm ảnh. Với các dữ liệu chiều sâu, có thể lấy được một khung xương của người đứng phía trước của cảm biến. Và với bộ xương đó, nó có thể nhận biết được cử chỉ của người sử dụng.

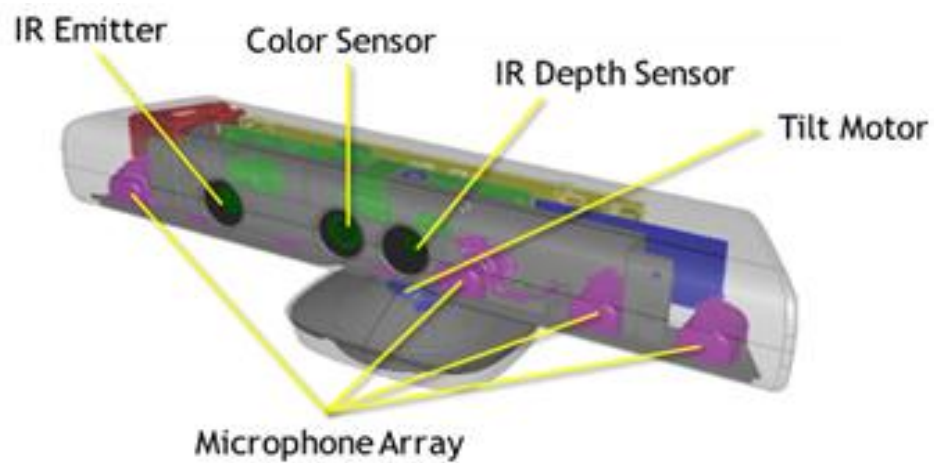
Một số hình ảnh về Kinect và Xbox 360

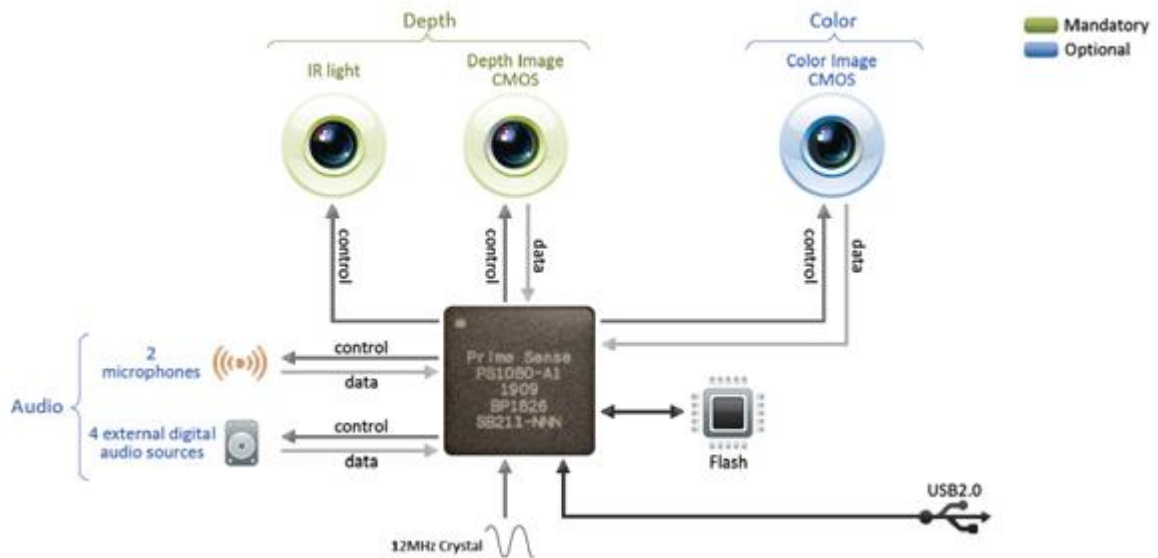


KINECT™
for  **XBOX 360.**



II. Cấu tạo





Sơ đồ khối

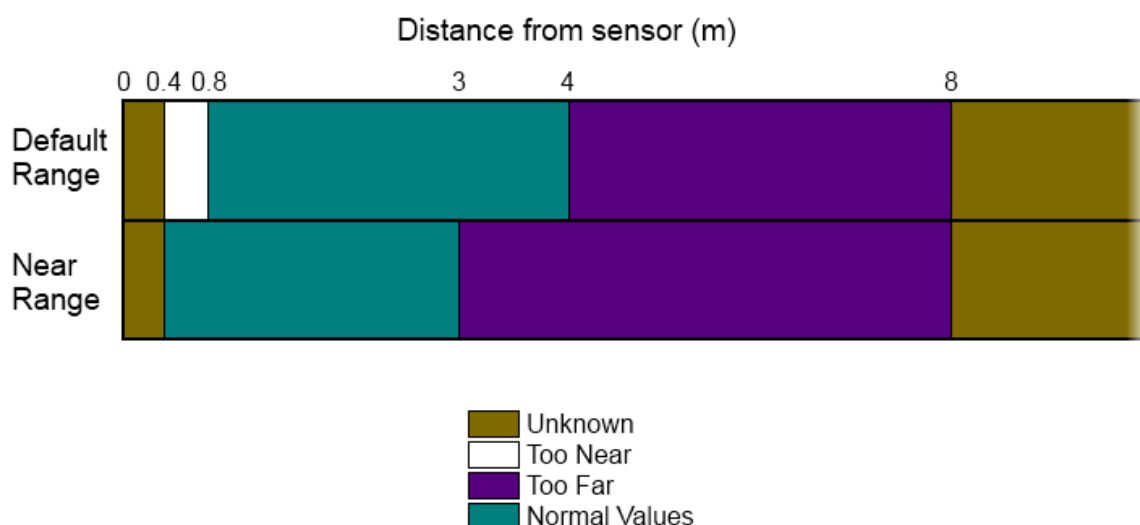
1. Hệ thống cảm biến chiều sâu

Bao gồm bộ phát hồng ngoại (Infrared light) và camera hồng ngoại (depth image CMOS). Kích cỡ ảnh là 640x480 pixel, tốc độ chụp 30 khung hình mỗi giây.

Có 2 phạm vi độ sâu:

+ Phạm vi mặc định có sẵn trong cả hai bộ cảm biến Kinect cho Windows và các cảm biến Kinect cho Xbox 360.

+ Phạm vi gần là chỉ có trong cảm biến Kinect cho Windows.



Để cảm biến hoạt động tốt nhất thì nên sử dụng khoảng cách từ 1,2m → 3,5m.

Hoạt động : Các tia hồng ngoại được chiếu qua bộ phát hồng ngoại đến đối tượng, sau đó camera hồng ngoại sẽ thu thập dữ liệu bị phản chiếu.

2. Camera RGB-D

