Occupant behavior modeling based on migration registration technique

J W Dziedzic[1], V Novakovic[1]

[1]Norwegian University of Science and Technology,
Department of Energy and Process Engineering,
Kolbjørn Hejes v 1B

jakub.w.dziedzic@ntnu.no

Abstract. The building occupant energy related behavior modelling is a key factor for building performance simulation that supports realistic forecasting of energy consumption. The current state of development of this topic delivers partial knowledge about occupant's behavior, but information about transient thermal comfort of each occupant is still, beyond the reach. Access to such a data would visualize outcome of the proposed HVAC design. With a use of proper measurement technique, it is possible to develop virtual bot that will recreate typical energy related behavior scenarios inside designed building and gather consistent information regarding thermal properties of the indoor environment. Such a tool will display potential faults of the design regarding occupants thermal comfort and point out potential paths of adjustment during design phase

1. Introduction

Re-creation of a building's occupant behavior has always been a challenge. The main reasons are the diversity of personal preferences among building users and various acceptance conditions. The above factors make modeling of occupant's behavior a troublesome subject. Since there are no significant input data, they have to be replaced by individual behavior approximations. Behavior models can be applied to all human-related scientific topics. They may even be used as a tool for monitoring and forecasting of energy consumption inside buildings. Several methods aiming at description of energy-related human behavior were developed in the past [1] [2]. They opened a new chapter in the research on indoor environment quality since they allowed evaluating sensation of comfort on a fundamental level. It was an essential step towards improvement of indoor air quality (IAQ) control and optimization of energy consumption for heating, ventilation and air conditioning (HVAC) systems.

The next step should lead towards personalized comfort detection. Reaching this goal will open a new space for automation in buildings. Without synergy of accuracy and reliability, it is impossible to reach the level of personalized simulations that could support preferential HVAC control. The current state-of-the-art instruments enable detecting occupants inside offices/zones with a low error margin [3]. It was possible to distinguish between two types of office/zone users: average daily users and guests. This simulation approach ensures accurate results, but it can only define behavior of occupant groups and does not offer a precise picture for a personalized indoor air model. This obstacle concerns both office and residential buildings. Occupants are described as a mass. It is difficult, within this solution, to distinguish individual actions of an occupant from occupants group. Insufficient accuracy has been emphasized in Yan's general review of occupant behavior modeling in buildings [4]. The results of his

research indicate that if higher simulation accuracy is to be achieved, it must be connected with detection of occupants' personal needs.

To improve energy simulation accuracy, input information about occupants behavior has to be more accurate/transparent [5]. With registration of each occupant's actions, it is possible to develop a pattern of their behavior. Organized chronologically, it will tell a "story" of their thermal sensation. Such data will generate a range of possible interpretations concerning an occupant's thermal needs. The collected feedback can later be interpreted as a habit or standard practice of each occupant. In order to reach the required level of detail in description, there is a need to improve the measurement accuracy. The collected data has to allow distinguishing each occupant's position and action in time. At the same time, it may not go as far as to cross the important boundary of recording data that can be considered sensitive from occupants' point of view. It has to provide sufficiently meaningful information so that all the future recordings could be attributed to a given person (for example Occupant A). The collected data has to provide specific metadata that enables interpretation of all the occupant's activities.

If all of the above requirements are fulfilled, the obtained data will allow taking the next step in development for personalized control for IAQ. With direct identification, it will be possible to build models that will support identification of occupants' needs based on actual thermal conditions. With this holistic approach, the accuracy of building performance simulation (BPS) results may reach the acceptable level. It may minimize the difference between the simulated and measured energy consumption of the building. This goal may be achieved with the use of the transient occupant behavior model. The first step towards its goal should aim at discovering patterns of occupants' movements that could later on be used to create simple yet robust movement model.

Improvement of the gathered data quality requires selection of a new measurement device or/and technique. To make them more reliable, measurements have to be conducted with the use of an in-situ method. This means that all solutions that require using a device, such as a wristband or smartphone, have to be excluded from this study as potentially leading to measurement errors with reference to some occupants. The errors may occur if some occupants lose their device or leave it in the monitored zone. Similar effects of under/overestimation of some occupants may appear if the system counts door opening as input data. It is hard to link the number of times when the door was opened to the number of occupants using it. Moreover, this solution give us no information on the direction of migration.

