A Human Tracking and Sensing Platform for Enabling Smart City Applications

Hirozumi Yamaguchi
Osaka University
Osaka, Japan
h-yamagu@ist.osaka-u.ac.jp

Akihito Hiromori
Osaka University
Osaka, Japan
hiromori@ist.osaka-u.ac.jp

Teruo Higashino
Osaka University
Osaka, Japan
higashino@ist.osaka-u.ac.jp

ABSTRACT

The progress of IoT technologies, which connect, control and operate everything in the physical world, is expected to realize secure and convenient societies and communities, and embodies a variety of smart city applications. However, not a few smart city applications are also "human-centric", which require individuals' and human crowds' locations and behavior. In this paper, we present our project that realizes a human tracking and sensing platform called *Hitonavi* for enabling location sharing and trajectory identification of people. The core technology is LIDAR-based highly-accurate tracking of people as well as human crowd detection, and smartphone-assisted trajectory identification for enabling "pinpoint" indoor location services in retail shops, exhibition halls, office buildings, commercial complex, train stations and so on. Leveraging these functions, we may build a variety of smart city applications on top of the Hitonavi platform, which would bring new values for smart communities, smart buildings and many other smart city applications. We introduce the concept and key technologies of the Hitonavi platform and discuss potential applications.

KEYWORDS

Human Tracking, Sensing, Smart City

ACM Reference Format:
Hirozumi Yamaguchi, Akihito Hiromori, and Teruo Higashino. 2018. A Human Tracking and Sensing Platform for Enabling Smart City Applications. In *Proceedings of The 19th International Conference on Distributed Computing and Networking (SCC-2018 Workshop).* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3170521.3170534

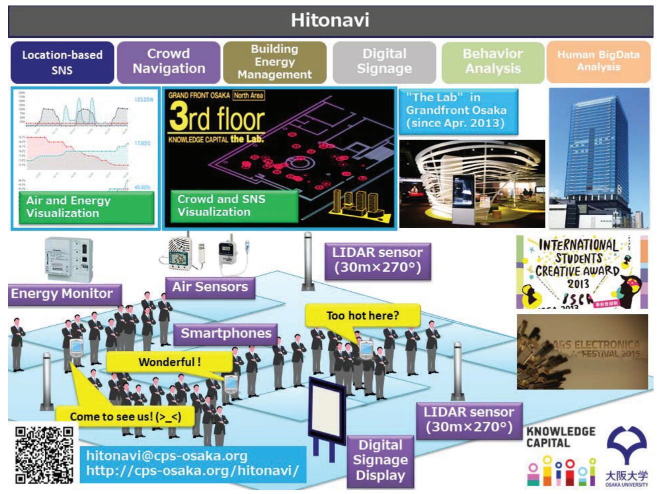


Figure 1: Hitonavi Platform Overview

1 INTRODUCTION

Smart computing technologies have been playing a key role to solve various social issues on environment, health, energy and transportation, and therefore, they always target humans in a broad sense. In cyber-physical systems, humans are users, who are beneficial from the system functionalities, and affect system decisions. Henceforth, the systems can become more strategic if we are able to precisely understand humans – what they are thinking, how and why they act, and how environment and things affect their behavior, as opposed to many of the current systems where uncertainty of humans brings a lot of inefficiency. Fortunately, in this IoT era, a variety of sensors are being available and less expensive, and huge amount of sensor data can be aggregated and utilized. Humans and their surroundings/environments should be sensed and understood for better and smarter feedback to humans themselves.

For example, suppose a booth in an exhibition event where only a few visitors are there but visitors think it is attractive. If such atmosphere and situations can be shared, it is beneficial to those who are seeking interesting booths while avoiding congestion. Social networking services or micro blogs play a major role for location-based information sharing, but they do not provide fine-grained location-based filtering for temporary places such as booths in an exhibition, and many smart city applications need such a function that enables spot information sharing. In case of smart buildings, electricity-saving and automated air/light control systems have recently appeared. Those systems, often called task-ambient air conditioning and lightning, embed environmental sensors which monitor temperature, humidity and illumination and human detection sensors such as cameras and Infra-red sensors in the buildings to adaptively control air and light. However, it is still challenging to track human locations and share personal (and subjective) thermal comfort to pursue the best balance between better comfort and more energy-efficient services.

