

PHÂN TÍCH DÁNG ĐI DỰA TRÊN THÔNG TIN ĐỘ SÂU

Hồ Việt Hà¹, Trần Anh Vũ², Ngô Văn Sỹ², Huỳnh Hữu Hưng³, Đặng Văn Đàng⁴

¹ Trường Cao đẳng Nghề Đà Nẵng

² Trung Tâm nghiên cứu Điện, Tin học, Tự động hóa miền Trung

³ Trường Đại học Bách khoa, Đại học Đà Nẵng

⁴ Khoa Công nghệ thông tin, Trường Đại học Cửu Long

{viethait, anhvu.tran1504, singo2016}@gmail.com, hhhung@dut.udn.vn, dangvandang@mku.edu.vn

TÓM TẮT— Ngày nay, các nghiên cứu trong lĩnh vực thị giác máy tính hỗ trợ rất nhiều cho công tác y tế, đặc biệt là các nghiên cứu phân tích dáng đi giúp quan sát bệnh nhân và có thể phát hiện sớm một số bệnh tật. Trong bài báo này chúng tôi đề xuất giải pháp nhận dạng hành vi té ngã thông qua ảnh chiều sâu thu nhận được từ thiết bị Kinect do Microsoft sản xuất. Quá trình xử lý bao gồm các bước: (1) thu nhận các đặc trưng về khung xương và đường biên cơ thể thông qua việc sử dụng Kinect; (2) tính toán các thông số đặc trưng gồm vị trí khớp và tốc độ khớp; (3) so sánh các giá trị thông số đặc trưng đó với các giá trị ngưỡng, nếu các đặc trưng đó vượt quá ngưỡng và không tiếp tục có sự biến động lớn, hành vi té ngã được giả định rằng đã xảy ra. Kết quả thực nghiệm trong môi trường thực tế và bộ cơ sở dữ liệu MOCAP về khung xương cho thấy giải pháp đề xuất mang lại hiệu quả hứa hẹn.

Từ khóa— Dáng đi, té ngã, độ sâu, Kinect.

I. GIỚI THIỆU

Theo các nghiên cứu về y tế hiện nay, té ngã là nguyên nhân chủ yếu gây ra các chấn thương ở người cao tuổi và trẻ em. Kết quả điều tra tai nạn thương tích ở Việt Nam năm 2010 cho thấy rằng, việc té ngã là nguyên nhân đứng thứ 2 dẫn đến các trường hợp tử vong. Vì vậy, việc nghiên cứu theo dõi dáng đi của con người để kịp thời phát hiện té ngã và phát ra cảnh báo cho những người phụ trách chăm sóc trở thành một nhu cầu hết sức cấp thiết và là một trong các vấn đề được quan tâm hàng đầu những năm gần đây. Các nghiên cứu này thường đi theo hai hướng chính: hướng thứ nhất sử dụng các thiết bị ngoại vi để cảnh báo và hướng thứ hai là sử dụng hệ thống camera để quan sát, phân tích dáng đi, phát hiện té ngã.

Hành vi té ngã của con người có thể được mô tả là sự thay đổi vị trí một cách nhanh chóng từ tư thế đứng hay ngồi sang tư thế mất thăng bằng, gần như là nằm. Té ngã là một hành động không được kiểm soát. Theo [1], hành vi té ngã có thể được chia thành 4 giai đoạn. Trong những năm gần đây, có nhiều nghiên cứu được đề xuất có thể được phân thành hai loại, tùy thuộc vào giai đoạn của quá trình té ngã mà nghiên cứu tập trung vào, hai giai đoạn được tập trung chủ yếu là: giai đoạn đang ngã và giai đoạn đã ngã.

