**ỨNG DỤNG CAMERA KINECT TRONG XÂY DỰNG MÔ HÌNH CHẤM ĐIỂM ĐỘNG TÁC VÕ CỔ TRUYỀN VIỆT NAM**

APPLICATION OF CAMERA KINECT IN GRADING GESTURE OF VIETNAMESE MARTIAL ARTS

***Nguyễn Tường Thành1, Nguyễn Đăng Tuyên2, Lê Dũng1, Phạm Thành Công1***

*1Trường Đại học Bách khoa Hà Nội;thanh1277@gmail.com,ledung1974@gmail.com, cong.phamthanh@hust.edu.vn*

*2Viện Công nghệ thông tin- Viện Hàn lâm công nghệ Việt Nam; nguyendangtuyen@quangtrung.edu.vn*

**Tóm tắt-** Bài báo đã đề xuất thực hiện đưa dữ liệu độ sâu ảnh vào công thức chấm điểm động tác người sử dụng camera Kinect [1] cho động tác võ cổ truyền Việt Nam, đã xây dựng cơ sở dữ liệu mẫu 3D của 36 động tác võ giữa giờ đang được đưa vào dạy tại các trường phổ thông của Việt Nam và phần mềm chấm điểm động tác tĩnh. Đây là bước mở đầu của việc thực hiện tự rèn luyện và đánh giá võ cổ truyền Việt Nam, góp phần cho việc bảo tồn và phát triển tinh hoa dân tộc. Chương trình hứa hẹn có nhiều hướng mở rộng khi có thể tích hợp kiểm tra đánh giá qua mạng sau này. Thống kê cho thấy chương trình bước đầu đã đánh giá tương đối chính xác các động tác võ. Hệ thống làm việc hiệu quả với độ phức tạp tính toán thấp và khả năng kháng nhiễu mạnh mẽ.

**Từ khóa-** Dữ liệu khung xương; chấm điểm động tác; véc tơ chi; tự rèn luyện; võ cổ truyền Việt Nam.

**Abstract-** The paper has proposed the inclusion of the data image depth into the formula grading camera movements Kinect users [1] to the traditional martial arts movements Vietnam, has built a database of 3D models 36 movements martial between now being put into teaching at schools in Vietnam and grading software static movements. This is the first step of the implementation of self-training and evaluation of traditional martial Vietnam, contributing to the preservation and development of ethnic elites. The program promises to have more expansive when can integrate online assessment later. Statistics show that the program was initially relatively accurate assessment of martial arts movements. System work effectively with low computational complexity and strong resistance to interference

**Key words-** skeleton data; gesture grading; limb vectors; self-training; martial Vietnam.

# Đặt vấn đề

Nhận dạng động tác người tự động là một trong những phần quan trọng trong các ứng dụng tương tác giữa Người- Máy[5]. Trong đó xác định thành công hình dáng người chắc chắn có thể giúp một máy tính có thể kiểm soát các động tác khác nhau trên cơ sở của sự tương tác với cùng một mẫu thu nhận trước đó. Trong những năm qua, đã có nhiều nghiên cứu sâu rộng đã được thực hiện trên cơ sở nhận dạng sinh trắc học, sử dụng các mẫu hành vi khác nhau như dáng đi, biểu diễn múa ba lê hay nhận dạng biểu hiện tâm lý người thông qua các biểu hiện môi, cơ mặt, giọng nói… là một trong những cách thức phổ biến của nhận dạng người [2][3].

Nhận dạng dáng đi của con người có lợi thế rất lớn trong sự nhận dạng từ các hình ảnh độ phân giải thấp, nơi các kỹ thuật sinh trắc học khác là không phù hợp vì không đủ thông tin điểm ảnh [4].

Võ cổ truyền Việt Nam là một đặc trưng của dân tộc Việt Nam hình thành qua quá trình đấu tranh giải phóng dân tộc. Có một đặc điểm cho đến nay là chưa có tính thống nhất cao giữa các môn phái nên vẫn chưa có hệ thống bài mẫu chuẩn hóa toàn thế giới giống các môn võ khác như karatedo hay taewondo...

