A Review of Hybrid Indoor Positioning Systems Employing WLAN Fingerprinting and Image Processing

Firdaus Firdaus
Universitas Islam Indonesia, Faculty of Industrial Technology
Department of Electrical Engineering, Jl. Kaliurang, Yogyakarta, Indonesia
firdaus@uii.ac.id

Noor Azurati Ahmad
Universiti Teknologi Malaysia, Razak Faculty of Technology and Informatics
Department of Advanced Informatics, Jalan Sultan Yahya Petra, Kuala Lumpur, Malaysia
azurati@utm.my

Shamsul Sahibuddin
Universiti Teknologi Malaysia, Razak Faculty of Technology and Informatics
Department of Advanced Informatics, Jalan Sultan Yahya Petra, Kuala Lumpur, Malaysia
shamsul@utm.my

Abstract – Location-based services (LBS) are a significant permissive technology. One of the main components in indoor LBS is the indoor positioning system (IPS). IPS utilizes many existing technologies such as radio frequency, images, acoustic signals, as well as magnetic sensors, thermal sensors, optical sensors, and other sensors that are usually installed in a mobile device. The radio frequency technologies used in IPS are WLAN, Bluetooth, ZigBee, RFID, frequency modulation, and ultra-wideband. This paper explores studies that have combined WLAN fingerprinting and image processing to build an IPS. The studies on combined WLAN fingerprinting and image processing techniques are divided based on the methods used. The first part explains the studies that have used WLAN fingerprinting to support image positioning. The second part examines works that have used image processing to support WLAN fingerprinting positioning. Then, image processing and WLAN fingerprinting are used in combination to build IPS in the third part. A new concept is proposed at the end for the future development of indoor positioning models based on WLAN fingerprinting and supported by image processing to solve the effect of people presence around users and the user orientation problem.

Keywords – indoor positioning, image processing, orientation, people presence, WLAN.

1. INTRODUCTION

Location-based services (LBSs) are a significant permissive technology that has wide-ranging applications in our life [1]. LBS are services that combine geographic location with other information to provide more helpful services [2]. The LBS market is growing rapidly [3]; a market report estimates the LBS market to generate up to USD 77.84 billion revenue by 2021 [4]. One of the main components of LBS is the positioning system, which can be divided into indoor or outdoor. Global Navigation Satellite Systems (GNSS) have been used over a wide range of applications for outdoor positioning.

However, GNSS cannot be used for indoor positioning because GNSS signals are not strong enough to penetrate buildings. The failure of GNSS to work indoors has caused people to develop the Indoor Positioning System (IPS) [5]. IPS-based service has great economic potential as well; a report estimates the global indoor location market to grow to $4,424.1 million by 2019, as shown in Fig. 1 [6]. IPS is any system that gives a precise position inside of buildings, such as smart buildings [7], hospitals [8], airports [9], [10], subways [11], construction sites [12], industrial sites [13], and university campuses [14]. IPS utilizes many existing technologies such as radio frequencies (RFs) [15], images [16], magnetic fields [17], acoustic signals [18], as well as thermal [19], infrared [20], and optical [21]–[23] sensors together with other sensory information collected by a mobile device (MD) [24]. Examples of RF technology used in IPS, among others, are WLAN (Wireless Local Area Network) or Wi-Fi [25], [26], Bluetooth [27], ZigBee [28], [29], RFID [30], frequency modulation (FM) [31], and Ultra-wideband (UWB) [32], [33].
WLAN has been highlighted as a preferred technology due to its accurate positioning results and minimal infrastructure cost and power consumption [34], [35]. WLAN location detection techniques can be categorized into three general categories: proximity, triangulation, and fingerprint [15]. Fingerprinting is based on a pattern recognition technique that combines radio frequency (RF) with location information, e.g. an environmental label to show the position of the MD.