Nghiên cứu [9] sử dụng một hệ thống nhiều camera để giám sát đối tượng nhận dạng; thí nghiệm trong nghiên cứu được thực hiện trên 300 chuỗi video tổng hợp từ nguồn dữ liệu CASIA-B được ghi trên 12 người khỏe mạnh với các hướng đi tự do theo 6 góc khác nhau (36°; 54°; 72°; 90°; 108°; 126°), từ đó có thể xây dựng được các thư viện về dáng đi của con người phục vụ cho nghiên cứu phát hiện té ngã. G. Debart và P. Karsmakers [10] sử dụng một camera đơn để tiến hành quan sát nhận biết các dấu hiệu có nguy cơ té ngã của bệnh nhân, sau đó sử dụng mô hình Support Vector Machine để phân loại đối tượng thành nhóm bình thường và nhóm bị té ngã.

Trong nghiên cứu [3], tác giả kết hợp Kinect và máy đo gia tốc để phát hiện té ngã. Để phát hiện té ngã, nhóm tác giả của nghiên cứu này đã phát hiện các đối tượng tiền cảnh (ở đây là đối tượng con người) bằng cách sử dụng nhiều kỹ thuật dựa trên ảnh chiều sâu. Trọng tâm của người và khoảng cách giữa trọng tâm của người với Kinect được tính toán. Việc xác định té ngã sẽ được thực hiện dựa trên các thông tin về trọng tâm và một số các quy tắc đã được định trước. Hiệu suất đem lại của phương pháp trong nghiên cứu này là khá khả quan. Tuy nhiên, nghiên cứu này lại không phân biệt được, đâu là té ngã và đâu là chủ động nằm hay ngồi trên sàn nhà.

Ưu điểm của hệ thống sử dụng camera là không cần có bất cứ thiết bị ngoại vi nào trên người người sử dụng, đồng thời thông tin về các chuyển động có thể được xác định chi tiết và kịp thời hơn. Với việc Kinect vừa có thể thu nhận ảnh màu như camera thông thường, vừa có thể thu nhận thông tin ảnh chiều sâu, chúng tôi dựa vào những dữ liệu này để theo vết khung xương của đối tượng. Do đó, Kinect được dùng trong nghiên cứu này để phát hiện té ngã.

Nội dung tiếp theo được trình bày như sau: phần II đề cập chi tiết giải pháp nhận dạng té ngã đề xuất; kết quả thực nghiệm được trình bày trong phần III; phần IV so sánh giải pháp đề xuất với một vài phương pháp đã có và cuối cùng phần V là kết luận cũng như đánh giá giải pháp.

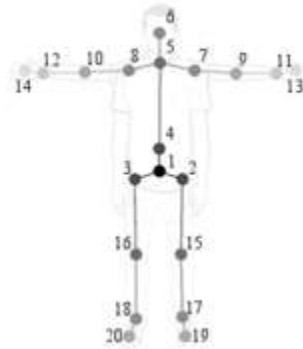
II. GIẢI PHÁP ĐỀ XUẤT

Trong nghiên cứu này, chúng tôi định nghĩa té ngã là quá trình mà vận tốc trung bình các khớp của cơ thể đạt đến mức lớn hơn một ngưỡng định trước. Điều này được đưa ra dựa trên thực tế rằng các hành vi bình thường ở người như ngồi, nằm diễn ra với một tốc độ chậm rãi (so với ngưỡng đưa ra). Dữ liệu vào của hệ thống là các khớp khung xương được thu nhận qua Kinect, dựa trên các khớp xương này, chúng tôi đưa ra 2 thuật toán để phát hiện té ngã. Thuật

toán đầu tiên sử dụng dữ liệu về vị trí của các khớp. Thuật toán thứ hai tính vận tốc của các khớp theo hướng vuông góc với mặt phẳng sàn, vận tốc trung bình được tính trên tất cả các khớp đã chọn và nhiều khung hình.

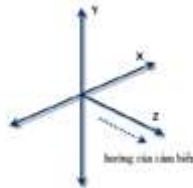
A. Đặc trưng khung xương

Hình ảnh khung xương Kinect nhận diện được bao gồm 20 vị trí khớp thể hiện ở Hình 1.