Other solutions – like infra-red (PIR) detectors – provide information on crossing its monitoring point. This measurement method can sometimes produce results that come close to the description of occupants' activity. However, this method does not allow capturing anticipated information. Using these detectors produces binary output. It can be enough for the lighting control or presence detection but the accuracy and resolution are not significant enough to support the development of personal thermal comfort sensing. The method has a great potential when it comes to the detection of CO_2 concentration. Unfortunately, due to the response delay in this measurement technique, it can only be used for areas, like an auditorium, where occupants stay for a longer period [6]. For residential solutions, such measurements can deliver only brief information regarding the occupancy state. Occupants activity and their degree of freedom is higher in residential buildings.

In order to obtain higher quality data, it seems that it is necessary is use video recording data. Supported by Kalman filtering or pre-trained model, it can map the paths of movement. The registered movement will be only a projection on an observed surface. The absence of perspective from the recording can be re-calculated. Unfortunately, applicability of video recording is limited due to ethical issues. It is most unlikely to be used in monitoring of residential buildings, but it can be used to track movement in wide, open areas, such as halls or lobbies. The only other solution for the measurement of occupant behavior that bypasses all of the obstacles is depth registration camera. Collected data do not allow direct identification of the monitored occupant, but it can provide information about the movement of each limb. Additionally, it can detect the outline of the observed body and thus identify individual persons. The main issue related to this type of observed depth recording is its limited range. The range boundaries are defined by the limitation of the power of signal amplifier and the price of the detector. There are many different producers currently on the market who are focusing on manufacturing this

kind of devices. Most of them find applicability in the automotive, personal aviation (drones) and entertainment sectors. Previous investigation into these measurement technique has shown promising results, and that is why using it will be explored further. [7] [8]

2. Aim

The aim of the presented work is to investigate the use of data collected in previous studies. The collected data offers high-resolution information, and its potential has not been fully used. With high-quality data, it is possible to produce an occupant behavior transition model that accurately recreates their movement. The developed model will become a foundation for building an occupant behavior simulator for model predictive control or annual simulations during the design phase. However, it is necessary to discover patterns of transition to reach this goal.

It will have to be seen if the already available data is sufficient to build a behavioral model. The first step towards development of a building occupant transient agent-based model (BOT-ABM) has to focus on observation regarding occupants' movement itself. The goal is to develop a self-sufficient model that will recreate pathways of transitions inside the investigated area. In the current state of development, drivers of occupants' actions regarding their transitions will be not investigated.

3. Methods

3.1. Measurements

Measurement data will be obtained with the use of a depth registration camera. For the purposes of this research, the Microsoft Xbox One Kinect device will be use. The choice of the measurement device was determined by the desired quality of the data to be collected, the sampling time, price and availability in the market. The selected device is capable of registering up to six people at one time and has the detection distance limitation of five meters. Regardless of these obstacles, this type of measurement can be successfully used in close spaces like corridors, entrances, small offices or residential buildings. Access to the full measurement device potential was obtained from the software development kit published by its manufacturer. Collection of this data is done with the use of a grid projected by a beam in an infra-red spectrum (to avoid discomfort of the observed occupants). Once someone crosses the observation spectrum, his or her transition/activity is registered. The obtained data is a skeleton model (SM) that has representation of twenty-five common points in three dimensions. Each SM is connected to an identification number, from one to six, by a semi-random procedure. During measurement, the device registers movements of each joint with the frequency up to thirty frames per second. A more detailed description of the measurement technique may be found in manufacturer webpage [9].

The whole recording was conducted with the use of self-developed software in the Matlab environment. The choice of the programming language was based on the availability of the existing library, the available third-party support and personal preference. The registered data will recreate paths of movement taken by each monitored occupant. Two Kinect devices were used during the measurement procedure. To make data more coherent, it was necessary to establish one common global coordination system .One device has to be selected as a 'master' device, and rest as a 'salves'. All recordings from the 'slave' devices has to be recalculated via referential 'master; device. It was done by measuring distances between both devices and setting up a translation vector to recalculate readings. Streamed information from SM provides three-dimensional data. The devices were placed so as to cover the whole office area, from one end to the other.