WLAN fingerprinting is usually conducted in two phases: offline and online. In the offline phase, a site survey is conducted to collect the value of the received signal strength indicator (RSSI) at many reference points (RPs) from all the detected access points (APs). Some researchers have proposed using indoor positioning technologies that do not require the construction of offline fingerprint maps [36] or just by updating the maps automatically [37]. In the online phase, a user samples or measures an RSSI vector at his/her position. Then, the system compares the received vector of RSSI with the stored fingerprints in the radio map (RM) database. The position is thus estimated based on the most similar “neighbors”, i.e. the set of RPs with RSSI vectors that closely match the RSSI of the target [38]–[40].

Image-based object localization tries to detect the location based on image processing [16]. It has two main approaches [41]. The first approach tries to estimate absolute object localization. This approach performs object detection, object tracking, and 2D-3D mapping [42]. The second approach gives a coarse estimation of object localization based on scene analysis.

Silva et al. [43] discussed 62 review papers in the field of indoor positioning. Khalajmehrabadi [44] made a specific review on WLAN fingerprinting IPS. However, no one has specifically made a review of the hybrid method between WLAN fingerprinting and IPS image positioning. These two methods were chosen because both combined can provide high accuracy and could very possibly be cheaply implemented. This paper explores the combination of WLAN fingerprinting and image processing for IPS. It discusses how this combination can improve location accuracy. Then, a new concept to enhance positioning accuracy is proposed.

Several studies have combined image processing with WLAN fingerprinting for IPS. In this review, the combination of WLAN fingerprinting with image processing is divided into 3 methods. The first part explained papers that have used WLAN fingerprinting to support image positioning. The second part consists of past work that has used image processing to support WLAN fingerprinting positioning. Then, the third part elucidates the works that have combined image processing and WLAN fingerprinting to build an indoor positioning system (IPS).

These methods are shown in Fig. 2. The first method is discussed in Section 2 while the second and third methods are discussed in Sections 3 and 4, respectively.

Fig. 1. Infographic on the future of IPS [6]

Fig. 2. The methods for image positioning, WLAN positioning, and Indoor positioning based on (a) WLAN fingerprinting, (b) Image Processing, (c) WLAN fingerprinting and Image Processing

2. IMAGE POSITIONING SUPPORTED BY WLAN FINGERPRINTING

In this method, image processing is used to determine position while WLAN Fingerprinting is used to support the process and improve system performance. Basically, this method is an image positioning or camera-based positioning technique that recognizes a location from an image taken of an indoor environment [21], [45]. Then, WLAN fingerprinting is used to improve the performance of the image positioning technique.

Ruiz [46] tried to find the location of some smartphones using the different sensors available in commodity smartphones. They used accelerometers, digital compasses, and Wi-Fi signals to determine the smartphone position based on motion estimators, probabilistic techniques, and fingerprinting methods. Then, they used images from the smartphone camera to determine the correspondences between 2D and 3D. Ruiz et al. [47] continued their research using fingerprinting to determine an initial cluster with a high probability of showing the device location. Then, the matching process was automatically performed for these images contained in the selected cluster.

Savvas [48] proposed a new system that combined visual detection with radio readings. The visual detec-
tion was captured using an available camera. This system is called Radio and Vision Enhanced Localization (RAVEL). RAVEL can improve the performance of visual trackers by overcoming issues such as occlusions and people entering/exiting the scene of visual detections.

Ito et al. [49] designed an IPS based on Wi-Fi and an RGB-D camera. First, they estimated a coarse global position using Wi-Fi fingerprinting and then computed a precise estimation using the data from the RGB-D camera based on the environmental floor plan.

Jiang [50] proposed a new method for indoor localization based on 3 parameters: Wi-Fi signal strength, orientation, and image. The main parameter was the image, while Wi-Fi signal strength and orientation were used to improve the technical accuracy and decrease the time required for image matching. They used Wi-Fi fingerprinting to determine an initial position, and then an inertial sensor to obtain user orientation. Finally, the user location was determined by matching candidate images with a certain tree image database, selected based on Wi-Fi signal strength and orientation.

