Novel Indoor Positioning Mechanism Via Spectral Compression

Jukka Talvitie, Markku Renfors, *Fellow, IEEE*, and Elena Simona Lohan, *Senior Member, IEEE*

***Abstract*—Received signal strength (RSS) measurements are important in indoor location solutions based on WiFi, cellular networks or Bluetooth. RSS-based positioning involves two phases, namely, learning and estimation. The database sizes required both for the learning and for the estimation phases grow rapidly as the network coverage areas and the number of access points number increase. Achieving large-scale/global localization solutions would be possible if the database size bottlenecks were solved. We present here an innovative approach based on spectral compression, which allows a tremendous reduction in the database sizes in both learning and estimation phases. We introduce the new concept of compressed RSS images. We show how, through an astute 2-D frequency analysis, only a fraction of the transform-domain components need to be stored and transferred to/from the mobiles. Our idea is validated with WiFi real-life measurements from five multistory buildings. We show that our method is able to provide comparable results with the traditional fingerprinting approach, but with up to 80% reduction in the database sizes.**

***Index Terms*—Fingerprinting, indoor positioning, path loss models, received signal strength, wireless local area networks.**

I. Introduction

INDOOR positioning using RSS measurements has been considered as one solution to the increasing demand of indoor position services [1]–[5]. The majority of RSS-based localization solutions are learning-based solutions, such as the fingerprinting [1], [2], [6]–[12], where the user position estimate is based on pre-collected learning data from the target area, which is then transmitted to the mobile during the estimation phase. To reduce the database size, and thus the data transfer to/from mobile, radio propagation Path Loss (PL) models have been proposed [3], [13], [15] as they require only a fraction of the database size compared to the fingerprinting approach. However, the positioning accuracy with the PL models falls often far behind the fingerprinting approach.

The problem solved in this letter is: how to achieve the accuracy of fingerprinting approaches with the low complexity of PL approaches. To achieve this goal we introduce the new idea of *compressed RSS images*, where the RSS power map per each Access Point (AP) is seen as an 'image' and analyzed in the spectral domain. Because these 'images' contain implicitly the measurement and shadowing noise, a lossy compression in transform domain would not only decrease the size of the information to be stored, but would also act as a low-pass filtering, by filtering out the noise. The compression of the RSS images is based on the obvious correlation existing between neighbor measurements taken from one AP. The compression is exploiting the transform-domain content of the RSS images and only a fraction of the components is stored in the database. The information 'lost' during the lossy compression stage is in fact mostly the noise part, and thus we are able to achieve very accurate results, similar to the traditional fingerprinting case. For the spectral analysis, we propose the use of the Discrete Cosine Transform (DCT), because of its ability to concentrate the frequency domain energy into a few DCT coefficients. Spectral analysis of the RSS measurements has also been studied in [16] and [17], but neither of the letters exploits the spatial correlation of the RSS measurements taken from the same AP as done in this letter.

The amount of compression, and thus, the quality of the compression, can be adjusted by choosing the number of stored DCT coefficients per image. Although the method can be directly extended to N-dimensional images by using N-dimensional DCT, in this letter we consider only 2-D images representing measurements from one AP and one floor and we address the full 3-D space through a set of these 2-D images.

Manuscript received August 25, 2015; revised October 19, 2015; accepted November 11, 2015. Date of publication November 26, 2015; date of current version February 12, 2016. This work was supported by the Academy of Finland (project 256175). The associate editor coordinating the review of this paper and approving it for publication was Y. Shen.

The authors are with the Department of Electronics and Communications Engineering, Tampere University of Technology, 33101 Tampere, Finland (e-mail: jukka.talvitie@tut.fi; markku.renfors@tut.fi; elena-simona.lohan@tut.fi).