Chức năng chính của camera là nhận biết 3 màu cơ bản là đỏ, xanh lá cây và xanh da trời (Red-Green-Blue). Quá trình chụp bao gồm việc chụp một ảnh màu (RGB) và thực hiện một phép đo độ sâu (D). Cảm biến hình ảnh kết hợp với cảm biến chiều sâu nằm ở gần nhau, cho phép sáp nhập bản đồ, cho ra hình ảnh 3D. Thông tin ảnh RGB-D được lưu trữ.

Với kích cỡ 1280x960 pixel, tốc độ 12 khung hình/giây. Với kích cỡ 640x480 pixel, tốc độ chụp 30 khung hình/giây. Từ đây cho thấy, khi thiết lập độ phân giải thấp hơn thì tốc độ chụp và truyền hình ảnh là cao hơn.

Với dữ liệu sâu thu được, nó sẽ tạo ra một bản đồ về bộ xương người đứng ở trước cảm biến. Và với bộ xương đó, nó có thể xác định được cử chỉ, hành động của người sử dụng.

3. động cơ, máy đo gia tốc, micro

- Động cơ : nằm ở đế của cảm biến Kinect. Nó có khả năng làm cho cảm biến hướng đầu lên trên cao và hạ xuống thấp. Với khả năng này sẽ giúp cho cảm biến có thể thu được hình ảnh cao hơn rất nhiều so với khi không có bộ động cơ.
- Máy đo gia tốc : Kinect sử dụng nó để xác định người đứng trước, đứng sau trong khi đo.
- Micro : bao gồm 4 micro nhỏ, hoạt động như nhau ở 16 bit với tốc độ lấy mẫu là 16kHz (số mẫu/giây).