Hình 1. Vị trí các khớp trong khung xương Kinect

Tọa độ 3D của mỗi khớp được Kinect biểu diễn bởi 3 thông số x , y , z lần lượt tương ứng với chiều ngang, chiều dọc và chiều sâu như trong Hình 2.



Hình 2. Hệ tọa độ không gian của Kinect

Bài báo này sử dụng tất cả 20 khớp xương do Kinect cung cấp để đánh giá khả năng xảy ra té ngã.

B. Thuật toán tính vị trí khớp

Kinect SDK xử lý và cung cấp cho chúng ta dữ liệu với tốc độ 30 khung hình mỗi giây. Mỗi khung hình sẽ được xử lý bởi giải pháp đề xuất và xác định trong hình đó có xuất hiện việc té ngã hay không. Đối với các khớp được theo dõi trên khung xương, sẽ có 4 mẫu dữ liệu tương ứng được thu thập từ Kinect: các tọa độ x , y , z của khớp và thông tin theo dõi của khớp đó. Các tọa độ x , y , z là tọa độ Descartes và thiết bị Kinect được đặt ở gốc tọa độ. Ngoài thông tin của các khớp, các thông tin của mặt phẳng sàn nhà sẽ được thu lại. Kinect sẽ cung cấp các thông tin của mặt phẳng sàn nhà thông qua phương trình

$$ax + by + cz + d = 0 \quad (1)$$

trong đó a , b và c lần lượt là các tọa độ của vector pháp tuyến của mặt phẳng sàn nhà, d được xác định bởi công thức

$$d = -ax_0 - by_0 - cz_0 \quad (2)$$

với x_0 , y_0 , z_0 là tọa độ một điểm xác định nằm trên mặt phẳng sàn nhà.

Từ phương trình mặt phẳng sàn, ta có thể tính được khoảng cách từ mỗi khớp có tọa độ (x, y, z) đến mặt phẳng sàn nhà bằng công thức

$$D = \frac{|ax+by+cz+d|}{\sqrt{a^2+b^2+c^2}} \quad (3)$$



Hình 3. Xác định vị trí khớp và khoảng cách từ các khớp đến mặt phẳng sàn nhà

Thuật toán sẽ chỉ tính toán khoảng cách của các khớp xương được chọn theo dõi bởi Kinect. Nếu từng khoảng cách từ các khớp xương này đến mặt phẳng sàn nhà nhỏ hơn các giá trị ngưỡng cho trước thì thuật toán sẽ kết luận đối tượng đang theo dõi bị té ngã.



Hình 4. Khoảng cách từ các khớp đến mặt phẳng sàn nhà dưới ngưỡng

C. Thuật toán tính vận tốc khớp

Kinect cung cấp thông tin khung hình với tốc độ 30 khung hình mỗi giây. Với mỗi khung hình, thông tin về thời gian và vị trí của mỗi khớp ứng với khung hình đó được lưu giữ. Chúng tôi giả định rằng Kinect được đặt tại một vị trí cố định và có góc độ không đổi so với mặt phẳng sàn trong suốt quá trình quan sát.

Thay vì sử dụng phương trình mặt phẳng sàn được cung cấp bởi Kinect, thuật toán này sẽ tính toán một mặt phẳng sàn riêng. Theo nghiên cứu [4] mặt phẳng sàn nhà sẽ được xác định theo phương trình

$$ax + by + cz + d = 0 \quad (4)$$

với $a = 0$, $b = \cos(\theta)$, $c = \sin(\theta)$, và $d = 3$, trong đó θ là góc tạo bởi Kinect và mặt phẳng sàn.