In this paper, we introduce our application-level platform called *Hitonavi* aimed at providing accurate anonymous human location, congestion and trajectories. On top of this platform, we have prototyped a function to support smartphones to obtain their own

Workshops ICDCN '18, January 4–7, 2018, Varanasi, India

ACM ISBN 978-1-4503-6397-6/18/01…$15.00
https://doi.org/10.1145/3170521.3170534

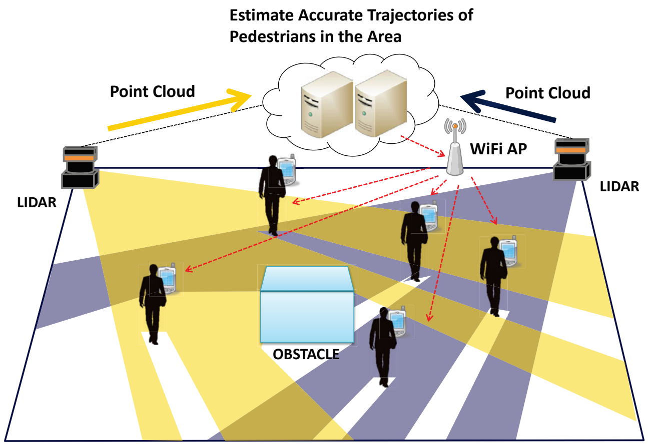


Figure 2: Hitonavi Architecture

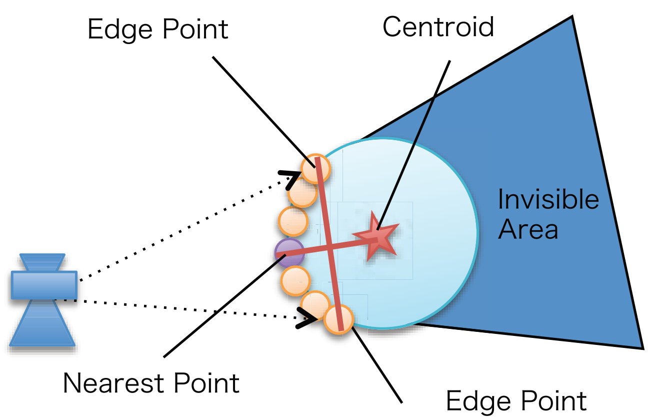


Figure 3: Human Body Detection by Point Clustering

pin-point locations as well as the surrounding crowd information, which can be used by each application to enable new services or enhancing the existing services such as geolocation-based tweets in indoor space. This concept is also applied to smart energy, where we need to control the indoor air environment more precisely based on the number of people and their locations. We note that localization of mobile devices has been widely and actively investigated in past decade and we have also presented collaborative localization of mobile phones [3, 4, 6, 9]. However, these approaches do not support human crowd recognition and accuracy is not usually sufficient.

We have installed Hitonavi platform to three real places, a real exhibition space in Osaka, an office building in China, and a retail store in Kyoto. We show application scenarios where innovative services can be embodied by the Hitonavi platform.

2 HITONAVI – SYSTEM ARCHITECTURE AND DEPLOYMENT

2.1 Architecture

The core technology of our Hitonavi platform is LIDAR-based human tracking. LIDAR (Light Detection and Ranging) sensors have been attracting significant attention as an enabler of accurate human behavior sensing. We assume LIDAR sensors are installed in indoor space such as exhibition halls, large shopping malls and commercial complex. Each LIDAR sensor scans the field and measures distance to the surrounding objects within a few tens of meters based on propagation delay of eye-safe laser pulses. Those sensors are connected via Ethernet, and the measurements are aggregated to a central server, which estimates the locations and trajectories of people in that area. The locations and trajectories can be fed back to the pedestrians and other users who may use it. Figure 2 illustrates this architecture.