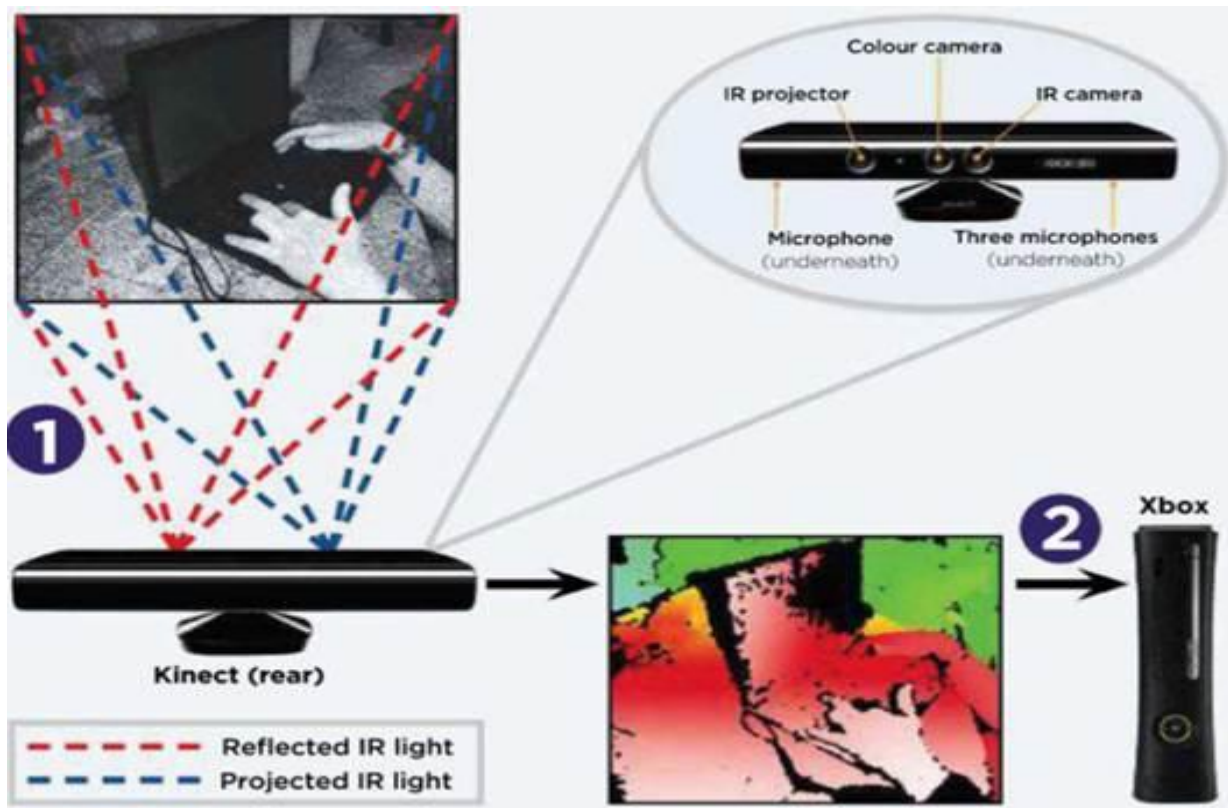
4. Sử dụng chip PrimeSensor của Microsoft sản xuất.

5. Phần mềm hỗ trợ :

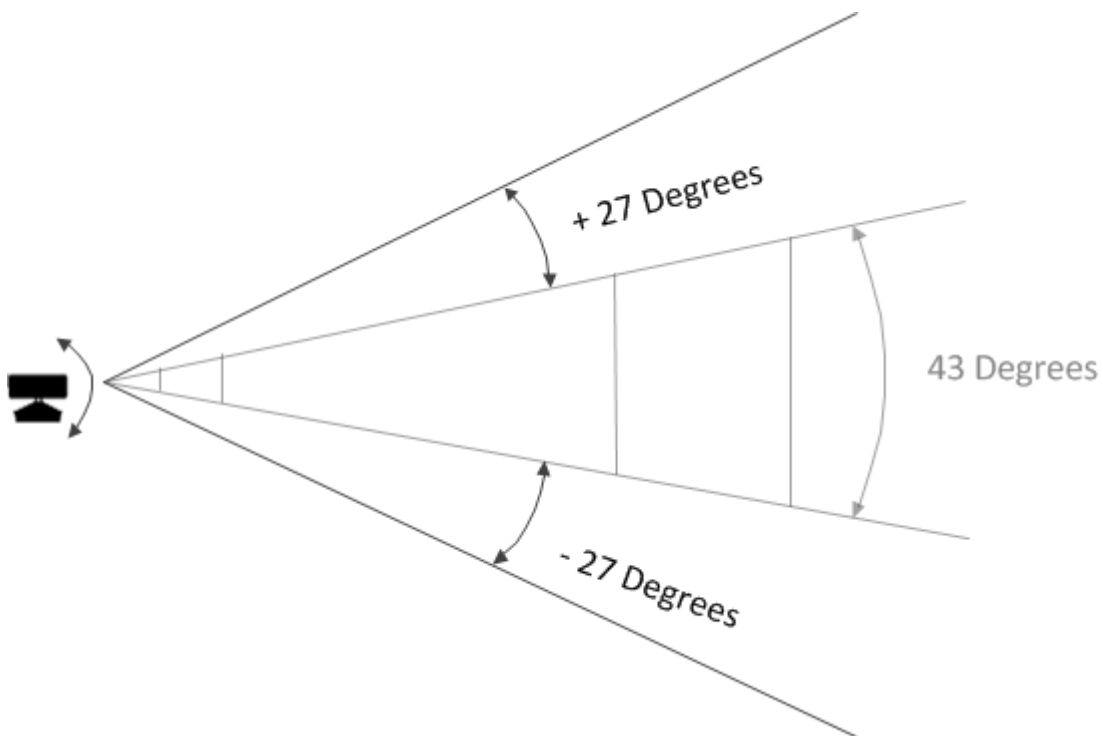
- + OpenNI : được sử dụng phổ biến nhất.
- + OpenKinect : do hãng Microsoft sản xuất.
- + Microsoft Kinect dành cho hệ điều hành Windows.