Khoảng cách từ mỗi khớp đến mặt phẳng sàn nhà được xác định theo công thức

$$D = \frac{|ax+by+cz+d|}{\sqrt{a^2+b^2+c^2}} \quad (5)$$

Với hai khung hình liên tiếp i và $i+1$, vận tốc của khớp j tại thời điểm đó được xác định dựa trên phương trình

$$v_j = 1000 \frac{d_i - d_{i+1}}{t_{i+1} - t_i} \quad (6)$$

với d_i là khoảng cách từ khớp j đến mặt phẳng sàn nhà xác định tại thời gian t_i của khung hình thứ i , d_{i+1} là khoảng cách từ khớp j đến mặt phẳng sàn nhà xác định tại thời gian t_{i+1} của khung hình thứ $i+1$, v_j là vận tốc khớp cần tính.

Vận tốc trung bình của khớp j trong n khung hình liên tiếp được Kinect thu nhận được tính theo công thức

$$v_m = \frac{1000}{n-1} \sum_{j=1}^{n-1} \frac{d_i - d_{i+1}}{t_{i+1} - t_i} \quad (7)$$

Cuối cùng, ta tính được vận tốc khớp trung bình từ các vận tốc trung bình thu được của 20 khớp như sau:

$$v_{full} = \frac{1}{20} \sum_{j=1}^{20} v_{m,j} \quad (3.5)$$

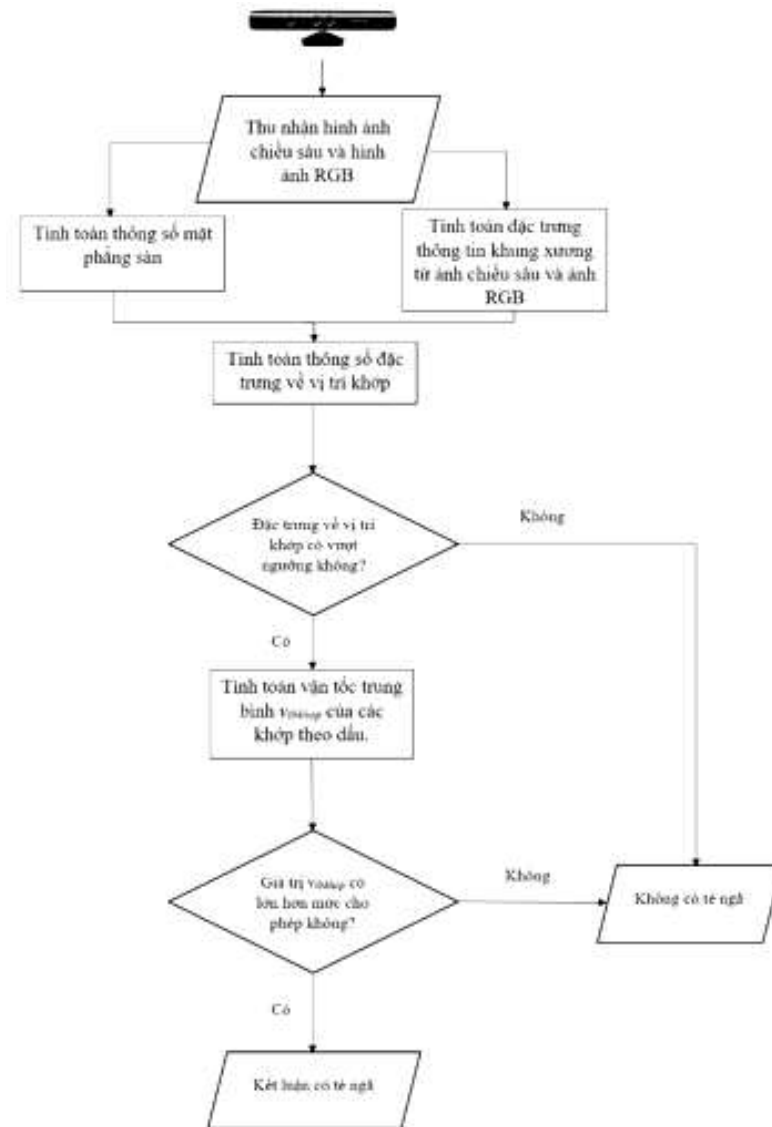
Nếu vận tốc trung bình của các khớp v_{full} lớn hơn ngưỡng, thì việc té ngã được xác định.