Trong bài báo này, chúng tôi đề xuất đưa dữ liệu độ sâu ảnh áp dụng vào hệ thống chấm điểm động tác tĩnh của con người dựa trên khung xương của camera Kinect của tác giả Linwan Liu [1] để thực hiện xây dựng chương trình chấm điểm động tác một bài Võ cổ truyền Việt Nam bằng cách lấy các dữ liệu khung xương thu nhận để tiến hành đối chiếu với bộ dữ liệu mẫu và tiến hành chấm điểm theo công thức so sánh độ sai lệch. Camera Kinect (Hình 1) là sản phẩm của Microsoft đang được bán rộng rãi với giá thành thấp cho phép giao tiếp với con người thông qua các cử chỉ dựa trên hai đặc tính chính là thông tin về độ sâu ảnh và khả năng phát hiện, bám theo đặc tính cơ thể người. Đóng góp chính của bài báo này được đưa ra dưới đây:

* Thực hiện đưa dữ liệu độ sâu ảnh vào để nhận dạng chấm điểm các động tác võ tĩnh tuần tự của một bài võ cổ truyền theo cơ sở dữ liệu mẫu 3D cho trước. Qua khảo sát từ trước đến nay chưa có tác giả nào tiến hành nhận diện động tác võ cổ truyền với camera Kinect.
* Xây dựng được cơ sở dữ liệu mẫu 3D để chấm điểm các bài võ cổ truyền dành cho chương trình thể dục giữa giờ được đưa vào chương trình học của các trường phổ thông từ năm 2016
* Xây dựng chương trình phần mềm để người dùng có thể chọn độ khó của việc chấm thông qua các thông số nhập vào.

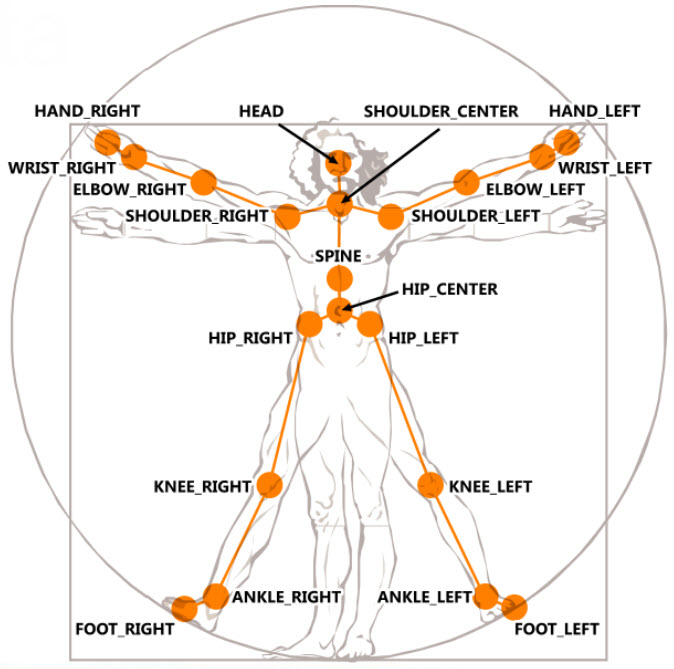


**Hình 1:** *Camera Kinect của Microsoft*

**2. Mô tả động tác người**

Việc xác định động tác người không phải là vấn đề đơn giản trong nhận dạng đối tượng. Chúng ta phải chọn một bộ cơ sở dữ liệu thích hợp để đại diện cho một hành động chính xác duy nhất. Khi đó cử chỉ mô tả cần chứa đầy đủ các đặc trưng của hành động để có thể phân biệt sự khác nhau của chúng và các mô tả phải bất biến khi vị trí đứng của người hay tầm vóc của người thay đổi khác

***2.1. Vị trí các khớp***



***Hình 2:*** *Vị trí các khớp xương camera Kinect cung cấp*

Camera Kinect cho kết quả tọa độ của 20 khớp xương xác định trong thời gian thực tạo thành một bản đồ khung xương của người biểu diễn [6][7] (Hình. 2). Các tọa độ của 20 khớp là duy nhất và hoàn toàn có thể đại diện cho một động tác, và các thiết lập dữ liệu là tương đối nhỏ cho máy tính. Nhưng việc mô tả thay đổi dựa trên thay đổi hình học các điểm ảnh là rất khó để xác định các số liệu khoảng cách.