III. Cảm biến đo chiều sâu (depth sensor)

Video2 : 2_depth sensor work

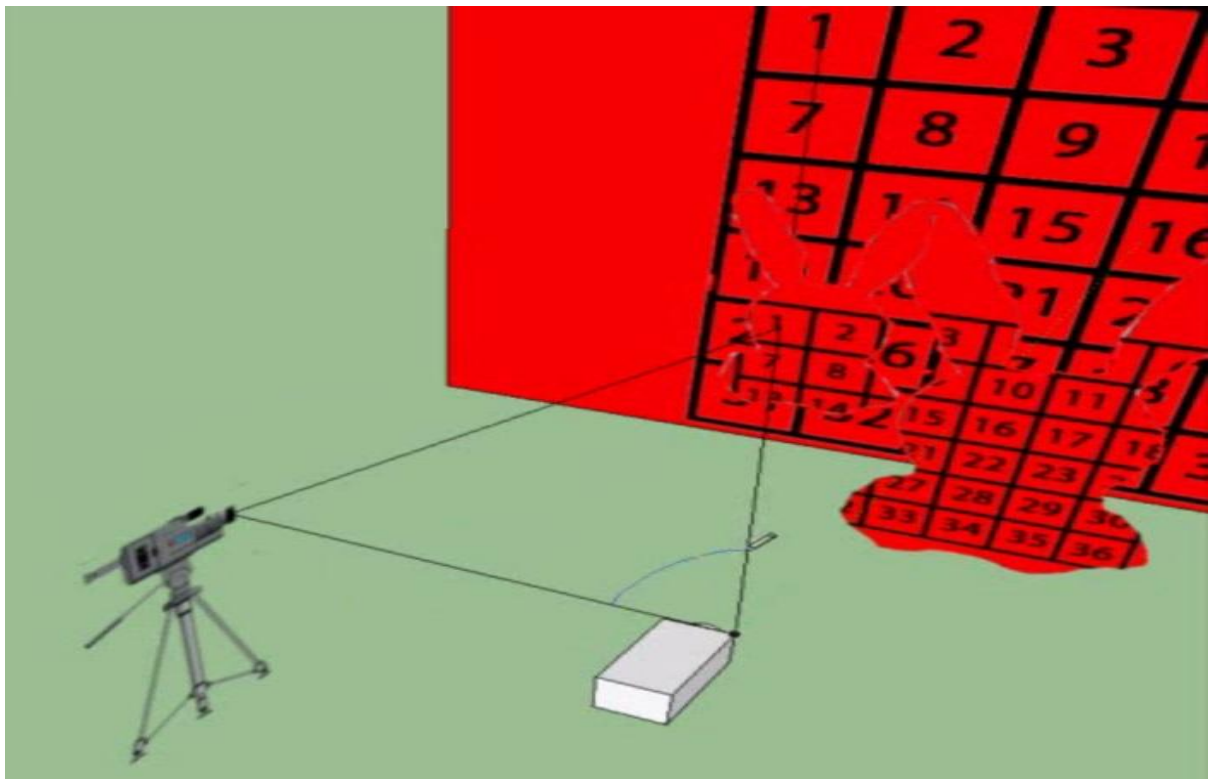
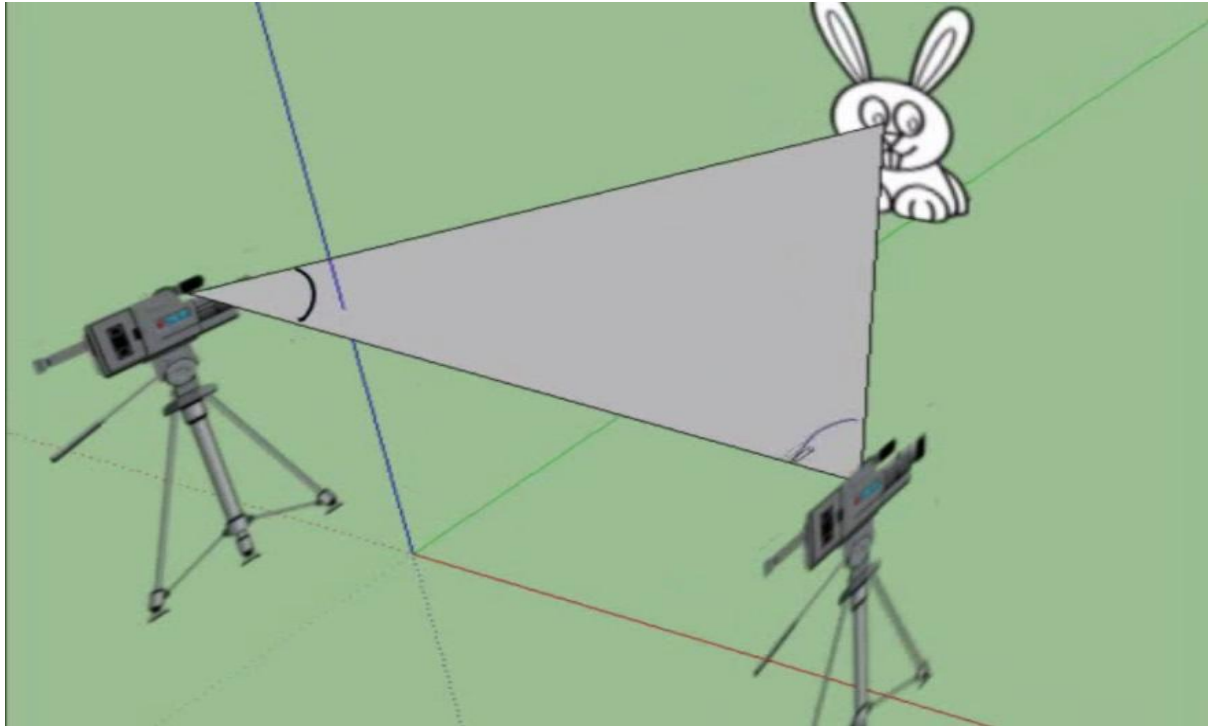