D. Quy trình nhận dạng hành vi té ngã

Hệ thống phát hiện hành vi té ngã dựa trên thông tin đặc trưng khung xương có quy trình thực hiện lần lượt như trong hình 5.

- **Bước 1:** Từ hình ảnh màu và hình ảnh chiều sâu thu nhận được qua Kinect, hệ thống sẽ tính toán các giá trị của mặt phẳng sàn nhà và thu nhận các khớp xương của khung xương đối tượng.
- **Bước 2:** Từ phương trình đặc trưng mặt phẳng sàn nhà và các khung xương trích xuất được, các giá trị khoảng cách D của các khớp đến mặt phẳng sàn nhà được xác định. Nếu D thay đổi bất thường thì tiếp tục tính toán vận tốc trung bình v_m của sự thay đổi đó.
- **Bước 3:** Tính toán vận tốc trung bình của một khớp dựa trên n khung hình liên tiếp, sau đó tính vận tốc trung bình của toàn bộ các khớp xương đang được theo dõi. So sánh kết quả thu được với giá trị vận tốc ngưỡng. Nếu v_{full} này lớn hơn ngưỡng thì có thể kết luận đó là hành vi té ngã.
- **Bước 4:** Kết luận về dáng đi của đối tượng đang được theo dõi dựa trên thông tin đặc trưng thu được ở trên.

Trong quá trình thực thi, hệ thống liên tục theo dõi tư thế của người trong tầm nhìn camera, đồng thời tính toán vận tốc trung bình các khớp trong một số khung hình liên tiếp gần nhất. Hệ thống dừng theo dõi và đưa ra cảnh báo khi vận tốc đó đạt giá trị vượt trên một ngưỡng định trước với giả định rằng việc té ngã đã xảy ra.



Hình 5. Quy trình phát hiện té ngã dựa trên đặc trưng khung xương

III. KẾT QUẢ THỰC NGHIỆM

Hệ thống thử nghiệm được xây dựng trên nền tảng ngôn ngữ lập trình C# và thực hiện với sự hỗ trợ của thiết bị Kinect 1.0. Việc thực nghiệm được tiến hành trong môi trường phòng ở bình thường. Hệ thống thử nghiệm đã khắc phục được tình trạng nhầm lẫn, phân biệt giữa hành vi bình thường của con người và hành vi té ngã. Khi có sự thay đổi vị trí các khớp xương nhưng không đi kèm với tốc độ thay đổi lớn, thì hệ thống không đánh giá đó là hành vi té ngã.

Hệ thống được thử nghiệm với 4 góc nhìn khác nhau: góc nhìn từ phía trước, từ phía sau, từ bên trái và từ bên phải. Với các góc nhìn này, khi phát hiện có sự thay đổi về vị trí các khớp xương, hệ thống sẽ kiểm tra xem sự thay đổi này có đi kèm với một tốc độ thay đổi lớn hay không, nếu có, hệ thống sẽ kết luận là có hành vi té ngã, ngược lại, đó là một hành vi bình thường của đối tượng. Vận tốc trung bình của cơ thể được tính trong 10 khung hình liên tiếp, tương ứng với xấp xỉ 1/3 giây (tốc độ ghi hình của Kinect xấp xỉ 30 khung hình mỗi giây). Vận tốc ngưỡng được thiết lập trong hệ thống là 2.5 m/s (gần gấp đôi vận tốc đi bộ ở người bình thường).



Hình 6. Phát hiện hành vi té ngã góc nhìn từ phía trước



Hình 7. Phát hiện hành vi té ngã góc nhìn từ phía sau



Hình 8. Phát hiện hành vi té ngã góc nhìn từ bên trái



Hình 9. Phát hiện hành vi té ngã góc nhìn từ bên phải

Để thử nghiệm hệ thống này, bài báo đã thử nghiệm khác nhiều lần các hành vi và các góc nhìn khác nhau. Kết quả thu được cho thấy hệ thống có khả năng phân biệt được hành vi té ngã với các hành vi bình thường như nằm và ngồi của đối tượng.