Accurate tracking of pedestrians in indoor environments has been a major research challenge in pervasive computing. One of the most popular approaches is vision-based systems, which employ cameras to capture heads, faces and bodies of pedestrians [1, 11]. LIDAR sensors are advantageous as they do not cause privacy concerns and enable smooth and continuous tracking that results in accurate trajectory estimation.

We assume that LIDAR sensors are placed at the waist level height in the target area. Since our LIDAR sensor measures 1,080 points within a 270 degree angle ranges with 25 ms intervals, a pedestrian can be detected as several points near by each other as shown in Fig. 3. The outline can be formed from the detected points according to a clustering algorithm. In addition, the centroid of the pedestrian can be derived from the outline. As shown in the figure, the line can be constructed between the two edge points from the outline. We can also obtain the perpendicular to the line. Thus, the point, whose distance to the nearest detected point is the radius of the outline, on the perpendicular can be the centroid of the pedestrian. Finally, we derive a trajectory for a pedestrian by connecting the centroids over time. Since LIDARs scan the points with 25 ms interval and the distances between centroids are very small, we can connect the centroids whose distance is smaller than a threshold. If we find a new centroid, we create a new pedestrian in the area.

If a flock (or a crowd) of pedestrians exists, it is detected in a different way as the LIDAR rays are obscured by human bodies in the crowd. Based on the experiments in our work [2], we coarsely identify a set of grid cells, which may possibly contain human crowds. For each LIDAR sensor, the server maintains a histogram of the recent distance samples in each ray direction. To mitigate the impact of short-term distance fluctuations due to people passing by between a LIDAR sensor and the nearest pedestrian crowd, we filter out the distance samples that come from such moving pedestrians. Since measurement intervals of LIDAR sensors are typically a few tens of milliseconds, the maximum distance that a pedestrian can move during the contiguous time steps is no more than 0.1m. It is usually much less than the minimum distance to the neighboring pedestrians, and thus we can accurately associate the human locations that belong to the same pedestrian by finding the nearest human location at the previous time step. If the total displacement

(a) Grandfront Osaka Building (b) “The Lab” in Grandfront Osaka (c) Hitonavi Display in “The Lab”

Figure 4: Hitonavi Tracking Technology Showcase

over the recent 0.5 seconds is larger than 0.3m, the server concludes that the pedestrian is moving and thus excludes the corresponding distance value in calculating the histogram. Then the server examines the presence of human crowds in each direction, and calculates the distance to a boundary of the nearest human crowd, if any. Finally, for each direction, the server locates “crowd edge point” and regards the cells that contain any crowd edge point as crowd cells and we estimate the density of pedestrians in each crowd cell. Interest readers may refer to [2] for the algorithm details.

2.2 Deployment

Hitonavi has been displayed since April 2013 in “the Grandfront Osaka” building (Fig. 4(a)), one of the largest commercial complex in Kansai region, Japan. It has newly opened since April 23rd 2013, and 1 million people have visited the place for the first three days. “The Lab” (Fig. 4(b)) is an exhibition space open for public, and we installed 8 LIDAR sensors to cover two floors of The Lab. Figure 4(c) shows the display we actually used in the first year. We also installed the Hitonavi system to the large office space owned by a large company in China (20 LIDAR sensors) and a retail shop in Kyoto (15 LIDAR sensors). We have tracked the office workers to analyze their behavior during the work time and manage the meeting space occupancy ratios. Throughout these deployment cases, we have come up with the use of accurate trajectories for each smartphone user, and new indoor services will be enabled by the support of Hitonavi. In the following section, we will introduce our technology to make Hitonavi a pinpoint locator, and we will introduce several application scenarios in Section 4.