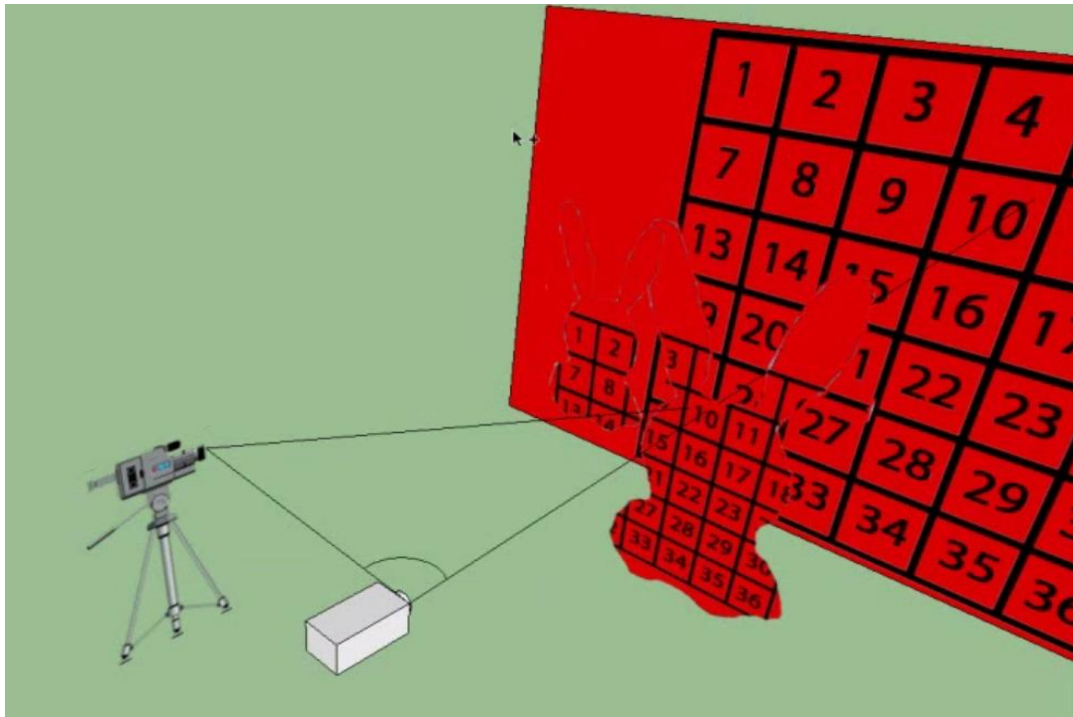
Vùng hoạt động của hệ thống là 43 độ dọc, 57 độ theo chiều ngang.



Cơ chế hoạt động :

Ban đầu, máy phát tia hồng ngoại sẽ phát ra tia hồng ngoại trong vùng hoạt động của nó, nó sẽ tạo ra một cấu trúc các điểm sáng trên vật thể, thông qua phản chiếu về hình ảnh, camera hồng ngoại sẽ thu thập cấu trúc này, cho kết quả là một hình ảnh vật thể với những chiều sâu khác nhau.