Bên cạnh các kết quả trực tiếp thu được, chúng tôi còn sử dụng cơ sở dữ liệu khung xương MOCAP [15] để kiểm thử hệ thống và thu được các kết quả như sau.

Bảng 1. Kết quả thử nghiệm với dữ liệu MOCAP

Tư thế kiểm thử	Số dữ liệu thử nghiệm	Tỉ lệ nhận dạng chính xác
Nằm	127	100%
Ngồi	118	100%
Ngã (góc nhìn trước)	136	93.4%
Ngã (góc nhìn sau)	130	95.4%
Ngã (góc nhìn trái)	125	96%
Ngã (góc nhìn phải)	158	94.3%

Bảng 2. So sánh giữa hệ thống triển khai với hệ thống đã có

	Hệ thống triển khai	Kepski, M. and B. Kwolek [3]
Phân biệt nằm/ ngồi với té ngã	X	
Không nhận dạng người không hoạt động là té ngã	X	X
Nhận dạng được nhiều góc nhìn	X	X
Nhận dạng khi bị khuất một phần nhỏ cơ thể	X	

Một số dạng té ngã bị hệ thống bỏ qua (nhận dạng sai) là do vận tốc khi ngã lớn, cùng với tư thế trong quá trình ngã gây khó khăn cho việc xác định vị trí các khớp cơ thể, lúc này vị trí mới của các khớp được nội suy từ quỹ đạo chuyển động trước đó. Điều này khiến cho hệ thống không phát hiện được sự thay đổi thực sự về vị trí các khớp, dẫn đến việc vận tốc trung bình của cơ thể không có nhiều biến đổi. Hạn chế này có thể được giải quyết khi sử dụng một hệ thống nhiều camera nhằm đảm bảo luôn tồn tại ít nhất một camera xác định được đúng vị trí khớp tại một thời điểm bất kỳ.

IV. KẾT LUẬN

Bài báo này trình bày phương pháp nhận dạng té ngã dựa trên dữ liệu khung xương được xác định trên ảnh độ sâu của Microsoft Kinect. Kết quả thực nghiệm cho thấy giải pháp đề xuất mang lại hiệu quả hứa hẹn trong bối cảnh nhà ở. Ngoài ra, giải pháp đề xuất có thể được tích hợp vào hệ thống giám sát sử dụng Kinect, chăm sóc sức khỏe tự động, là công cụ để theo dõi và phân tích dáng đi của con người để phát hiện hành vi té ngã của người được chăm sóc và đưa ra cảnh báo kịp thời khi có sự cố. Điểm hạn chế của giải pháp là hệ thống chưa đánh giá chính xác được một số hành vi có tốc độ thay đổi cao như nhảy nhót, chạy tốc độ cao.

Trong tương lai, giải pháp đề xuất có thể được cải tiến bằng cách kết hợp các dữ liệu khung xương và dữ liệu đường biên của cơ thể để xây dựng mô hình 3D của đối tượng cần theo dõi, từ đó có thể thu được những kết quả chính xác nhất phục vụ cho việc đánh giá hành vi cũng như chăm sóc sức khỏe có liên quan đến dáng đi của con người.

LỜI CẢM ƠN

Nghiên cứu này do Dự án UK – ASEAN Research Hub, Viện Nghiên cứu và Đào tạo Việt - Anh, Đại học Đà Nẵng tài trợ (07/HD-UARH).