It was detected that there were certain specific points within the monitored area that were more actively used. All of these points on this surface can be treated as nodes of communication or points of interest (POI). POIs have to be understood as being limited by boundary spaces/zones which can be defined as points. POI location was detected with the use of the Pareto filter. While "typical" enterers were exposed, it could be used to generate a hidden Markov chain model. This part of study regarding occupancy modeling was put aside for future investigation. At the present current stage, movement inside the measurement zone and discovery of its patterns was treated as a priority.

3.2. Raw data

The recording process produces a vast amount of data. Each minute of recording generates an 1800x457 matrix. Each row of the matrix represents one frame, 30 frames per second multiplied by 60 seconds give 1800 rows. In one exposition, the Kinect device can capture up to six occupants standing within the device's range. Each registered occupant provides information about the movement of 25 points in three dimensions, 75 columns in the registration matrix. Based on this information, the recording device has to reserve recording space for six people – six multiplied by 75 gives 450.The last seven columns are reserved for the clock data that gives information about the year, month, day, hour, minute, second and frame number.

The software built into the device is unable to distinguish occupants directly, and it does not link any observed occupant with the same SM number. After it loses track of a specific person, it may link it to the same SM, but the odds of this happening are one to six. The process of connecting the SM with an observed human is semi-random. It links an occupant with the next available SM number. Once the software reaches the SM number six, it resets numbering and starts from SM one. That is why it is necessary to record all the possible SMs simultaneously. During the registration procedure, each of device's clocks was synchronized. Recorded data packages provide an overwhelming amount of data, far too much to process it manually, so it was necessary to develop software to do it automatically. It is essential to develop filtering and linking software that will transform the acquired data into a suitable form enabling its correct interpretation

3.3. Processing

The first step of data processing is initialized during recording. To save the hard drive space and improve performance of the recording software, minimization of the file size was required. To discard rows with zero activity, all empty rows have to be deleted before saving the data on the hard drive. The developed software was checking the sum of the first 450 columns from a row and deleted it if it was equal to zero (no movement detected), so that it would not be included in the saved matrix. Because the reporting (saving) protocol is resolved every minute, most of the saved files are entirely blank, due to the absence of movement inside the observed field. If there is any detected action, it immediately provides a significant data stream. It makes the files with recorded activity stand out, regarding its file size, from the rest of the files, which are blank. Simple filtering of files by their size allow quick determination which file has data with registered activities.

Afterwards, the data is successfully connected and calculated by the transition vector to one global coordinate system. The collected records require extra processing steps to identify whether there exists any POIs. The merged data still requires connecting pathways because of the existing discontinuities caused by the following two factors:

•   The devices lose track of an occupant;
•   The second source of discontinuities in the recorded data is dependent on some characteristics of the device used in the experiment.

Each device binds SM data by itself, with no communication between them. It is highly unlikely that both devices will link a monitored occupant to the same identification number of SM. To connect detached paths, it is necessary to sort records by time and calculate the velocity of each step. The next step requires investigation of the SM body ratio.

Each SM provides a complete set of information about the movement of each monitored limb. It can be used to distinguish each occupant regardless of the connected SM identification number. Registered occupants may differ from each other regarding their body dimensions, height, width etc. They can be distinguished by their body ratio factor. For example, the value of height divided by the horizontal distance between shoulders. The choice of the representative ratio or ratios will be based on its reading stability and transparency. If this layer of filtering does not provide a reliable sorting criterion, limb data can be applied to use its swing (pendulum) movement analysis. Each limb has its own specific harmonic

move during walking. Observation of its repeatability can also provide some "marker" information. Frequency and deviation angle of limb movement can provide significant support in the personal \\\identification process.