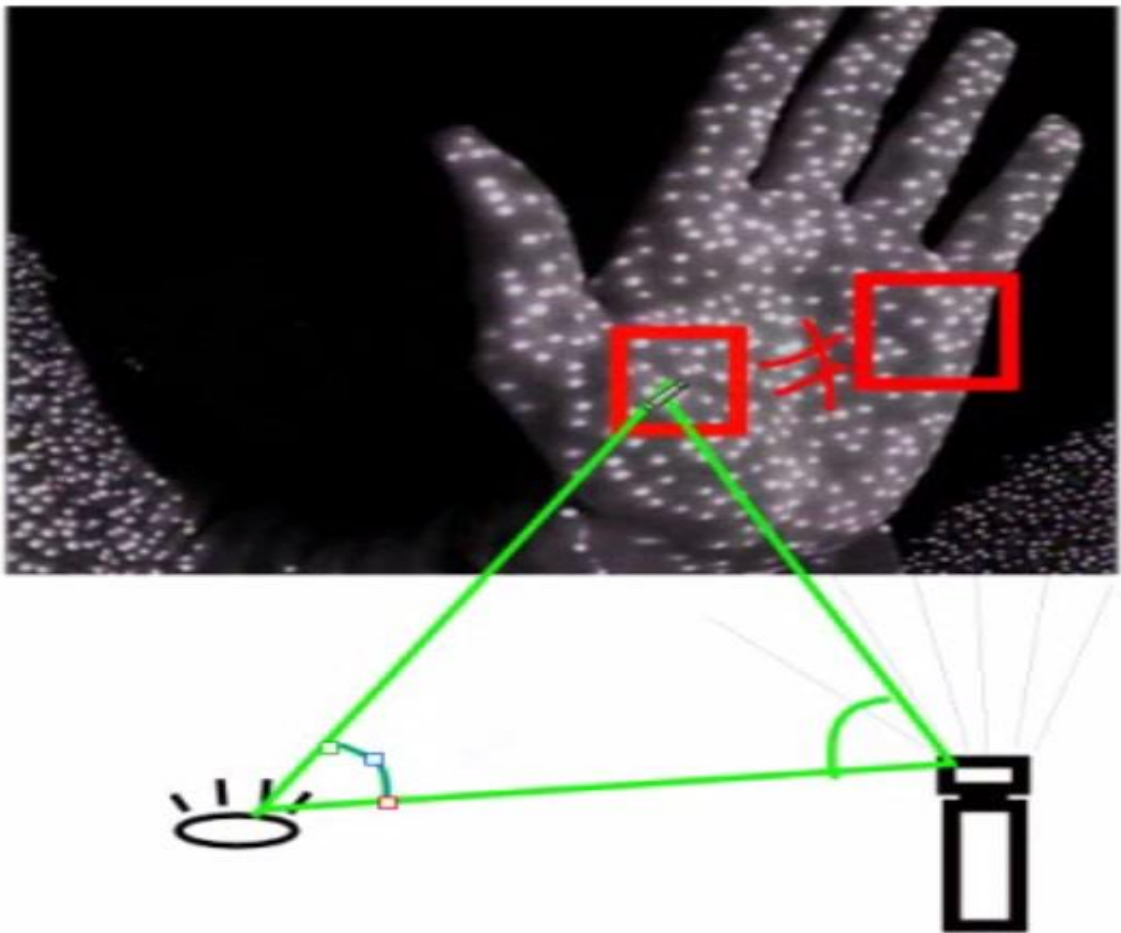


IV. Cảm biến hình ảnh (RGB-D)

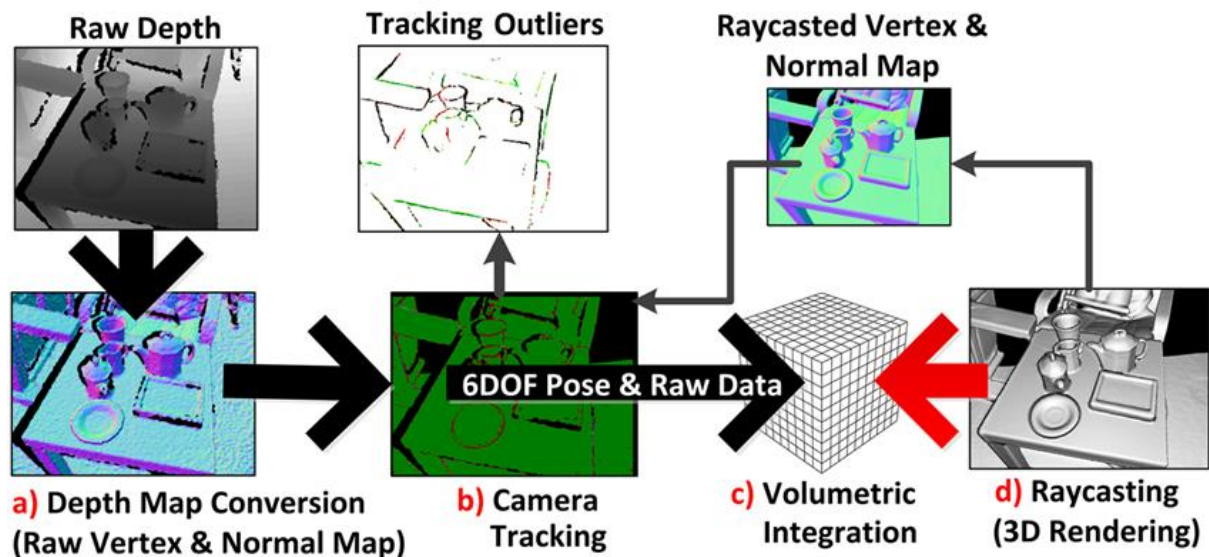


Bao gồm việc chụp và xử lý ảnh.

+ Chụp ảnh : bao gồm việc chụp ảnh RGB và quét 3D. Camera chụp cho ra ảnh RGB có kích cỡ mặc định là 640x480 pixel, tốc độ chụp là 30fps. Nhưng camera màu này có thể chụp với định dạng cao là 1280x960 pixel, tốc độ chụp 12fps. Khi chụp, nó kết hợp với camera đo độ sâu để tạo ra được ảnh có độ sâu (3D). Đo độ sâu được thực hiện bằng cách sử dụng 1 kỹ thuật ánh sáng có cấu trúc. Ánh sáng có cấu trúc được thực hiện bằng cách chiếu 1 mô hình các điểm ảnh lên vùng hoạt động, những điểm ảnh gập bề mặt khác nhau của vật thể sẽ cho phép hệ thống quan sát tính toán độ sâu của đối tượng.