Xiang [51] presented a pose estimation using 6-degrees-of-freedoms based on portable 3D visuals. First, in the offline phase, they built a 3D and signal strength model for an indoor environment. They then used a Wi-Fi signal to locate the device in a 3D sub-model. Next, they applied feature matching between the online-captured 2D image and the keyframe images to build a 3D model. Finally, the estimation process in the 3D domain was done using the iterative closest point (ICP) and the RANSAC algorithm.

Jiao [52] improved smartphone camera-based positioning using TC-OFDM (Time and Code Division Orthogonal Frequency Division Multiplexing). They developed FAST-SURF to compute a detailed location while reducing computational time at the same time. Speeded Up Robust Features (SURF) is a new invariant interest point detector and descriptor that is scalable and rotatable [53].

Wang et al. [54] proposed an indoor location that combined Wi-Fi signal and real-time images, as shown in Fig. 3. In the offline stage, Wi-Fi and image data were collected at the location via feature extraction application on the data. A total of 300 Wi-Fi fingerprints and 50 photos from different angles were collected at each sampling point. In the online stage, the Wi-Fi fingerprint acquired by the user's mobile phone was matched to obtain a rough position interval, then, the distance compensation and SURF point similarity matching of the real-time captured image were used to obtain the similarity with each pre-captured image, so that the coordinates of the sampling points corresponding to the most similar image could be obtained. However, the accuracy of the system was not mentioned.

The above-method was then improved by Wang et al. [55], who used AlexNet [56] to perform supervised learning classification modeling on the image for regional scoping in the online positioning stage. However, the position error was still quite high, at 2.2 m (90% of cumulative probability).

Simultaneous Localization and Mapping (SLAM) is an extensively researched topic in robotics. However, visual SLAM algorithms face several challenges including perceptual aliasing and high computational cost. Hashemifar [57] proposed a method to utilize Wi-Fi received signal strength to alleviate the challenges faced by visual SLAM algorithms, as shown in Fig. 4. The approach was able to improve the accuracy of visual SLAM algorithms by 11% on average and reduced computation time on average by 15% to 25%.

To improve the accuracy of SLAM, the Wi-Fi signal strength and the RGB-D images had to be combined [58]. The Extended Karman Filter (EKF) employs the Wi-Fi signal strength information to estimate the pose of a robot and the locations of the APs, while the graph-optimization part employs the RGB-D images to estimate the poses of the robot.

Image positioning techniques provide good accuracy. However, because they are based on image processing, these techniques require high computational power, processing time, and consume a lot of battery life. A list of past research related to Image Positioning Supported by WLAN Fingerprinting is shown in Table 1.
3. WLAN FINGERPRINTING POSITIONING SUPPORTED BY IMAGE PROCESSING

In this method, WLAN Fingerprinting was used to determine the position whereas image processing was used to support the process and improve performance. Mirowski [59] built a navigation robot using multiple sensors, including signal strength (RSSI) and odometry. He used image processing in the offline phase to construct a physical map of a building. In the online phase, WLAN fingerprinting was used to determine the location.

Ching [60] proposed a new framework based on Wi-Fi Localization and image processing. The system consisted of client and server components. The client-side consisted of three modules: Camera, Wi-Fi Scanner, and Gyroscope. The server consisted of four modules: Image Feature Extraction, Wi-Fi Localization, Image Feature Comparison, and Database Search. The study used JOpenSurf to extract and compare image features. JOpenSurf is a variation of the SURF (Speeded Up Robust Features) method in Java. The precision of this method was 94%—less than 5 m—and the accuracy was 1.64 m.

Niu [61] also proposed WLAN and image combination for IPS, similar to the solution given by Ching [60], but did not use a gyroscope. The study demonstrated that by using the proposed method, they were able to reduce the search time by up to 83%, but it did not mention the error position.