3 HITONAVI AS A PINPOINT LOCATOR OF SMARTPHONE

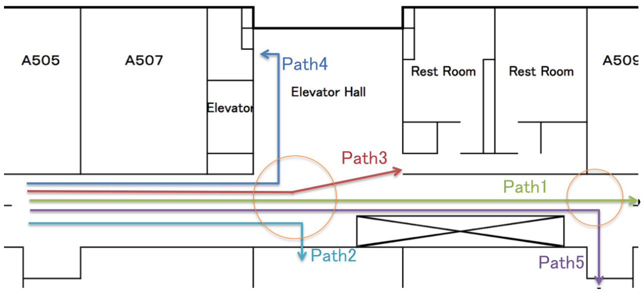
As shown in Figure 2, LIDAR sensors can derive pedestrians’ accurate trajectories in the area of interest, and such a set of trajectories can be fed-back to the pedestrians’ mobile devices (basically, smartphones) for crowd navigation or some other application purposes. However, those smartphones cannot identify their own trajectories as they are anonymous. If it is possible to identify the owners of trajectories, it means the smartphones can obtain their accurate locations and trajectories, which would be beneficial to many types of location-based services.

We leverage the smartphones’ capabilities of detecting their movement by inertial sensors (accelerometers and gyroscope), and introduce a matching algorithm to assess the similarity between each trajectory sent from the central server and its own movement (distance and directions). Each smartphone estimates the moving distance and direction from the data obtained by these two sensors, and finds the trajectory that has the best similarity with the movement. In this way, Hitonavi can give these pedestrians their highly-accurate positions and trajectories, which enable new smart and intelligent applications.

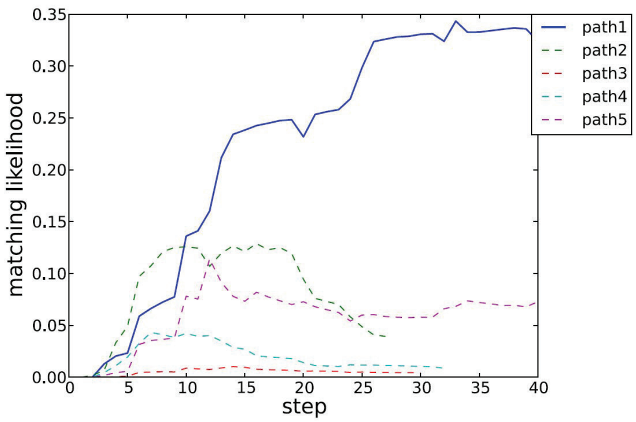
The moving distance and direction using the accelerometer and gyroscope are obtained as follows. While a pedestrian is walking, we can see fluctuations in vertical accelerations. Thus, we detect walking steps for each pedestrian based on this observation. The moving distance for a pedestrian is calculated from the number of steps and the distances for walking steps. As for the direction, we calculate the variation of the moving direction in each walking step by the integral of angular velocity from a gyroscope over the time for the step.

After we derive a rough trajectory composed of moving distances and directions, we match it with one of LIDAR trajectories. Since both LIDAR sensors and phone sensors can detect pedestrians who do not move in the area, each phone firstly finds correspondence between the phone trajectory and each of the LIDAR trajectories based on stop and go information in addition to the geographical information. Then we can remove the LIDAR trajectories that contradict the stop and go information from the phone. After that, we calculate the similarity of moving distances and directions in the candidates, and find an appropriate trajectory. Figure 5 shows the experimental map and the matching likelihood result for “path1” of the map. We can see that the likelihood of path 1 grows as the number of steps increases faster than the other 4 paths illustrated in the map. For the technical details, the readers may refer to Ref. [10].

0



(a) Experimental Map



(b) Matching Likelihood for Path1

Figure 5: Trajectory Identification Evaluation

(a) Crowd Detection by LIDAR Sensors in "The Lab"

(b) HumanS Simulator for LIDAR Placement Design []

Figure 6: Hitonavi Crowd Detection Case Study and LIDAR Placement Design Support

4 APPLICATION SCENARIOS

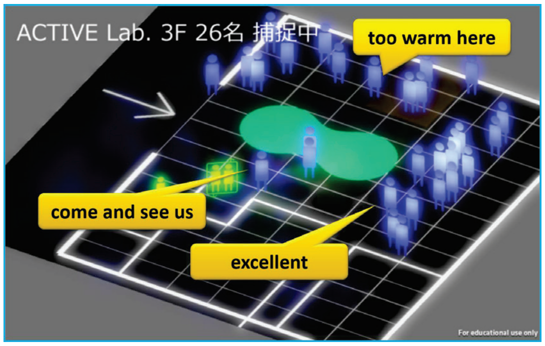
We introduce some of smart city applications, which may be designed and implemented over the Hitonavi platform and benefit from its functionalities.