Digital Object Identifier 10.1109/LCOMM.2015.2504097

II. Learning Database Structure

As learning data we assume a set of fingerprints (FP), where the i^{th} FP in the k^{th} floor is given as $\{x_{i,k}, y_{i,k}, \mathbf{ID}_{i,k}, \mathbf{P}_{FP,i,k} : i = 0, \ldots, N_k^{(FP)} - 1, k = 0, .., N_f - 1\}$, where $x_{i,k}$ and $y_{i,k}$ are the x coordinate and y coordinate of the measurement, $\mathbf{ID}_{i,k}$ is a $1 \times N_{i,k}^{(AP)}$ vector including the identity ID_j (e.g. the MAC address) for the j^{th} AP, and $\mathbf{P}_{FP,i,k}$ is a vector of the same size including the corresponding RSS measurements for each AP in dBm. Moreover, N_f is the number of floors, $N_k^{(FP)}$ is the FP number in the k^{th} floor, $N_{i,k}^{(AP)}$ is the number of RSS measurements in the i^{th} FP and k^{th} floor, and $j = 0 : N_{AP} - 1$, where N_{AP} is the total number of APs in the database. FPs are mapped into a rectangular grid with grid interval of d_{grid} meters as done in [6]. If multiple RSS measurements from the same AP are observed in the same grid point, the arithmetic mean of the RSS values is used. The FP can also be organized AP-wise for each floor so that for the j^{th} AP (with ID_j) and the k^{th} floor, measurements are given as $\{x_{m,k},\ y_{m,k},\ P_{m,k} :\ \text{m} \in \Omega_{j,k}\}_{j,k}$, where the $\Omega_{j,k}$ includes all those FP indices i in the k^{th} floor, where the j^{th} AP has been heard (i.e. when ID_j is found from $\mathbf{ID}_i$), and $P_{m,k}$ is the stored RSS value.

III. Compression and Positioning With RSS Images

A. Definition and Compression of RSS Images

The RSS images are built and compressed separately for each AP and floor. Hence, for the j^{th} AP and k^{th} floor we have a set of $N_{j,k}^{(RSS)} = \left|\Omega_{j,k}\right|$ (i.e., the number of elements

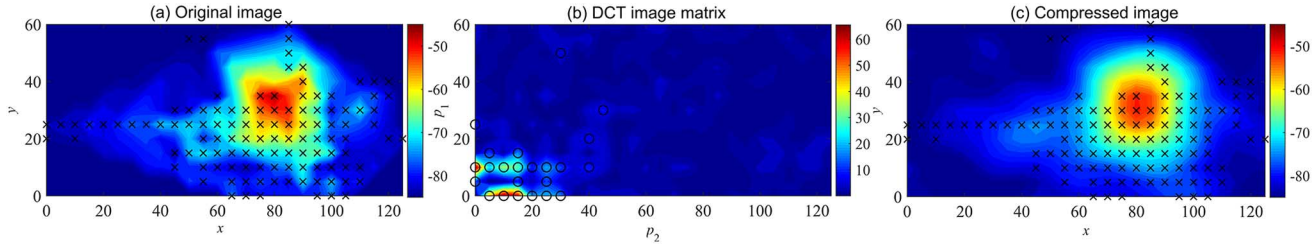

Fig. 1. Example of compression of a real life measurement based (interpolated and extrapolated) RSS image. The original image is shown in (a) with the fingerprint locations illustrated as "**x**", the DCT of the original image (used to build (c)) is shown in (b), where "**o**" show the stored DCT coefficients, and finally the corresponding compressed image is shown in (c).

in $\Omega_{j,k}$) different coordinates and RSS values described as $\{x_{m,k},\ y_{m,k},\ P_{m,k}\}_{j,k}$. First, the empty grid points in the coverage area of the AP are interpolated and extrapolated, e.g. via linear interpolation and minimum value extrapolation [15]. Secondly, we denote the resulted full RSS image for the j^{th} AP and k^{th} floor as $\mathbf{G}_{j,k}$, which is a matrix of size $M_{j,k} \times N_{j,k}$ providing RSS values for each x coordinate (in columns) and y coordinate (in rows) (see part (a) in Fig. 1). Thirdly, we apply the DCT transform to the RSS image (see part (b) in Fig. 1) and obtain the DCT image matrix $\mathbf{H}_{j,k}$. Here, it is worth of noticing that $\mathbf{G}_{j,k}$, includes naturally both the PL model and the locally varying shadowing model. Thus, compared to the conventional PL models, which are only able to present the local average of the RSS values, the proposed approach offers much more accurate RSS modeling including the shadowing process. The energy in $\mathbf{H}_{j,k}$ is concentrated around the low frequencies and the magnitude of the DC (Direct Current) coefficient is typically much larger than with the other DCT coefficients. This is because the DC coefficient represents the mean of the RSS image which is usually considerably larger than the RSS variations inside the image. Then, *the zero mean RSS image* is defined as

$$\mathbf{G}_{j,k}^{(0)} = \mathbf{G}_{j,k} - \mu_{j,k}, \text{ where } \mu_{j,k} = \frac{1}{M_{j,k} \cdot N_{j,k}} \sum_{s=0}^{M_{j,k} \cdot N_{j,k} - 1} \mathbf{G}_{j,k}(s) \tag{1}$$