***2.2. Véc tơ chi***

Camera Kinect thu nhận và cho chúng ta tọa độ các khớp xương nên chúng ta sử dụng các vector chi để biểu diễn hình ảnh khung xương của người biểu diễn. Một vector chi được định nghĩa là các đoạn đường nối giữa 2 khớp xương với nhau (Hình 3). Như các vector hình học, vector mô tả đại diện cho duy nhất một cử chỉ mà không dịch chuyển hay thay đổi tỷ lệ. Chuyển động cơ thể cũng có thể được mô phỏng hoàn hảo bằng chuyển động quay của mỗi vector xung quanh khớp ban đầu của nó.

|  |  |
| --- | --- |
|  | (b) |

***Hình 3:*** *Dữ liệu khung xương Kinect và véc tơ chân tay.*

Xem xét các đặc điểm của chuyển động của con người và cấu trúc cơ thể, như [1] đề xuất: chúng ta có thể giảm dữ liệu dư thừa bằng cách phân tích các vectơ chi trong các nhóm khác nhau.

1) Véc tơ đầu mình: véc tơ đầu mình bao gồm các đoạn bên trong đầu, vai và hông (Hình 3, với điểm đầu cuối màu đỏ).. véc tơ này ít khi thể hiện chuyển động riêng lẻ và mạnh. Xoay và uốn cong của phần thân thể này gắn chung với tứ chi. Do đó, chúng tôi loại bỏ các véc tơ này khỏi nhóm đại diện.

2) Véc tơ cấp độ1: véc tơ cấp độ 1 bao gồm khuỷu tay và đùi (Hình 3, với điểm đầu cuối màu cam).. véc tơ cấp độ 1 chứa rất nhiều thông tin của những chuyển động và cử chỉ. Do đó, chúng tôi sắp xếp chúng trong nhóm đại diện.

3) Véc tơ cấp độ 2: véc tơ cấp độ 2 bao gồm cánh tay và cẳng chân(Hình 3, với điểm đầu cuối màu xanh).. véc tơ cấp độ hai được kéo dài ra hơn so với vector cấp độ 1, và chúng tạo ấn tượng thị giác đáng kể. Do đó, chúng tôi xếp chúng trong nhóm đại diện.

4) Bàn tay và bàn chân: Kinect có thể theo dõi bàn tay và bàn chân nút (hình 3, các nút màu đen), nhưng trong quá trình thu nhận thường không ổn định.. Chuyển động của bàn tay và bàn chân liên quan đến cổ tay và mắt cá chân thường không đáng kể. Do đó trong giai đoạn ban đầu này, chúng tôi loại bỏ bàn tay và bàn chân ra khỏi nhóm đại diện để đảm bảo sự vững mạnh của mô tả.

Tóm lại, nhóm đại diện bao gồm tám vectơ chi (Bảng 1). Như chúng ta có thể thấy, kích thước dữ liệu được giảm đáng kể trong khi các thông tin cần thiết của động tác vẫn được giữ lại tốt.