4.1 Crowd Navigation

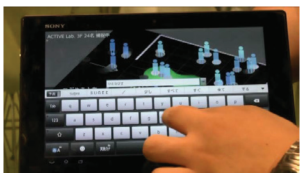
Crowd detection is an essential technology to provide context-aware services in crowded public space like a complex commercial buildings. To observe wide areas with less number of devices, LIDAR sensors are reasonable solutions, each of which can monitor a few tens of meters with over 270 degrees. As we briefly explained in Section 2.1, our crowd detection has been applied to the most crowded day of "The Lab" as shown in Figure 6(a). Hitonavi can work with low density scenarios where the trajectory of each pedestrian can be identified, as well as with the high density scenarios like Figure 6(a) where crowd edge is detected to estimate the human density. We note that if sufficient coverage is not obtained by LIDAR sensors only, we may incorporate the idea of participatory sensing, which would offer a strong alternative solution. Particularly, we have developed an algorithm to classify the surrounding congestion levels of a pedestrian by microphone and accelerometers [7]. We note that we have presented our tool to support the optimal placement of LIDAR sensors knowing pedestrian flows (Fig. 6(b)), which is presented in Ref. [5].

4.2 Indoor Geotagged Tweets

Each smartphone with pinpoint location provided by Hitonavi can post tweets about the visited or current places. The check-in services in SNS provide a similar function, but they only support the listed places (e.g. points of interest). Therefore, they do not support to point geographical coordinates in indoor space. Such function would be necessary to ask residents to tweet their thermal comfort to better control the air conditioning in smart building applications and to share comments that are annotated to some particular locations in event exhibitions or some other indoor places. Figure 7(a) shows our concept of geotagged tweets with pinpoint location in the share space, and we have implemented a micro blog interface on Hitonavi (Fig. 7(b) and Fig. 7(b)). We have also presented a concept of event mining from the geotagged tweets in Ref. [2].



(a) Indoor Geotagged Tweets: Illustration of Concept



(b) Prototyped System Interface

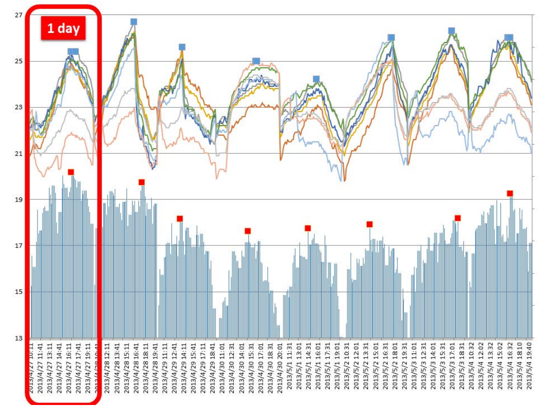
(c) Displayed Geotagged Tweet

Figure 7: Indoor Geotagged Tweets

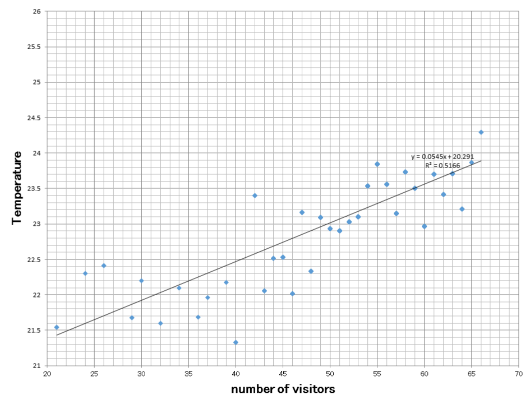
4.3 Smart Building

It has been reported sensing-based task-ambient control substantially reduces wasteful power consumption (10% or more in average [8]). However, most of those systems are not able to capture the real comfort of occupants as sensors are not around them. Also, the thermal effect by human bodies is not ignorable in crowded situation. To confirm the fact, we installed 8 temperature sensors in "The Lab", counted the number of visitors and measured the temperature. Figure 8(a) shows the number of visitors (shown by blue bars using the right-Y axis) and temperature values from the 8 sensors (shown by colored lines using the left-Y axis) measured for 8 days (the site was most heavily crowded with a number of visitors). We can see clear correlation between the temperature and the number of visitors where peak time of both values (red and blue small boxes) in each day is quite similar. Actually, the correlation coefficient by linear approximation shown in Fig. 8(b) is 0.71, which proves such relation between them.