Hou [62] used the RANSAC algorithm for AP selection in the online stage. Hou’s [62] solution did not use images but instead used an algorithm usually used in image processing. The method was able to filter and remove the APs impacted by environmental variation. As a result, not only was the amount of calculation reduced but also the positioning accuracy was improved. The average error of the positioning system was 1.63 m, and the minimum error was 0.76 m.

Ali et al. [63], [64] used image processing and IoT sensors (ESP8266 module) in the offline stage to generate and update a radio map database automatically. They used WLAN Fingerprinting to find the position. The experimental results showed that the proposed system achieved an average accuracy of 2 m.

Travi-Navi [65] captured images and collected Wi-Fi samples and IMU-sensed data to pack them into navigation traces, as these traces can be shared with the user directly or through a cloud server. To achieve this aim, some lightweight trace-merging algorithms were employed in the study to discover the overlapping segments and merge multiple traces. Gu et al. [63] proposed an indoor localization system that combined Wi-Fi and magnetic fingerprinting, image-matching, and people co-occurrence (WAIPO). A list of past research into WLAN Fingerprinting Positioning supported by Image Processing is shown in Table 2.
The average image database to define location estimation. SIFT [70] combined a Wi-Fi fingerprint database and an

4. INDOOR POSITIONING BASED ON IMAGE PROCESSING AND WLAN FINGERPRINTING

The combined WLAN and image processing techniques used data from RSSI, WLAN, and images to determine location. Usually, this technique has two databases, an RSSI database, and an image database. The initial method that combined Wi-Fi and camera was introduced by Hattori [67]. The study used a camera to get information from a two-dimensional marker placed on the floor. When they could not find the marker, they used Wi-Fi fingerprinting to define the marker’s position. They used two types of fingerprints, RSSI from APs and RSSI from neighbor MDs (ad-hoc communication).

Nathan [68] introduced WHLocator, which provides location information based on three technologies: Wi-Fi, altimeter, and images. The Wi-Fi component was used to obtain MAC addresses and RSSI values from APs. Then, an algorithm was used to determine a list of possible locations. The location decision component requests data from the altimeter component and the scene analyzer to determine an estimated location. A scene classifier calculates the probability of a scene being a particular type. This method can be used to distinguish positions on the same floor that are difficult to distinguish based on fingerprints.

EV-Loc is another example of the same combination of techniques. Teng [69] proposed a localization technique called EV-Loc. They used a matching engine to find the correspondence between an electronic identifier (E) and its visual appearance (V). The visual signals improved the accuracy of the system. Levcchev [70] combined a Wi-Fi fingerprint database and an image database to define location estimation. SIFT features were used for image processing. The average on-demand success rate was 86.3%, higher than that of continuous capture (66.9%). The mean error was 2.46 m, the maximum error was 10.55 m, and the standard deviation was 1.35 m.

Bejuri [71] proposed an indoor positioning system based on wireless LAN and a camera. This method used grey-world-based feature detection and matching. The author also used a model fitting approach to combine the data from WLAN and the camera. First, the image was segmented and the corner detected. Then, a model fitting approach was used to match the point of interest in the image with the coordinated information and to get the location estimation. This method achieved a positioning accuracy of around 1.5 m in 22% of trials. Xu et al. [72] proposed Argus, an image-based localization system. This system works by extracting geometric constraints from crowd-sourced photos. Argus uses image and WLAN fingerprints to find the location inside a building. The study obtained the image from a smartphone camera by shooting a point of interest (POI) near their position. Then, they sent the image of the POI and the WLAN fingerprint to a server. The 80-percentile error was 1.38 m in the Mall and 2.30 m in the Plaza. The mean localization error using Argus was 1.29 m.

In 2017, Jiau et al. [73] proposed a model that considered the effect of people. They proposed a new signal attenuation model to compensate for propagation loss, based on the population density. Image processing was used to detect a human body using a deep CNN [42] of local humans to which the number of humans was then calculated. This technique is based on Fully Convoluted Localization Neural Network (FCLN) architecture without the Recursive Neural Network language model to speed up the processing.