3.4. Moment knowledge
Post-process data allows displaying all chosen pathways separately. All of the pathways were projected
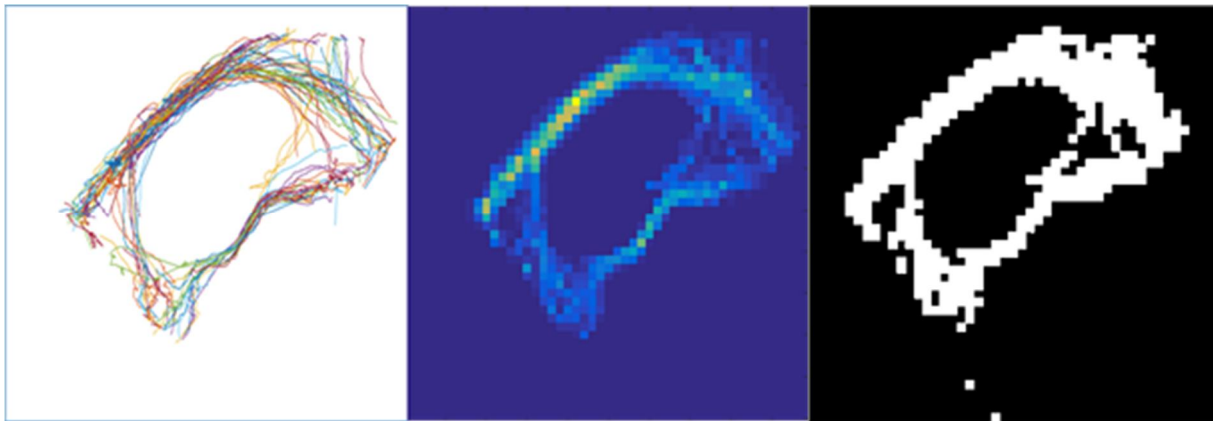

Figure 1. Pathway registration, the Heat map of space usage, transition outline plots

on a heat map to discover the frequently used tracks. Displaying these results allows swift identification of the monitored area layout (Figure 1). Based on ground truth, the uncrossed area represents the position of the table and chairs. It shows that the transition model has to take into account all possible connotations of the layout of the building. It is impossible to achieve it with just an overview of the transition and that is why the next part of the analysis has to focus on each separate step of the transition. As has already been mentioned, all monitored pathways have to be plotted on a normalized shape, where it is possible to discover patterns of transition from point A to B. Conceptual analyses of the movement were presented in Figure 2. It shows the possibilities of transition from point 1 to point 2. Transition steps and the angularity of consecutive steps are unequal. The probability of the step selection was illustrated by a gradient cloud, which represents the density of the selected steps from the starting to the end position. To understand the whole transition phenomenon, all of the movement transition has to be observed separately..
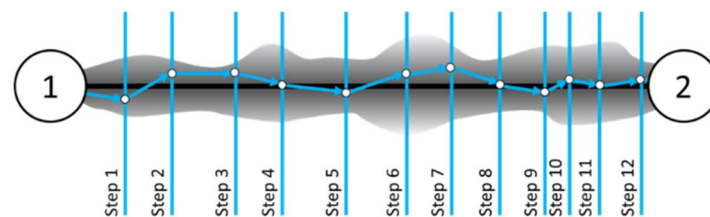

Figure 2. Movement pathway distribution illustration

3.5. Movement model
Investigation into the transition patterns requires to include all of the recoding inputs in the scope. The most interesting observation, regarding the model development potential, was the distribution of step lengths and its angularity, which is understood as an angle between the previous step and the next step turn. The distribution of the step lengths and angularity was displayed in histograms (Figure 3). Both of the graphs show that each value is characterized by natural distribution in a limited range. It demonstrates the existence of actual patterns that can be explored further. The displayed data shows a

Gaussian semi-bell-shaped curve. Both distributions were investigated, and distribution data, such as range and peek, were used to develop a movement.
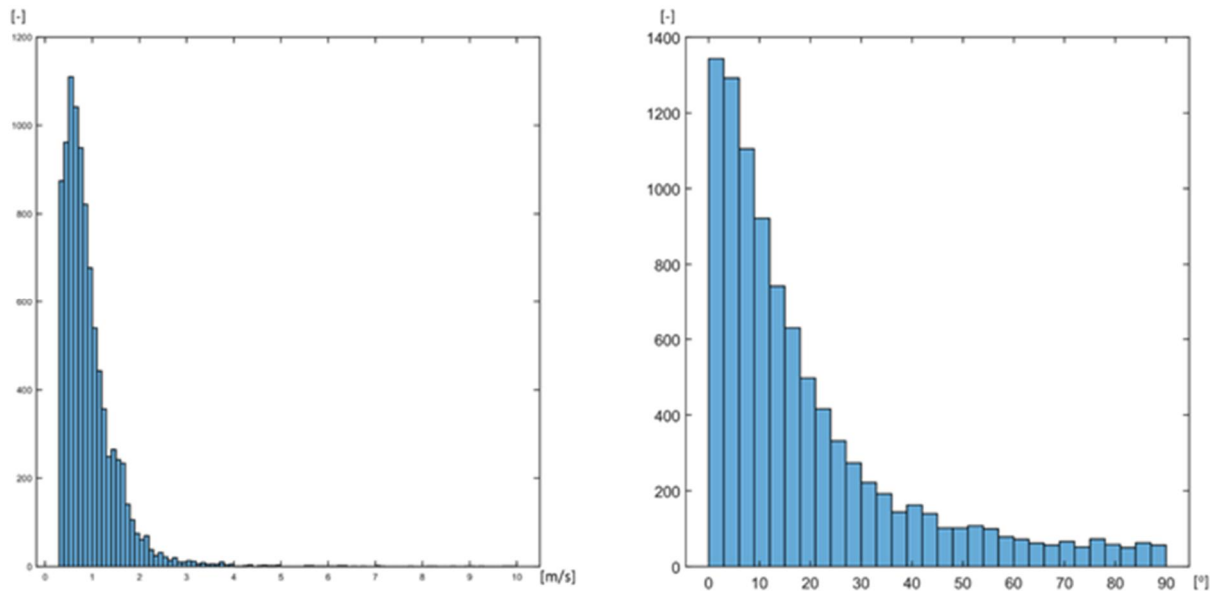