TÀI LIỆU THAM KHẢO

- [1] N, N. , et al. , “A proposal for the classification and evaluation of fall detectors”, Biomed Eng Res (IRBM), 29(6): pp.340-349, 2008.
- [2] Rougier, C. , et al., “Fall Detection from Depth Map Video Sequences”, in ICOST, pp. 121-128, 2011.
- [3] Kepski, M. and B. Kwolek, “Fall Detection on Embedded Platform Using Kinect and Wireless Accelerometer”, in ICCHP2012, pp.407-414, 2012.
- [4] Kawatsu, c., I. Li, and C. Chung, “Development of a Fall Detection System with Microsoft Kinect”, in Robot Intelligence Technology and Applications, pp.623-630, 2012.
- [5] Zoran Zivkovic, “Improved Adaptive Gaussian Mixture Model for Background Subtraction”, In Proc. ICPR, pp.28-31, 2004.
- [6] Muhammad Jamil Khan and Hafiz Adnan Habib, “Video Analytic for Fall Detection from Shape Features and Motion Gradients”, WCECS 2009, San Francisco, USA, pp.20-22, 2009.
- [7] Chen, M., Huang, B., & Xu, Y., “Human Abnormal Gait Modeling via Hidden Markov Model,” In Proceedings of International Conference on Information Acquisition 2007 (ICIA'07), IEEE, pp. 517-522, 2007.
- [8] Greene, B. R., Donovan, A. O., Romero-Ortuno, R., Cogan, L., Scanail, C. N., & Kenny, R. A., “Quantitative Falls Risk Assessment using the Timed Up and Go Test,” IEEE Transactions on Biomedical Engineering, 57(12), pp. 2918-2926, 2010.
- [9] Goffredo, M., Bouchrika, I., Carter, J. N., & Nixon, M. S., “Performance analysis for gait in camera networks”, In Proceedings of the 1st ACM workshop on Analysis and retrieval of events/actions and workflows in video streams, ACM, pp. 73-80, 2008.
- [10] Debard, G., Karsmakers, P., Deschodt, M., Vlaeyen, E., Van den Bergh, J., Dejaeger, E., ... & Vanrumste, B., “Camera based fall detection using multiple features validated with real life video”, In Workshop Proceedings of the 7th International Conference on Intelligent Environments, vol. 10, pp. 441-450, 2011.
- [11] Clark, R. A., Pua, Y. H., Bryant, A. L., & Hunt, M. A., “Validity of the Microsoft Kinect for Providing Lateral Trunk Lean Feedback during Gait Retraining,” Gait & posture, Elsevier, 38(4), pp. 1064-1066, 2013.
- [12] Nguyen, H. A., Auvinet, E., Mignotte, M., de Guise, J. A., & Meunier, J., “Analyzing Gait Pathologies using a Depth Camera,” In Proceedings of Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBC), IEEE, pp. 4835-4838, 2012.
- [13] Stone, E. E., & Skubic, M., “Passive in-home Measurement of Stride-to-Stride Gait Variability Comparing Vision and Kinect Sensing,” 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, Boston, Massachusetts, pp. 6491-6494, 2011.
- [14] Thi-Thanh-Hai Tran, Thi-Lan Le. An analysis on human fall detection using skeleton from Microsoft Kinect. IEEE Bordeaux, France, pp. 484 – 489, 2014.
- [15] <http://mocap.cs.cmu.edu/info.php>.

GAIT ANALYSIS USING DEPTH INFORMATION

Ho Viet Ha, Tran Anh Vu, Ngo Van Sy, Huynh Huu Hung, Dang Van Dang

ABSTRACT – Nowadays, the research in the field of computer vision supports a lot for public health, especially the gait analysis help observe the patient and can detect a disease early. In this paper we propose behavioral recognition solution falls through the depth image acquired by the Microsoft Kinect device. Treatment process comprising the steps of: (1) acquire the characteristics of the boundary frame and body through the use of Kinect; (2) calculate the characteristic parameters including location and speed joints joints; (3) comparing the characteristic parameter value with the threshold value which, if it exceeds the threshold characteristic and not continue to have large fluctuations, the behavior falls is assumed that happened. Experimental results in real environment and the database of the skeleton suggests MOCAP proposed solutions bring promising results.