Then, they improved the positioning performance of the system by combining wireless signals and RGB images, as shown in Fig. 5 [74]. LBP is a local binary pattern. The proposed method was able to provide an accuracy of up to 0.83 m (RGB, Wi-Fi, and Bluetooth), and 2.88 m (RGB and Wi-Fi). However, the method requires significant time investment and effort in the offline stage, as well as a large amount of computing power in the online stage. Redzic et al. [75] presented an indoor localization system that combined both imagery and WLAN data to support a wide variety of location-based applications. The WLAN Localization component computes a location by matching the input RSS fingerprint against the fingerprint database. The Image Localization component compares the input image with the image database.

Then, the Fusion Engine employs the Threshold-based component that combines the WLAN-based and image-based locations to output the final user location, as shown in Fig. 6. The study used an extended Naive-Bayes approach for WLAN Localization, and a hierarchical SURF (Speeded Up Robust Features) vocabulary tree of descriptors for Image Localization. The

**Table 2. Research related to WLAN Fingerprinting Positioning Supported by Image Processing**

<table>
<thead>
<tr>
<th>Author</th>
<th>Algorithm &amp; Performance</th>
<th>Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mirowski et al. [59] (2012)</td>
<td>• SLAM</td>
<td>RGB-D sensors</td>
</tr>
<tr>
<td>Ching et al. [60] (2013)</td>
<td>• SURF</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Niu et al. [61] (2014)</td>
<td>• Probabilistic Filter Model, SURF</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Hou et al. [62]</td>
<td>RANSAC algorithm for AP selection in the online stage</td>
<td>Image Algorithm</td>
</tr>
<tr>
<td>Gu et al. 2017 [63] (2017)</td>
<td>• Wi-Fi and magnetic fingerprinting, image-matching</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Trav-Navi [64] (2017)</td>
<td>• Wi-Fi and IMU</td>
<td>Server</td>
</tr>
</tbody>
</table>

The average image database to define location estimation. SIFT [70] combined a Wi-Fi fingerprint database and an
positioning error was 1.908 m for the particle filter and 1.904 m for the threshold-based engine. The average computational time was 5.68 s for the particle filter and 8.87 s for the threshold-based engine.

Image processing in combination with WLAN Fingerprinting provides good accuracy. However, these methods run both image matching and fingerprint matching, so more computational power and time are needed. For example, Argus consumed 171 mW additional power than the WLAN fingerprint method. A list of past research into Indoor Positioning based on Image Processing and WLAN Fingerprinting is shown in Table 3.

Table 3. Research related to Indoor Positioning based on Image Processing and WLAN Fingerprinting

<table>
<thead>
<tr>
<th>Author</th>
<th>Algorithm &amp; Performance</th>
<th>Device</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hattori et al.</td>
<td>Two types of fingerprints: RSSI from APs neighbor MDs.</td>
<td>Phone</td>
</tr>
<tr>
<td>Nathan et al.</td>
<td>A scene classifier</td>
<td>Phone</td>
</tr>
<tr>
<td>Teng et al.</td>
<td>Matching engine to find the correspondence between an electronic identifier and its visual appearance.</td>
<td>Phone</td>
</tr>
<tr>
<td>Levchev et al.</td>
<td>SIFT</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Bejur and Mohamad</td>
<td>Feature detection &amp; matching using Grey-world</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Xu et al.</td>
<td>The mean error was 1.2m</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Jiau et al.</td>
<td>Deep CNN</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Jiau et al.</td>
<td>A Fully Convoluted Localization Neural Network</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Jiau et al.</td>
<td>0.83 m (RGB, Wi-Fi, and Bluetooth), and 2.88 m (RGB and Wi-Fi)</td>
<td>RGB images</td>
</tr>
<tr>
<td>Redzic et al.</td>
<td>An extended Naive-Bayes approach for WLAN Localization</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Redzic et al.</td>
<td>Hierarchical vocabulary tree of SURF descriptors for Image Localization</td>
<td>Phone camera</td>
</tr>
<tr>
<td>Jiao et al.</td>
<td>Positioning error was 1.908 m</td>
<td>RGB image</td>
</tr>
<tr>
<td>Jiao et al.</td>
<td>Combination of images/Wi-Fi/magnetic/inertial</td>
<td>RGB image</td>
</tr>
</tbody>
</table>

Positioning accuracy was less than 1.23 m.