Figure 3. Histogram of movement speed and angle difference between steps

For the development of the Agent-Based Model that will recreate a selected pathway, all of the necessary data has been obtained. The essential model parts are the current location, the step length distribution, the step angle distribution and the turn vector (Figure 4). The step length was directly translated into the natural distribution model, where its metadata, such as the position of the semi-bell peek, range and standard deviation, were used in its description. Each step has been resolved with a random number that indicated the step deviation from the natural model. The step turn, was resolved in the nearly the same fashion as the step length. Data from angular distribution was re-calculated to absolute values to take into account all possible changes of the angle. It produced a semi-bell shaped natural distribution curve which includes data on step angle differences and the maximum difference between all the necessary statistical values. The maximum angle difference allowed locating the left and right arc ends. The turn vector always followed the general turn of the simulated agent from the previous step, and that is why the whole distribution arc was continually updating its step turn per time. It seems that the angular distribution was split into two semi arcs (the left with a red outline, the right one with a blue outline). For each step, it is necessary to decide which part of it will be taken into account. Selection of the semi arc part was made by measuring the distance from the intended goal to the arc ends. The smaller distance has always been selected. If both distances were even, selection of the semi-arc was random.

The designed movement model has been tested, and it has produced similar pathways in a normalized transition scenario. That is why it has been decided to test its capabilities in a virtual environment. For this purpose, four different layout cases were investigated. The developed movement model requires information on the transition towards the intended goal. If the randomly generated start and end point
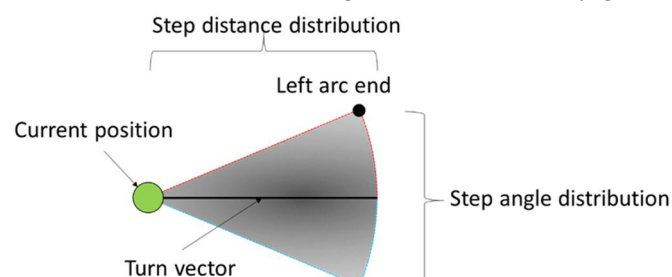

Figure 4.Drawing of principal element of the transient movement simulator

have some obstacle between them, the developed model will not be able to compile it. The model requires more information regarding the specific position of the investigated case. That is why test layouts have to be supplemented with additional information on local transition points. The points play the role of navigation points inside the tested cases. Location of navigation points was detected with a use of self-designed software. Each corner of the layout has been taken into account. During that phase, it was tested if other currents are also visible. If so, each pair provides one middle point between the visible corners. All middle points on the layout outline were discarded. Next, all the middle points identify connections between them. From now on these points are called orientation points. The identification process uses once again visibility analysis. Once the matrix of connection is established, the layout is ready to be tested for movement simulations. Each layout had one designed purpose of the testing. Scenarios naming is adequate to the figure included in results (Figure 5).

•   Scenario A was investigating behavior in close corners;
•   Scenario B was a testing transition between rooms in cross-like connections;
•   Scenario C was focused on testing transition in an advanced room connection;
•   Scenario D has tested the reaction of the model to the occurrence of possible circular patterns in the simulated space.