Con đường tạo ra ảnh 3D:



Thứ nhất :

Phác thảo bản đồ → định hướng bề mặt

Thứ hai : camera theo dõi

Camera tính toán tổng quát (vị trí và hướng của đối tượng) → kết hợp kiểm tra hướng và kiểm tra các đỉnh. Tính toán liên tục, sắp xếp lặp đi lặp lại để cho ra ảnh tốt nhất.

Thứ ba : hội nhập

Kết hợp dữ liệu sâu từ cảm biến đã biết sẽ đưa ra biểu diễn một diện tích duy nhất của không gian xung quanh máy ảnh.

Một cảm biến chuyển động quan sát một bề mặt từ những điểm hơi khác nhau, nhiều những khoảng trống hay lỗ, nơi dữ liệu độ sâu không có trong hình ảnh Kinect ban đầu cũng có thể được lấp đầy (ví dụ như bạn có thể di chuyển cảm biến xung quanh một đối tượng để điền vào những điểm trước đó chưa được xác định phía sau của đối tượng) và bề mặt liên tục được cải tiến với các dữ liệu ở độ phân giải cao.

Thứ tư : dựng hình 3D

Khối lượng tái tạo được dò tia (Raycast) từ cảm biến, điều này cho kết quả là đám mây điểm được tô bóng hơn, kết hợp lại thành hình ảnh 3 chiều.

Ghi chú: Dò tia là kỹ thuật tạo ra một hình ảnh bằng cách dò đường đi của ánh sáng thông qua các điểm ảnh trên một tấm kính ảnh. Kỹ thuật này có khả năng đạt tới một mức độ rất cao như ảnh thật, thậm chí hơn các phương pháp kết xuất dạng quét dòng thông thường.

(Theo <http://vi.wikipedia.org>)

+ Xử lý ảnh : các hình ảnh RGB và dữ liệu độ sâu có một số vấn đề về nhiễu và các lỗ (do độ phân giải của máy ảnh kém nên tạo ra các điểm ảnh chưa được xử lý). Có rất nhiều giải pháp ở đây để xử lý chúng. Tôi sẽ đưa ra một giải pháp đơn giản mà hiệu quả. Đó là lấy giá trị trung bình của các điểm ảnh xung quanh các lỗ đó.

V. Một vài ứng dụng quan trọng

1. Tương tác :

suy luận vị trí cơ thể là một quá trình 2 giai đoạn.

+ tính toán một bản đồ độ sâu (sử dụng ánh sáng có cấu trúc).

+ sau đó suy ra vị trí cơ thể (sử dụng học máy).



Học máy (learning machine) là gì : Học máy, có tài liệu gọi là Máy học là một lĩnh vực của trí tuệ nhân tạo liên quan đến việc phát triển các kỹ thuật cho phép các máy tính có thể "học". Cụ thể hơn, học máy là một phương pháp để tạo ra các chương trình máy tính bằng việc phân tích các tập dữ liệu. Học máy có liên quan lớn đến thống kê, vì cả hai lĩnh vực đều nghiên cứu việc phân tích dữ liệu, nhưng khác với thống kê, học máy tập trung vào sự phức tạp của các giải thuật trong việc thực thi tính toán. Nhiều bài toán suy luận được xếp vào loại

bài toán NP-khó, vì thế một phần của học máy là nghiên cứu sự phát triển các giải thuật suy luận xấp xỉ mà có thể xử lý được.

Học máy có tính ứng dụng rất cao bao gồm máy truy tìm dữ liệu, chẩn đoán y khoa, phát hiện thẻ tín dụng giả, phân tích thị trường chứng khoán, phân loại các chuỗi DNA, nhận dạng tiếng nói và chữ viết, dịch tự động, chơi trò chơi và cử động rô-bốt (robot locomotion).

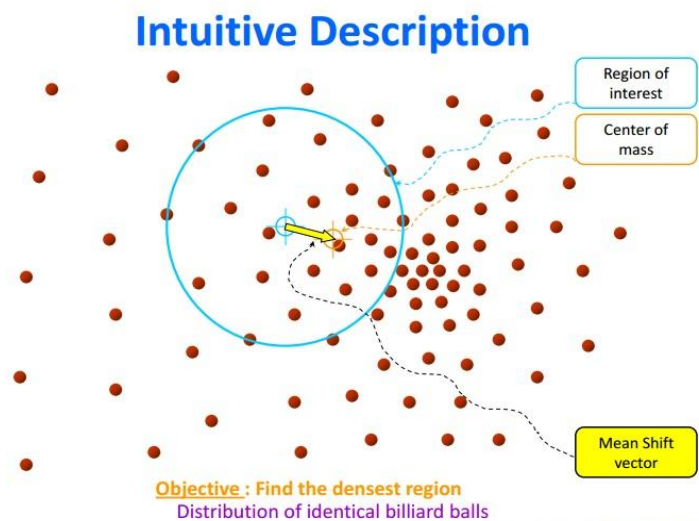
(Theo <http://vi.wikipedia.org>)

Giai đoạn 1: được trình bày như cơ chế hoạt động ở trên. Máy chiếu sẽ chiếu các điểm ảnh lên vật thể để đo độ sâu.