Jiao et al. [76] leveraged on a CNN framework and a particle filter to build a two-level high accuracy positioning architecture fusing on the image/Wi-Fi/magnetic/inertial. They proposed a new image named WMV, which fused 1-D signals (Wi-Fi/magnetic) with an RGB image. It was noted that the magnetic field, the image, and the inertial information were infrastructure-free and ubiquitous. The core of the fusion architecture was an improved four-layer deep neural network that integrated a Convolutional Neural Network (CNN) and an improved particle filter. The ubiquitous positioning accuracy of the proposed algorithm was less than 1.23 m.

5. FUTURE DEVELOPMENT

One of the performance indicators of IPS is accuracy, which is the error distance between the actual location and the estimated location. Many applications require precise IPS, among others, for emergency cases and patient monitoring. It is essential for the American emergency telephone number, 911, to know the location of the caller as precisely as possible to control delays in emergency response because delays in response can lead to a loss of lives. Therefore, 911 defined a new
standard called "next-generation 911 (NG911)" [77], [78]. The US National Aeronautics and Space Administration Jet Propulsion Laboratory (NASA JPL) cooperated with the Department of Homeland Security Science and Technology Directorate (S&T) to develop a precision outdoor and indoor navigation and tracking for emergency responders [79]. It is also important to build an indoor navigation system that can track firefighters within a meter [80].

WLAN-based RSSI Fingerprinting can provide highly accurate position estimations [81]. However, its accuracy decreases when there are many environmental changes. The RSSI of WLAN is affected by the environment. Examples of obstacles that can cause fluctuations in RSSI are walls, ceilings, and people [15], [38]. The effect of walls and ceilings have been discussed in [82]–[84]. The effect of people on signal strength was investigated in [85] for 60 GHz [86], 868 MHz, and [87] [88] 2.4 GHz. The result showed that people’s presence in the Line of Sight (LOS) between the AP and the Mobile Device (MD) decreased the RSSI by 2 dBm to 5 dBm. This decline in RSSI could result in a position error of more than 2 m. Hence, an adaptive IPS that can adapt to environmental changes, including user orientation [89] and people’s presence [90], [91], is needed to improve the accuracy of IPS.

The question then becomes, how do we detect people around users inside a building? In this case, the user’s MD can be used to detect the people around him/her [92], while others in the MD neighborhood could be found based on many protocols of communication such as WLAN direct [93], WLAN aware [94], near-field communication (NFC) [95], sound waves [96], and Bluetooth [97]. These solutions are less practical because they require many sensors or devices and it is assumed that the people around the user have brought their MD in an active condition. If the people did not bring an MD or they did not enable the communication protocol, then, they cannot detect their neighbor. Operating this function on MD would also consume a lot of battery.

Alternative ways to use the device installed in the building, therefore, have to be found. One of the devices that can be used and has been widely installed in buildings is the Closed-circuit television (CCTV). A report mentioned that 245 million CCTVs had been installed in 2014, with 65 percent of them being installed in Asia [98]. A market report from www.marketsandmarkets.com estimated the total market of video surveillance applications to reach $25.43 billion by 2016. There are even more installed CCTVs in Kuala Lumpur than in Australia [99], and more than 30% of Malaysian buildings have 21–30 CCTVs [100].