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | Khủy tay trái | 5 | Cánh tay trái |
| 2 | Khủy tay phải | 6 | Cánh tay phải |
| 3 | Đùi trái | 7 | Bắp chân trái |
| 4 | Đùi phải | 8 | Bắp chân phải |

***Bảng 1:*** *Tám véc tơ chi*

***2.3. Đánh giá động tác***

Trong tài liệu tham khảo [1] là dùng cho hành động có thể biểu diễn 2D nhưng với động tác võ thì động tác luôn có chiều sâu thay đổi và động tác che khuất nên tác giả đã đề xuất đưa dữ liệu độ sâu ảnh vào công thức tính toán góc nằm giữa vectơ chi tương ứng thu được thực tế và các động tác chuẩn sử dụng công thức cosin, được định nghĩa là:

cos α = (1)

Trong đó () biểu thị tọa độ của một vectơ chi của các cử chỉ chuẩn, () biểu thị tọa độ của một vectơ chân tay của động tác trong thời gian thực. Chúng tôi xác định tập hợp bao gồm tám góc như là mảng khoảng cách AngleDiff = {} (xem Bảng 1), nó chứa hầu hết các thông tin về khoảng cách giữa động tác trong thời gian thực và động tác chuẩn[1].

**3. Công thức chấm điểm**

Xem xét thực tế, chúng tôi quan sát được từ các thí nghiệm, chúng tôi gán trọng số khác nhau cho từng thành viên trong mảng khoảng cách và sử dụng tổng trọng để bù đắp cho sự chênh lệch thị giác. Sử dụng công thức chấm điểm được [1] đề xuất có bổ sung dữ liệu độ sâu ảnh từ (1):

D=[(] (2)

Trong đó là biểu thị một thành viên của AngleDiff, D biểu thị các tham số khoảng cách duy nhất tính từ mảng, , biểu thị các giá trị trọng số của khuỷu tay, đùi, cánh tay, và cẳng chân .

Để tính toán các giá trị trọng số, hệ thống sẽ tự động thu thập 10 cặp cử chỉ gần đây nhất được xếp loại từ cơ sở dữ liệu mà có số điểm trong khoảng ± 15% và tham khảo trung bình cộng của mỗi αi. Giá trị trọng số được tính như

(3)

Trong đó Avg1, Avg2, Avg3, Avg4 biểu thị trung bình cộng của khuỷu tay (α1 + α2), đùi (α3 + α4 ), cánh tay (α5 + α6) và cẳng chân (α7 + α8). Thao tác này sẽ gán trọng số lớn hơn tới các chi mà được đánh giá chặt chẽ hơn theo người chấm.

Sau khi chúng tôi đã nhận giá trị D, mục tiêu của chúng tôi là để D vào chia phần trăm điểm. Nhiệm vụ là tương đối đơn giản bằng cách sử dụng biến đổi tuyến tính. Chúng tôi cho phép người dùng nhập vào một giá trị chuẩn Dst và số điểm tham chiếu của nó Sst.

*Score = f*().[(Dst –D) x + Sst (4)

Trong đó *f*(αmax) biểu thị hàm giới hạn độ lệch.

Dst có thể tự thiết lập bởi người sử dụng theo nhu cầu của họ, Dst nhỏ hơn chỉ ra tiêu chuẩn cao hơn về phân loại

Sst cung cấp cho người dùng một tùy chọn để kiểm soát các điểm trong phạm vi mong muốn.

Trong trường hợp phán đoán cử chỉ một người biểu diễn thường không được phép đi chệch khỏi những cử chỉ chuẩn quá nhiều. Chúng tôi xác định một hàm giới hạn độ lệch cho phép người dùng nhập vào một giá trị ngưỡng M (độ) để hạn chế những cử chỉ độ lệch không thể chấp nhận:

f(αmax ) = 1 – (5)

Trong đó αmax biểu thị tối đa trong AngleDiff. Hàm có một số tính năng:

• f (0) = 1. Khi cử chỉ thời gian thực giống hệt với những cử chỉ tiêu chuẩn, giá trị hàm là 1.