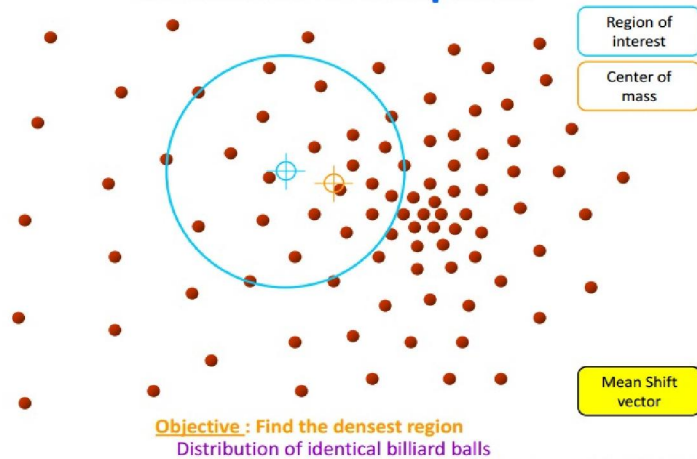
Giai đoạn 2: bộ phận cơ thể được tính toán độ sâu, chiều cao ... rồi được so sánh với những mẫu người tương ứng trong bộ nhớ, cho ra một mẫu cơ thể ảo chính xác với cơ thể thật. Các mẫu được tổng hợp từ 1 triệu mẫu được lưu trước đó.

Các thuật toán chuyển đổi trung bình được sử dụng để tính toán mạnh mẽ phương thức phân bố xác suất các điểm trên cơ thể, nhằm mục đích tìm thấy những khu vực dày đặc nhất.

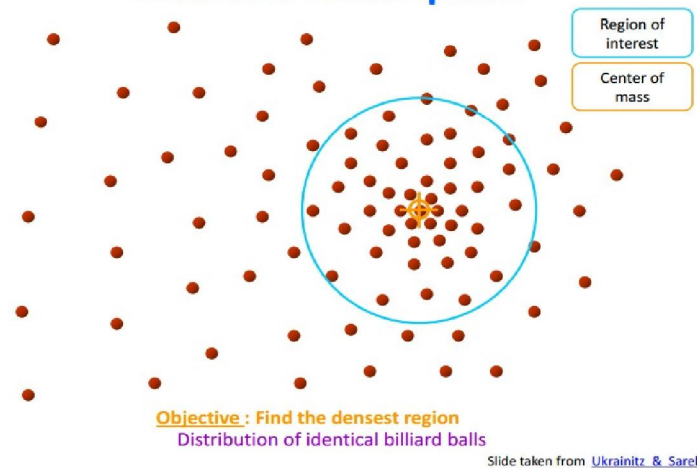
Mô tả trực quan



Intuitive Description



Intuitive Description



The results are great!



Sau khi có bản đồ bắt các điểm trên cơ thể người, ứng dụng của nó rất phong phú. Có thể nhập vai vào các nhân vật trong game để chơi game, được hướng dẫn tập thể dục tăng sức khỏe,... nó có ở trong Xbox 360.

2. Ứng dụng tạo bản đồ 3D

Sử dụng hệ thống Kinect Fusion để tạo bản đồ 3d. Nó sử dụng cả 2 camera để ghi lại hình ảnh và dữ liệu chiều sâu, dữ liệu được lưu trữ, ngay lập tức nó tạo ra được bản đồ 3d.

Ứng dụng này rất hữu ích nếu như bạn cần ghi lại bản đồ ngôi nhà của bạn. nó rất sinh động.

Video: [depth camera_1](#), [depth camera_2](#)



Frontal View (elevated, FOV=60)

Rear View (elevated, FOV=60)



3. ứng dụng vào robot và nghiên cứu khoa học

Kinect được ghép với máy tính cộng với việc viết code cho nó, có nhiều ứng dụng như xác định vật cản khi nó di chuyển, xác định màu sắc, khoảng cách với vật thể ...

Nó cũng có nhiều ứng dụng quan trọng trong nghiên cứu khoa học.

Khi được kết nối với Xbox 360, nó có thể bắt wifi, kết nối dữ liệu thông qua cổng usb của Xbox. Việc kết nối bạn bè trở nên dễ dàng hơn.

Tài liệu tham khảo

<http://msdn.microsoft.com/en-us/library/dn188670.aspx>

<http://blogs.howstuffworks.com/2010/11/05/how-microsoft-kinect-works-an-amazing-use-of-infrared-light/>

<http://www.aurelien.plyer.fr/the-kinect/how-the-kinect-work/>

<http://www.youtube.com/>

<http://www.wikipedia.org/>