where $\mu_{j,k}$ is the DC coefficient of $\mathbf{H}_{j,k}$ and $\mathbf{G}_{j,k}(\mathrm{s})$ is the s^{th} element of the matrix $\mathbf{G}_{j,k}$ (of size $M_{j,k} \times N_{j,k}$) with arbitrary index order. Each element of the DCT matrix of $\mathbf{G}_{j,k}^{(0)}$ is:

$$\mathbf{H}_{j,k}(p_1, p_2) = w_{p_1} w_{p_2} \sum_{n_1=0}^{M_{j,k}-1} \sum_{n_2=0}^{N_{j,k}-1} \mathbf{G}_{j,k}^{(0)}(n_1, n_2) \cos\left(\frac{\pi (2n_1+1) p_1}{2M_{j,k}}\right) \cos\left(\frac{\pi (2n_2+1) p_2}{2N_{j,k}}\right), \tag{2}$$

where

$$w_{p_1} = \begin{cases} 1/\sqrt{M_{j,k}}, & p_1 = 0 \\ \sqrt{2/M_{j,k}}, & 1 \le p_1 \le M_{j,k} \end{cases}, \quad w_{p_2} = \begin{cases} 1/\sqrt{N_{j,k}}, & p_2 = 0 \\ \sqrt{2/N_{j,k}}, & 1 \le p_2 \le N_{j,k} \end{cases},$$

where $p_1^{\text{th}} = 0, \ldots M_{j,k} - 1$ is the row index and $p_2^{\text{th}} = 0, \ldots N_{j,k} - 1$ is the column index. As the matrix dimensions are defined, we can use a single index q to represent the elements of $\mathbf{H}_{j,k}$ as $\mathbf{H}_{j,k}(q)$, which are further referred to as the DCT coefficients. Thus, by storing only the most significant coefficients of $\mathbf{H}_{j,k}$ (e.g., in part (b) in Fig. 1), it is possible to have an approximate representation of the original image. In order to ensure desirable data compression for each image, we choose to select the number of stored DCT coefficients with respect to $N_{j,k}^{(RSS)}$ and the number of stored DCT coefficients is given as

$$N_{j,k}^{(coef)} = \left\lfloor \eta N_{j,k}^{(RSS)} \right\rfloor, \tag{3}$$

where $0 \le \eta \le 1$ is a design parameter defining the ratio between the $N_{j,k}^{(coef)}$and $N_{j,k}^{(RSS)}$, and $\lfloor \cdot \rfloor$ is a floor function. Thus, the stored DCT coefficients are then the $N_{j,k}^{(coef)}$ largest magnitude elements taken from $\mathbf{H}_{j,k}$ denoted as $\left\{\mathbf{H}_{j,k}(q_l) : l = 0, \ldots, N_{j,k}^{(coef)} - 1\right\}$, where q_l define the indices for the largest magnitude elements.

Finally, the image is reconstructed from the reduced set of DCT coefficients (see part (c) in Fig. 1). In order to reconstruct the images, we need to store the following parameters in the database for each image: image (or AP) identity (ID_j), floor index k, image coordinates (e.g., xy-coordinates of one corner of $\mathbf{G}_{j,k}$), image dimensions ($M_{j,k}$ and $N_{j,k}$), DC coefficient ($\mu_{j,k}$), and DCT coefficients values ($\mathbf{H}_{j,k}(q_l)$) with their corresponding indices q_l with $l = 0, \ldots, N_{j,k}^{(coef)} - 1$. The reconstruction begins by defining the compressed version of the DCT matrix of size $M_{j,k} \times N_{j,k}$ as

$$\hat{\mathbf{H}}_{j,k}(q) = \begin{cases} \mathbf{H}_{j,k}(q_l), & \text{when } q = q_l \\ 0, & \text{otherwise} \end{cases}. \tag{4}$$