To reveal what happens in such cases, We have conducted CFD simulation with six different scenarios by combinations of air intake temperature (22 or 18) and crowd density (0.1, 1.0 or 2.7). Fig. 9 show the screenshots of the simulation tool where the colors on the human body surface show PMV (Predicted Mean Vote) values and those of the surface of ground, walls and ceiling show the air temperature. The arrows in the screenshots illustrate air flows, and it is observed that air is heated and then rises. As seen in Fig. 9(d) with air intake temperature 18°C and density 0.1, PMV values range between 0.0 and 1.0, and it is higher near human heads. This is caused by rise of air by human bodies and seen in any other figures. Based on this observation, we have designed a thermal comfort estimation model from the density of crowd, intake air temperature and the distance to HVAC, leveraging the accurate location of occupants.



(a) Number of Visitors (Blue Bars) and Temperature at 8 points (Lines) in "The Lab"



(b) Temperature and the Number of People Correlation

Figure 8: Temperature and Crowd Correlation in The Lab.

5 CONCLUSION

In this paper, we have introduce our Hitonavi platform, which enables innovative services in urban indoor space. We have shown some human-centric smart city application scenarios to show the significance of the platform.

ACKNOWLEDGEMENT

This work was supported in part by JSPS KAKENHI JP15H02690 and JP26220001.

REFERENCES

[1] M. Enzweiler and D. M. Gavrila. 2009. Monocular Pedestrian Detection: Survey and Experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 31, 12 (Dec 2009), 2179–2195. https://doi.org/10.1109/TPAMI.2008.260

[2] K. Fujita, T. Higuchi, A. Hiromori, H. Yamaguchi, T. Higashino, and S. Shimojo. 2015. Human crowd detection for physical sensing assisted geo-social multimedia mining. In *2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*. 642–647. https://doi.org/10.1109/INFCOMW.2015.7179458

[3] T. Higuchi, S. Fujii, H. Yamaguchi, and T. Higashino. 2014. Mobile Node Localization Focusing on Stop-and-Go Behavior of Indoor Pedestrians. *IEEE Transactions on Mobile Computing* 13, 7 (July 2014), 1564–1578. https://doi.org/10.1109/TMC.2013.139