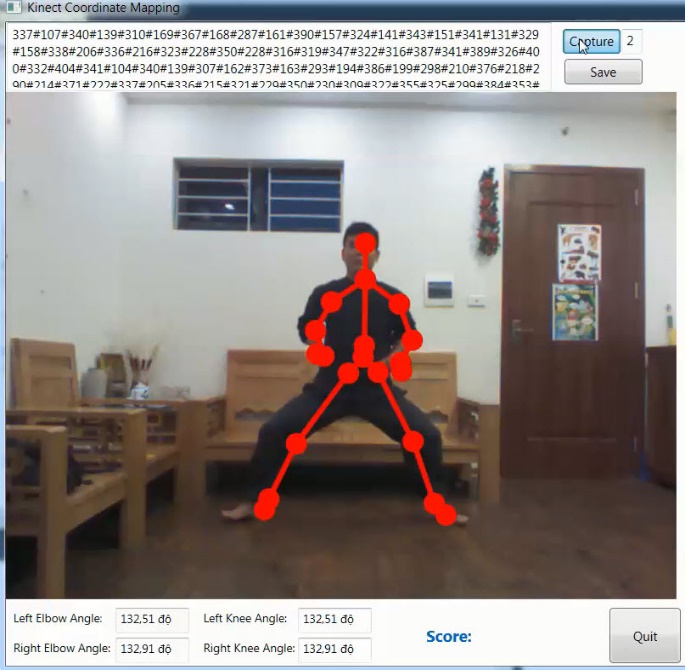
• f '(αmax ) <0, f "(αmax ) <0, khi mà góc lệch tăng tối đa, giá trị hàm giảm mạnh.

• f (M) = 0, khi góc tiếp cận lớn nhất hoặc vượt quá giá trị ngưỡng, giá trị hàm tiếp cận bằng 0 hoặc trở nên vô cực.

Với các tính năng trên, điểm số sẽ được hạn chế một cách hợp lý khi độ lệch giữa thời gian thực cử chỉ và điệu bộ tiêu chuẩn, tức là số điểm cao đòi hỏi phải thực hiện trơn tru của tất cả chân tay

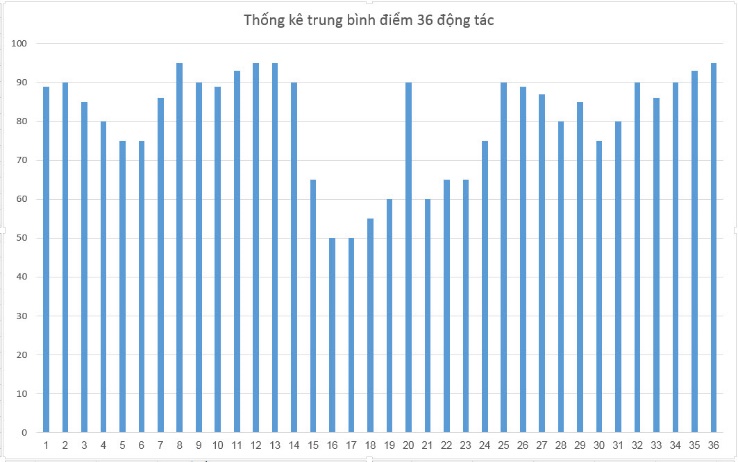
**4. Kết quả thực nghiệm**

***4.1. Kết quả chấm điểm:***



***Hình 4:*** *Thu nhận dữ liệu chuẩn từ võ sư*

Chúng tôi thu thập được 36 động tác được đưa ra bởi một võ sư mà được cho là có cùng một mức độ hiệu suất (Hình 4). Chúng tôi thiết lập các đầu vào người sử dụng Dst = 50, Sst = 80, M = 35, và 36 lớp cử chỉ bằng cách sử dụng (4). Bảng 2 cho thấy một số kết quả trong đó các động tác mà người biểu diễn quay lưng lại với camera thì điểm số không cao như các động tác trực diện. Cũng như một số động tác có độ che khuất cao thì độ chính xác cũng như điểm chấm giảm đi đáng kể.