CCTVs and image processing are an alternative technique to detect people around users, so no extra cost is needed and the accuracy of IPS can finally be improved. Generally, people counting in image processing are classified as direct or indirect [101], as shown in Fig. 7. The direct method can be divided into model-based and trajectory-clustering-based [102]. The direct method was done by segmenting and detecting each individual and then counting each using some classifiers [103]. The direct method is also called object detection.

The model-based method detects every single person using a model or human shapes [104]. An example of the model-based approach can be seen in Fig. 8, which used head shape detection. In contrast, the trajectory-clustering-based approach detects every independent motion [105]. In the indirect approach, the people can be counted based on some features in the foreground image, which are then processed using learning algorithms and statistical analysis [106]. It has been proven that detecting features are simpler than detecting persons. So the estimation process can be done using linear or non-linear regression functions such as support vector regression, as well as Bayesian and Gaussian regressions [107].

Yoshinaga et al. [108] tried to count the number of pedestrians in real-time using a blob descriptor and neural networks. Chan et al. [107] used a mixture of dynamic-texture motion models to divide a crowd into homogeneous motion components. Their method extracted each seg-
mented region into a set of holistic low-level features and then employed a Bayesian regression to map the features into the number of people per segment. This technique enabled the counting of crowds moving in different directions, but it was tested for an outdoor environment.

Another method proposed by Foroughi [103] was based on image retrieval, a simple global image descriptor, sparse representation, and random projection [110]. This method yielded better performance than regression methods, especially when the dataset was large. This technique, however, was also tested for outdoor environments.

Zhang [111] used a label distribution learning (LDL) strategy for crowd counting. Sivabalakrishnan [112] proposed a feedback update-by-detection scheme in foreground segmentation and attempted to balance between background noise removal and stationary people segmentation. A pair of collaborative Gaussian process models (GP) with different kernels was used by Zaydi [113] to estimate the number of people based on image processing. Ma [114] attempted to count the number of people in a video of them crossing a line using regression and local-level features. This method was able to estimate both cumulative and instantaneous counts.

A new concept in which an indoor positioning model based on WLAN fingerprinting and supported by image processing can be proposed to count the people around a user while a digital compass could be used to determine the user orientation. The data from image processing and the digital compass then serves as input for the WLAN IPS to improve the system accuracy without adding any extra cost, as seen in Fig. 9.

Finally, a new, accurate adaptive indoor positioning model based on WLAN Fingerprinting is proposed, using existing common devices (Access Points and CCTVs) installed in a building. This model was proposed to overcome the effect of any environmental change in a dynamic environment, especially the effect of people presence around a user and user orientation.

6. CONCLUSION

Several studies have combined image processing with WLAN fingerprinting to build an IPS. This review proposed dividing the combined techniques of WLAN fingerprinting and image processing into 3. The first part examined the use of WLAN fingerprinting to support image positioning in the literature. The second part assessed works that have employed image processing to support WLAN fingerprinting positioning. Then, studies that had combined image processing and WLAN fingerprinting were reviewed to build IPS in the third method. This combination aims to improve system performance, especially the accuracy of the system. The second method gave more accurate position estimation, but it required more computational power and memory, as it mainly used the image processing technique to define user position.

The accuracy of IPS is significantly affected by the environment, especially the effect of people presence around the user and user orientation. WLAN fingerprinting and image processing techniques can be used in combination to solve this problem. A new concept for the future development of an indoor positioning model based on WLAN fingerprinting and supported by image processing was proposed for counting the people around a user while a digital compass could be used to determine the user orientation. The data from image processing and the digital compass could then serve as input for the WLAN IPS to improve system accuracy without adding any extra cost.

ACKNOWLEDGEMENT

This work has been supported by the UTM Grant (UTM-TDR 9.2(T1)). Hence, the authors would like to thank Universiti Teknologi Malaysia (UTM) for their support.

7. REFERENCES:


on advanced intelligent mechatronics, Zurich, Switzerland, 4-7 September 2007, pp. 1–6.