Next, the compressed version of $\mathbf{G}_{j,k}^{(0)}$ is defined at n_1^{th} row and n_2^{th} column via the inverse DCT as

$$\hat{\mathbf{G}}_{j,k}^{(0)}(n_1, n_2) = \sum_{p_1=0}^{M_{j,k}-1} \sum_{p_2=0}^{N_{j,k}-1} w_{p_1} w_{p_2} \hat{\mathbf{H}}_{j,k}(p_1, p_2) \cos\left(\frac{\pi (2n_1+1) p_1}{2M_{j,k}}\right) \cos\left(\frac{\pi (2n_2+1) p_2}{2N_{j,k}}\right), \tag{5}$$

where w_{p1} and w_{p1} were earlier given in Eq. (2), and $n_1^{\text{th}} = 0, \ldots M_{j,k} - 1$ is the row index and $n_2^{\text{th}} = 0, \ldots N_{j,k} - 1$ is

the column index. Finally, the *compressed RSS image* can be obtained by adding the image mean (i.e., the DC coefficient) as

$$\hat{\mathbf{G}}_{j,k} = \hat{\mathbf{G}}^{(0)}_{j,k} + \mu_{j,k} \tag{6}$$

It can be seen (part (c) in Fig. 1) that the reconstructed image is a smoother representation of the original image as the high frequency DCT coefficients have been removed from the data.

B. Database Size Computations

The database size is computed with respect to the quantity of real valued numbers required to represent the data. The database size affects not only the required storage space, but also the data transfer rate to/from the mobile, which is crucial to be kept low. The size of the FP database can be given as

$$B_{FP} = \sum_{k=0}^{N_f-1} \sum_{i=0}^{N_{FP,k}-1} 3 + 2N^{(AP)}_{i,k}, \tag{7}$$

since each FP in each floor requires information on the fingerprint location regarding the x, y and z coordinates and the RSS data with the AP identities and their corresponding RSS values. The size of the database with the compressed images $\hat{\mathbf{G}}_{j,k}$ can be calculated as

$$B_{IM} = \sum_{k=0}^{N_F-1} \sum_{j=0}^{N_{AP}-1} 7 + 2N^{(coef)}_{j,k}, \tag{8}$$

since each image requires information on the image identity, image xyz-location, image dimensions, image mean, and the stored DCT coefficient values and their indices.

C. Position Estimation With RSS Images

Based on the previous discussion, the RSS measurements from the j^{th} AP in k^{th} floor at location (x,y) can be modeled as

$$R^{(x,y)}_{j,k} = \hat{\mathbf{G}}_{j,k}(T(x,y)) + W \tag{9}$$

where $T(x,y)$ maps the coordinates (x,y) into a corresponding image index q using the stored information of the image location and image dimensions, and W is the noise term including the measurement noise and image compression error. For simplicity we assume W as a white Gaussian noise with variance σ^2_W. However, in practice the noise distribution might differ from the Gaussian, and moreover, W is colored, since at least the compression error is correlated between nearby locations. Further in Section IV when deriving the positioning results, we have assumed $\sigma^2_W = 6$ dB for all the considered approaches. Nonetheless, assuming that a user observes a set of independent RSS measurements from separate APs, as $\left\{R^{(user)}_j : j \in \Omega_{AP}\right\}$, where Ω_{AP} is the set of heard AP indices j, the overall log-likelihood of the user position can be given as

$$L(x,y,z) = \sum_{j\in\Omega_{AP}} \log p(R_{j,k^{(z)}}|x,y), \text{ where}$$

$$p(R_{j,k^{(z)}}|x,y) = \frac{1}{\sqrt{2\pi\sigma^2_W}} \exp\left(-\frac{\left(R^{(user)}_j - \hat{\mathbf{G}}_{j,k^{(z)}}(T(x,y))\right)^2}{2\sigma^2_W}\right), \tag{10}$$

where $k^{(z)}$ is the index of the floor whose z coordinate is closest to the coordinate z. One often used approach to estimate the user position is the Maximum Likelihood (ML) principle, where the user position estimate is simply given as: $\left\{\hat{x}_{ML}, \hat{y}_{ML}, \hat{z}_{ML}\right\} = \arg\max_{x,y,z}(L(x,y,z))$. However, based on our studies, better results are obtained by taking the mean of the likelihood as

$$\hat{\mathbf{x}}_{MEAN} = \sum_{v\in\Omega_{area}} \frac{L(\mathbf{x}_v)}{\sum_{v\in\Omega_{area}} L(\mathbf{x}_v)} \mathbf{x}_v \tag{11}$$

where $\hat{\mathbf{x}}_{mean} = \left[\hat{x}_{mean}\ \hat{y}_{mean}\ \hat{z}_{mean}\right]$ is a vector of x, y and z coordinate estimates, Ω_{area} is a set of all grid point indices v in the target area where $L(x,y,z)$ is defined, and $\mathbf{x}_v$ is a vector including the coordinates of v^{th} grid point as $\mathbf{x}_v = \left[x_v\ y_v\ z_v\right]$.