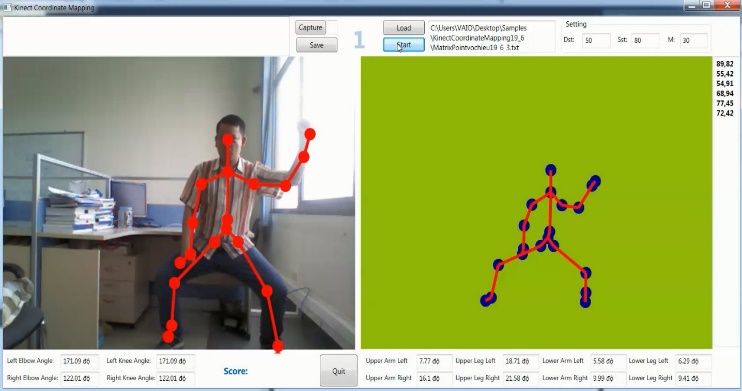
**

***Bảng 2:****Thống kê trung bình điểm chấm 36 động tác võ cổ truyền.*

***4.2. Thực hiện hệ thống chấm điểm liên tục***

Các hệ thống cử chỉ chấm điểm tĩnh được phát triển trên nền tảng Visual Studio 2013 sử dụng C# và Matlab. Người dùng sẽ được yêu cầu nhập vào các thông số trong hộp chỉnh sửa và chọn tập dữ liệu động tác chuẩn từ thư viện trước khi bắt đầu. Các động tác chuẩn sẽ lần lượt được hiển thị trên màn hình dưới dạng hình ảnh tĩnh của khung xương (Hình 5, bên tay phải)..

Người tập có khoảng thời gian 5 giây để biểu diễn động tác của mình theo đúng như trên mẫu hiện ra hoặc nếu quên thì có thể nhìn theo mẫu để biểu diễn cho đúng. Sau khi máy tính thu nhận dữ liệu của người biểu diễn thì sẽ so sánh và chấm điểm ngay, sau đó là cho hình ảnh của động tác tiếp theo cho đến hết bài. Cuối chương trình sẽ cho tổng điểm của người biểu diễn.

****

***Hình 5:*** *Giao diện chương trình chấm điểm.*

**5. Kết luận**

Việc đề xuất đưa dữ liệu độ sâu ảnh áp dụng vào công thức chấm điểm động tác tĩnh [1] của dữ liệu khung xương thu được từ camera Kinect để xây dựng phần mềm đánh giá động tác tĩnh võ cổ truyền Việt Nam đã mở đầu cho việc áp dụng đánh giá tự chấm điểm người học; Camera Kinect được phát triển bởi Microsoft với nhiều chương trình ứng dụng và được bán với giá thành rẻ từ đó người học có thể tự rèn luyện các động tác võ ban đầu và tự đánh giá mình thông qua camera Kinect góp phần vào việc bảo tồn và phát triển võ cổ truyền Việt Nam.

Từ bài báo đã xây dựng được cơ sở dữ liệu mẫu 3D để chấm điểm các bài võ cổ truyền dành cho chương trình thể dục giữa giờ được đưa vào chương trình học của các trường phổ thông từ năm 2016, từ đó có nhiều hướng mở ra cho việc đẩy mạnh tự rèn luyện và có thể đánh giá tự động qua mạng chỉ cần có camera Kinect.

Tuy nhiên vẫn còn nhiều hướng để tiếp tục hoàn thiện như tiến hành chấm điểm với chuỗi hoạt động liên tục chứ không yêu cầu người tập dừng lại như hiện nay. Người dùng có thể tương tác với hệ thống bằng cách tự thiết lập các thông số theo nhu cầu cụ thể và làm sao giải quyết được tốc độ thu nhận các động tác nhanh hơn để đáp ứng nhu cầu của bài võ thực tế là một đòi hỏi còn nhiều hướng để phát triển.