Once the location of beginning (A) and ending (B) of the pathway is established, the shortest pathway algorithm is used to describe the step of the agent. The procedure prescribes a schedule of the points that the agent should reach to achieve the end position. For example, it describes that to reach goal (B), the agent has a scheduled vector [A,1,3,6, B]. It means that from point (A) it should be focused on aiming towards and reaching orientation point no.1. First of all, in each step, the algorithm checks if end point (B) is reached. If not, the software movement procedure is initialized. While progressing towards this orientation point, it will detect that orientation point no. 3 is visible. From that event, the aim of the agent is updated to reach orientation point no. 3. This procedure continues until the end of the pathway has been reached. To avoid extension of the simulation time of the agent movement, if endpoint (B) is within reach of the agent, it finishes the movement simulation with a transition towards it. If during the simulation procedure the next step taken by the agent will cause over-crossing the space outline, it resolves the step selection once again. In critical condition, if it is not possible to resolve a movement without crossing the layout border, the simulation software cancels the last step, and it is resolved once again. To check the frequency of appearance of such situation, a specific counter has been applied in the code. Once all of the monitoring systems have been set up to track all the potential misconducts of the movement model, it is possible to initialize simulation procedure.

4. Results

For the model test trials, four different layout scenarios were designed in a black and white bitmap. Once all the necessary assumptions had been formulated, each layout was tested for generation of fifty pathways. The start and end points were generated at random. Due to that reason, it was impossible to predict if all of the selected areas layouts will be visited by the designed agent. The results of the simulated pathways were displayed in Figure 5.
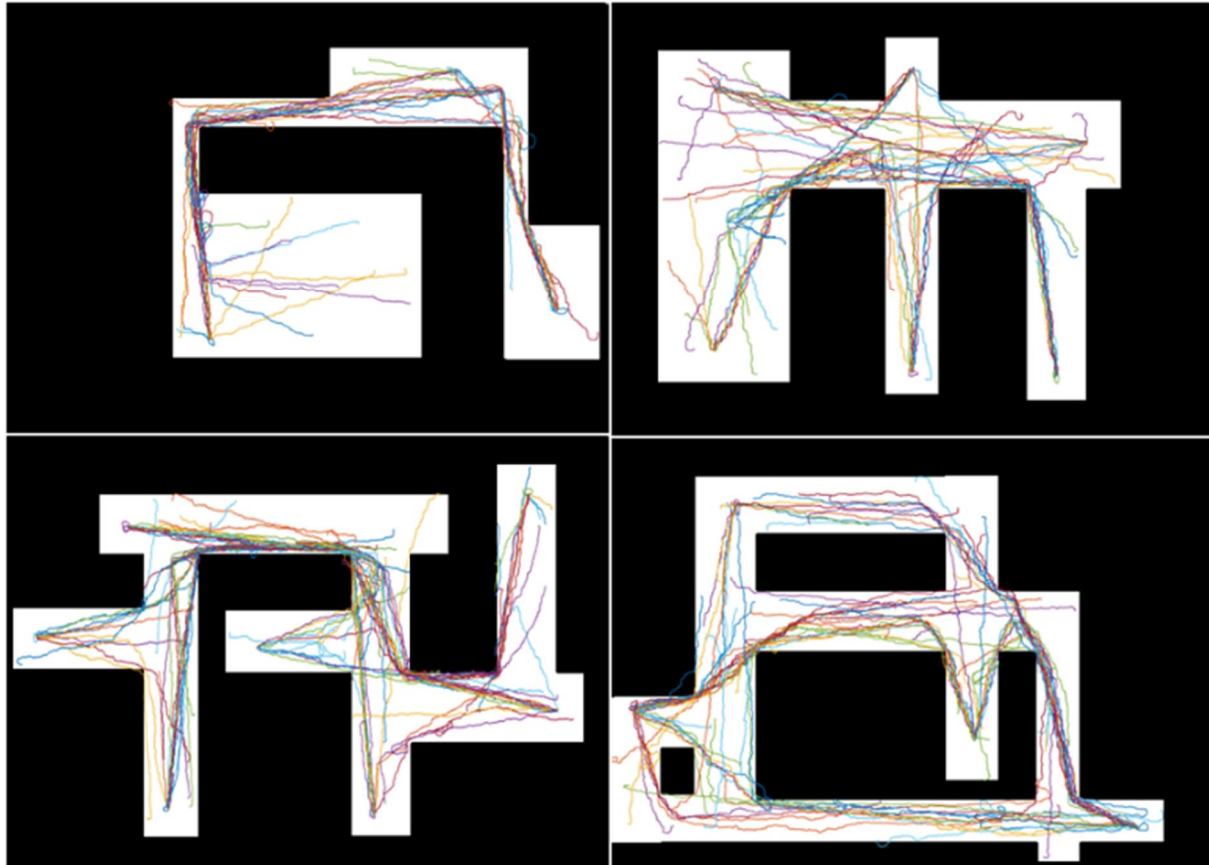
Figure 5.Movement simulation results on selected floor layouts