In practice, the RSS images do not overlap perfectly, and therefore, when adding the likelihoods from different APs together, for each image we must define the value of the likelihood in locations where the image is not defined. Hence, we define the likelihood outside the image region as the average value of the likelihood inside the image subtracted with a penalty value of 6 dB (in logarithmic scale). This value was chosen based on a brief optimization, in which the positioning accuracy was shown to be rather insensitive to the chosen value.

IV. Positioning Results

We have collected data from 5 different multi-storey buildings. Four of the buildings are in Tampere, Finland, including one shopping mall, one office building, and two university buildings. The remaining building is a large shopping mall in Berlin, Germany. The measurements have been taken from 802.11 networks at 2.4 GHz carrier using a Nexus 7 tablet by ASUSTeK Computer Inc. (ASUS, Android 4.3.1 OS) with proprietary measurement software. In each measurement location (x,y,z), which is manually inserted by the measurer based on the building floor maps, the heard AP identities and their corresponding RSS values are stored. In addition to this, several independent user tracks for testing the positioning performance were collected. Here, we compare the positioning accuracy between the conventional fingerprinting, the compressed RSS image based approach, and the simple PL model based approach. In the PL model approach only the AP identity, AP coordinates, and a single-slope PL model parameter estimates [5], has to be stored in the database. To achieve a fair comparison between different approaches, we have used the same likelihood principle with all the approaches. Hence, with the conventional FP database we use the positioning method shown in [15], which is a counterpart for the likelihood method used with the compressed images, given in Eq. (10) and Eq. (11). Similarly with the PL models we exploit the same principle by using recreated RSS images discussed in [5].

The results are given jointly over all user tracks in all buildings considering the overall database size to be the sum of building-wise database sizes. The *mean FP modeling error,* the *mean 3D positioning error* and the *floor detection probability* are given in Fig. 2 as a function of database compression ratio $(1 - B_{IM}/B_{FP})$. Here, with the FP modeling error we refer

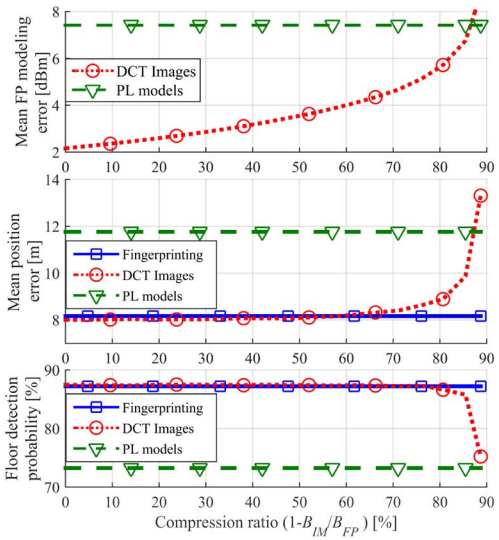

Fig. 2. The mean FP modeling error, mean 3D positioning error and floor detection probability as a function of database compression ratio.

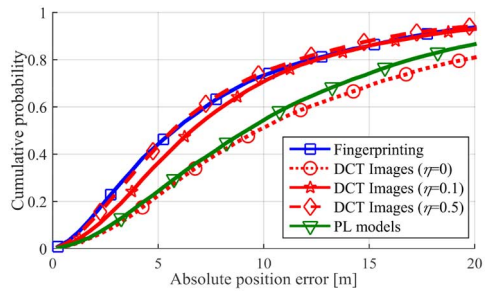

Fig. 3. Cumulative distribution of 3D positioning error.

to the error of the RSS values in each FP due to the learning database compression. From here it can be seen that the modeling error with the DCT images reduces along with the compression ratio, but the positioning error does not follow the same pattern, as it saturates around the compression ratio of 60%. This implies that the original FP are noisy and it is not necessary to aim at too accurate DCT images. Nonetheless, the RSS modeling error should be minimized, since it affects the variance of the likelihood function σ_W^2 given earlier in (10). In Fig. 2, different compression ratios are achieved by simply tuning the design parameter η (i.e., the ratio between the number of stored DCT coefficients and the number of original FPs in each image). The floor detection is assumed to be correct, if the estimated z coordinate is closer to the true user floor than to any other floor. It can be seen that **up to 80% compression ratio, the image based approach achieves comparable results with the FP approach.**

In Fig. 3, the cumulative distribution of the positioning error is given for the considered approaches with three different values of η as $\eta = 0$, $\eta = 0.1$ and $\eta = 0.5$, which result in compression ratios of 89%, 81% and 42%, respectively. Basically, with 81% compression ratio, we are almost able to reach the performance of the conventional fingerprinting.