**Tài liệu tham khảo**

1. Linwan Liu, X.Wu, L.Wu, Static human gesture grading based on Kinect, 5th International congress on Image and signal processing 2012.
2. Zoe. M, Sebastian.K, Isabel.P, Joao.B, Super Miror: A Kinect interface for ballet dancers, CHI May 5-10, Austin, Texas, USA, 2012.
3. Michalis.R, Darko.K, Hugues.H, Real-Time classification of dance gesture from skeleton animation, ACM SIGGRAPH Symposium on computer animation, 2011.
4. D. Kim and J. Paik, "Gait recognition using active shape model and motion prediction”, , vol. 4, pp. 25-36,IET Computer Vision 2010
5. J. Lazar, J.H. Feng, and H. Hochheiser, “Research methods in human computer interaction,” John Wiley & Sons, 2010
6. J. Shotton, A.Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, A. Blake, “Realtime human pose recognition in parts from single depth images,” In Proc. Conf. Computer Vision and Pattern Recognition, 2011
7. Kinect Natural User Interface Overview, Microsoft Research, 2011

(BBT nhận bài: …/…/2016, phản biện xong: …/…/2016)**Thông tin về tác giả**

|  |  |
| --- | --- |
|  | Nguyễn Tường Thành:  - Tốt nghiệp ĐHBK Hà Nội năm 2002 chuyên ngành Điện tử Viễn thông;  - Năm 2008: Tốt nghiệp Thạc sỹ chuyên ngành Kỹ thuật Điện tử, Trường Đại học Giao thông vận tải Hà Nội;  - Năm 2012-2016: Nghiên cứu sinh chuyên ngành Kỹ thuật điện tử, Trường Đại học Bách khoa Hà Nội;.  - Hiện nay Nguyễn Tường Thành đang công tác tại Khoa Kỹ thuật& Công nghệ, trường ĐH Quy Nhơn;  - Điện thoại: 0914092020 |
|  | Lê Dũng:  - Năm 1998: Tốt nghiệp Thạc Sĩ chuyên ngành Điện tử Viễn thông, Trường Đại học Bách khoa Hà Nội;  - Năm 2009: Tốt nghiệp Tiến Sĩ, ngành, Trường Đại học Shibaura Nhật Bản, Công việc hiện tại: Giảng viên, Trường Đại học Bách khoa Hà Nội  - Các nghiên cứu chính là Xử lý ảnh 2D, 3D và video, nhận dạng mẫu với mạng nơron, Giao tiếp thông minh giữa người-robot, Thiết kế trên FPGA và DSP.  - Điện thoại: 0904982398. |
| thay cong | Phạm Thành Công:  - Năm 1998: Tốt nghiệp Thạc Sĩ chuyên ngành Điện tử Viễn thông, Trường Đại học Bách khoa Hà Nội;  - Năm 2010: Tốt nghiệp Tiến Sĩ, ngành Điện tử Viễn thông, Trường Đại học Bách khoa Turin, Italia;  - Công việc hiện tại: Giảng viên, Viện Điện tử Viễn thông, Trường Đại học Bách khoa Hà Nội ;  - Lĩnh vực quan tâm: Kỹ thuật siêu cao tần, Anten, Hệ thống viễn thông.  - Điện thoại: 0978808583. |
| hinh the tuyen | Nguyễn Đăng Tuyên:  -Năm 2008: Tốt nghiệp Học viện Bưu chính Viễn thông, TP HCM năm chuyên ngành Công nghệ thông tin  -Năm 2013: Tốt nghiệp Thạc sỹ chuyên ngành Công nghệ máy tính, Trường Đại học Đà Nẵng.  - Hiện nay là NCS của Viện Công nghệ Thông tin, Viện Hàn lâm khoa học công nghệ Việt Nam.  - Công việc hiện tại: Giảng viên khoa Công nghệ Thông tin, trường Đại học Quang Trung  - Điện thoại: 0914783357 |

**Thông tin cụ thể:**

1. Họ và tên: Nguyễn Tường Thành

2. Học hàm, học vị: Thạc sĩ

3. Tên cơ quan: Trường Đại học Quy Nhơn

4. Liên hệ số mobile: 0914092020, thanh1277@gmail.com