5. Discussion & Conclusion

The performed simulation test shows that the suggested modeling approach has its potential applicability in building performance simulation. This study has to be treated as the first step towards transient occupant behavior modeling. The results of simulated pathways show that a few parts of occupant transition have to be improved. The next improvement steps should aim at a transition from the raster to vector data. Re-calculation of the whole matrix into a raster matrix can be overwhelming, for CPU or GPU. That is why it is important to deliver this calculation switch. Additionally, all of the designing software uses vector drawing as its basis. If the potential designers are to support their work with such a tool, it is critical that their workflow will not be disturbed by additional unnecessary tasks. Simulation of the transient occupant behavior has to reach the level of a complementary feature. During the simulation trials, none of the monitored errors were observed. The movement simulator respected the limitations imposed on all the four test scenarios.

With continuous work aimed at improvement of this method, it will eventually be able to obtain more precise data regarding occupant thermal sensation. With the use of this method, it may be coupled with

a converging CFD simulation. If it holds information of the indoor air thermal condition, it is possible to track the history of an occupant's exposition to local conditions. Such asset can be used for evaluation of the already existing local thermal comfort models. Once the accuracy of the previously developed comfort models is taken into account, it will be discovered what their transient accuracy is. It can indicate the potential direction for future improvement towards development of the personal transient thermal comfort model. To reach that level of confidence, the proposed approach has to be put to more extensive trial. To do so, it is necessary to conduct a new series of measurements with wider spectrum of occupants included in the study. It has to be checked what are the pathway decision-making factors which influence the selection of a specific pathway. This knowledge can be obtained from investigating simple movement scenarios. Each future scenario will have to include the choice of at least two different pathways. Additionally, it will have to be checked how the beginning position influences the selection of a pathway. Once the above-mentioned parameters have been investigated, data collected in those experiments may be used for improvements in the movement model. The current state of development shows promise of a potential breakthrough in building performance simulations. If all the missing parts of the knowledge are verified and receive positive feedback, it will mean that this model could be treated as the first milestone in development of the Building Occupant Transient Agent-Based Model (BOT-ABM).

References

[1]   A K. Parsons, 2014, *Human Thermal Enviroments*, (CRC Press)
[2]   P. Fanger, 1970, *Thermal comfort*, (Danish Technical Press)
[3]   J. Hviid and M. Baun Kjærgaard, 2018, *Activity-Tracking Service for Building Operating Systems*, (IEEE International Conference on Pervasive Computing and Communications Workshops)
[4]   D. Yan, W. O'Brien, T. Hong, X. Feng, H. B. Gunay, F. Tahmasebi and A. Mahdavi, 2015, *Occupant behavior modeling for building performance simulation: Current state and future challenges*, (Energy and Buildings)
[5]   T. Hong, J. Langevin and K. Sun, 2018, *Building simulation: Ten challenges*, (Building Simulation)
[6]   I. B. Arief-Ang , M. Hamilton and F. D. Salim, 2018, *RUP: Large Room Utilisation Prediction with carbon dioxide sensor*, (Pervasive and Mobile Computing)
[7]   J. Dziedzic, Y. Da and V. Novakovic, 2017, Occupant migration monitoring in residential buildings with the use of a depth registration camera, (Procedia Engineering)
[8]   J. Dziedzic, Y. Da and V. Novakovic, 2018 , *Real Time Measurement of Dynamic Metabolic Factor (D-MET)*, (Springer Proceedings in Energy)
[9]   Micorsof Xbox One, "http://www.xbox.com/," [Online], Available: http://www.xbox.com/en-US/xbox-one/accessories/kinect. [Accessed 10 1 2018].