V. Conclusion

We have proposed a novel approach on RSS-based indoor positioning by using RSS images and compressed spectral analysis. We showed how the size of the learning database can be considerably reduced by storing only the most significant coefficients from the DCT transformed images. The results showed that the proposed method can reach the performance of the conventional fingerprinting with a database size reduction up to 80% compression rate.

Acknowledgment

The authors would like to express their warm thanks to Jari Syrjärinne and Lauri Wirola for constructive cooperation and to the colleagues at TUT for helping with the measurements.

References

[1] S. A. Zakavat and R. M. Buehrer, Eds., *Handbook of Position Location: Theory, Practice and Advances*. Hoboken, NJ, USA: Wiley/IEEE Press, 2011.

[2] Y. Gu, A. Lo, and I. Niemegeers, "A survey of indoor positioning systems for wireless personal networks," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 13–32, Mar. 2009.

[3] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, and J. Sievanen, "A probabilistic approach to WLAN user location estimation," *Int. J. Wireless Inf. Netw.*, vol. 9, no. 3, pp. 155–163, Jul. 2002.

[4] A. Arzullaev, W. P. Park, and J. Hoyoul, "Accurate signal strength prediction based positioning for indoor WLAN systems," in *Proc. IEEE/ION Position Location Navigat. Symp.*, Monterey, CA, USA, 2008, pp. 685–688.

[5] S. Shrestha, J. Talvitie, and E. S. Lohan, "Deconvolution-based indoor localization with WLAN signals and unknown access point locations," in *Proc. Int. Conf. Localization GNSS*, Turin, Italy, 2013, pp. 1–6.

[6] S. Khodayari, M. Maleki, and E. Hamedi, "A RSS-based fingerprinting method for positioning based on historical data," in *Proc. Int. Symp. Perform. Eval. Comput. Telecommun. Syst.*, Ottawa, ON, Canada, 2010, pp. 306–310.

[7] J. Machaj, P. Brida, and R. Piché, "Rank based fingerprinting algorithm for indoor positioning," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Guimaraes, Portugal, 2011, pp. 1–6.

[8] C. Koweerawong, K. Wipusitwarakun, and K. Kaemarungsi, "Indoor localization improvement via adaptive RSS fingerprinting database," in *Proc. Int. Conf. Inf. Netw.*, Bangkok, Thailand, 2013, pp. 412–416.

[9] R. Kubota, S. Tagashira, Y. Arakawa, T. Kitasuka, and A. Fukuda, "Efficient survey database construction using location fingerprinting interpolation," in *Proc. IEEE 27th Int. Conf. Adv. Inf. Netw. Appl.*, Barcelona, Spain, 2013, pp. 469–476.

[10] N. Swangmuang and P. Krishnamurthy, "Location fingerprint analyses toward efficient indoor positioning," in *Proc. 6th Annu. IEEE Int. Conf. Pervasive Comput. Commun.*, Hong Kong, 2008, pp. 100–109.

[11] C. Xiaoyong and Y. Qiang, "Reducing the calibration effort for probabilistic indoor location estimation," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 649–662, Jun. 2007.

[12] V. Moghtadaiee, A. G. Dempster, and L. Binghao, "Accuracy indicator for fingerprinting localization systems," in *Proc. IEEE/ION Position Location Navigat.*, Myrtle Beach, SC, USA, 2012, pp. 1204–1208.

[13] H. Nurminen *et al.*, "Statistical path loss parameter estimation and positioning using RSS measurements in indoor wireless networks," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat.*, Sydney, NSW, Australia, 2012, pp. 1–9.

[14] A. Bazhyna, "Image compression in digital cameras," Ph.D. dissertation, Dept. Signal Process., Tampere Univ. Technol., Tampere, Finland, 2009.

[15] J. Talvitie, M. Renfors, and E. S. Lohan, "Distance-based interpolation and extrapolation methods for RSS-based localization with indoor wireless signals," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1340–1353, Apr. 2015.

[16] Z. Minghua, Z. Shensheng, C. Jian, and M. Haibin, "A novel indoor localization method based on received signal strength using discrete Fourier transform," in *Proc. 1st Int. Conf. Commun. Netw.*, 2006, pp. 1–5.

[17] F. Shih-Hau, L. Tsung-Nan, and L. Pochiang, "Location fingerprinting in a decorrelated space," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 5, pp. 685–691, May 2008.