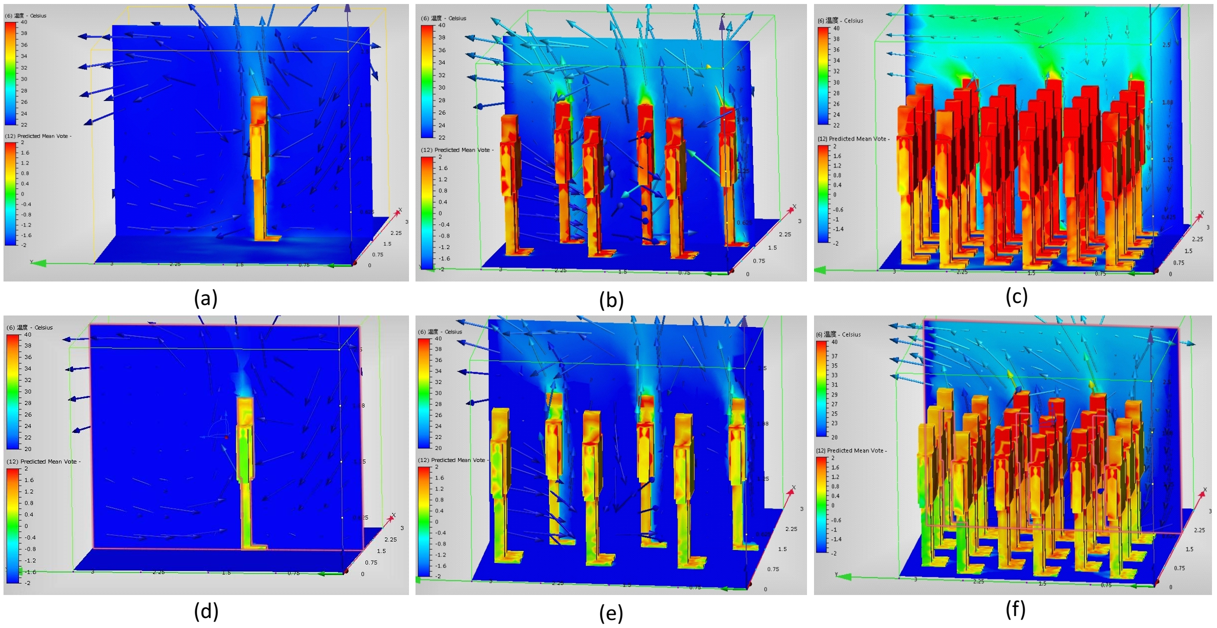


Figure 9: CFD Simulation with Human Bodies as Thermal Sources; (a)-(c):22°C, (d)-(f):18°C, (a)(d):Low Density, (b)(e):Medium Density and (c)(f):High Density

[4] Takamasa Higuchi, Hirozumi Yamaguchi, and Teruo Higashino. 2014. Context-supported local crowd mapping via collaborative sensing with mobile phones. *Pervasive and Mobile Computing* 13, Supplement C (2014), 26 – 51. https://doi.org/10.1016/j.pmcj.2013.10.012

[5] A. Hiromori, H. Yamaguchi, and T. Higashino. 2013. Sensor Placement Optimization Method for People Tracking. In *2013 Seventh International Conference on Next Generation Mobile Apps, Services and Technologies*. 62–67. https://doi.org/10.1109/NGMAST.2013.20

[6] Yuki Maekawa, Akira Uchiyama, Hirozumi Yamaguchi, and Teruo Higashino. 2014. Car-level Congestion and Position Estimation for Railway Trips Using Mobile Phones. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 939–950. https://doi.org/10.1145/2632048.2636062

[7] Tomohiro Nishimura, Takamasa Higuchi, Hirozumi Yamaguchi, and Teruo Higashino. 2014. Detecting Smoothness of Pedestrian Flows by Participatory Sensing with Mobile Phones. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers (ISWC '14)*. ACM, New York, NY, USA, 15–18. https://doi.org/10.1145/2634317.2642869

[8] K. Niwa, S. Tsukami, K. Kobayashi, and R. Hu. 2010. Measurement and simulation of task/ambient air conditioning system. *Journal of the Society of Heating, Air-Conditioning and Sanitary Engineers of Japan* 84, 8 (2010), 637–642.

[9] Y. Noh, H. Yamaguchi, and U. Lee. 2017. Infrastructure-Free Collaborative Indoor Positioning Scheme for Time-Critical Team Operations. *IEEE Transactions on Systems, Man, and Cybernetics: Systems* PP, 99 (2017), 1–15. https://doi.org/10.1109/TSMC.2016.2615652

[10] T. Takafuji, K. Fujita, T. Higuchi, A. Hiromori, H. Yamaguchi, and T. Higashino. 2014. Indoor Localization Utilizing Tracking Scanners and Motion Sensors. In *2014 IEEE 11th Intl Conf on Ubiquitous Intelligence and Computing and 2014 IEEE 11th Intl Conf on Autonomic and Trusted Computing and 2014 IEEE 14th Intl Conf on Scalable Computing and Communications and Its Associated Workshops*. 112–119. https://doi.org/10.1109/UIC-ATC-ScalCom.2014.103

[11] Beibei Zhan, Dorothy N. Monekosso, Paolo Remagnino, Sergio A. Velastin, and Li-Qun Xu. 2008. Crowd analysis: a survey. *Machine Vision and Applications* 19, 5 (01 Oct 2008), 345–357. https://doi.org/10.1007/s00138-008-0132-4