# Accurate and Low-cost Location Estimation Using Kernels

**Jeffery Junfeng Pan**[a] **James T. Kwok**[a] **Qiang Yang**[a] **Yiqiang Chen**[b]
[a] Department of Computer Science, The Hong Kong University of Science and Technology, Hong Kong
[b] Institute of Computer Technology, Chinese Academy of Sciences, China
{panjf,jamesk,qyang}@cs.ust.hk, yqchen@ict.ac.cn

## Abstract

We present a novel method for indoor-location estimation using a vector-space model based on signals received from a wireless client. Our aim is to obtain an accurate mapping between the signal space and the physical space without incurring too much human calibration effort. This problem has traditionally been tackled through probabilistic models trained on manually labeled data, which are expensive to obtain. In this paper, we present a novel approach to building a mapping between the signal-vector space and the physical location space using kernel canonical correlation analysis (KCCA). Its training requires much less human labor. Moreover, unlike traditional location-estimation systems that treat grid points as independent and discrete target classes during training, we use the physical location as a continuous feedback to build a similarity mapping using KCCA. We test our algorithm in a 802.11 wireless LAN environment, and demonstrate the advantage of our method in both accuracy and its ability to utilize a much smaller set of labeled training data than previous methods.

## 1 Introduction

Location estimation is a major area of pervasive computing applications that range from context-dependent content delivery to people monitoring . However, accurate location information may not be always available. For example, many buildings cannot be reached by location-determination devices such as GPS. In such environments, a commonly-used technique is to leverage the signal strength values of a wireless network such as the IEEE 802.11b wireless LAN network (WLAN). Many systems utilize the signal strength values received from the access points to infer the location of mobile devices, based on deterministic or probabilistic techniques [Ladd *et al.*, 2002; Gentile and Berndt, 2004; Youssef *et al.*, 2003; Bahl and Padmanabhan, 2000; Ni *et al.*, 2003]. Similarly, in a sensor network, the location information must also be inferred from signals received from various deployed sensors. A major issue in these applications is how to build an accurate mapping between the signal space and the physical space while incurring minimal human effort.

In general, location-estimation systems using (RF)-based signal-strength values function in two phases: an *offline training phase* and an *online, localization phase*. In the offline phase, a *probabilistic model* is trained by considering the signal strength values received from the access points at selected locations in the area of interest. These values comprise the training data gathered from a physical region, which are used to calibrate a probabilistic location-estimation system. In the online localization phase, the real-time signal strength samples received from the access points are used to estimate the current location based on the learned model.

In the training phase, the learned model is essentially a mapping function $f(\mathcal{S},\mathcal{P})$ between the signal space $\mathcal{S}$ and the physical location space $\mathcal{P}$. Deterministic techniques build such a mapping function by assuming a signal propagation model based on properties of electro-magnetic fields. These signal-propagation models often cannot accurately describe a realistic environment due to uncertainty in the devices and the obstacles in the signal paths. Probabilistic techniques are often more accurate by directly handling the uncertainty. These techniques construct the function $f(\cdot,\cdot)$ by maximizing the likelihood of the observation given the probabilistic distribution of the training data used for calibration.

However, the probabilistic methods suffer from several major shortcomings. First, these techniques require a huge amount of labeled data for the trained model to be accurate. In reality, the data collection is a time-consuming process. In our experience, for example, in our earlier test it took many hours to collect and label the signal-strength data in a limited indoor environment. Second, the mapping between the signal and physical spaces are very difficult to construct due to nonlinearity. Nonlinearity exists when similar locations have very different signal signatures. Therefore, a direct mapping between the two spaces may not always work well, even when much data are collected. Third, many probabilistic techniques adopt a classification-based approach, by treating the locations as discrete classes to be classified. However, this approach ignores the important information inherent in the physical proximity of the locations. Furthermore, previous methods that apply regression based techniques treat the physical dimensions $(x,y)$ separately, they cannot fully exploit the geometric information imparted by considering $x$ and $y$ together.

In this paper, we present a similarity-based approach to building a mapping function. Our main intuition is to perform a kernel-based transformation of the signal and physical spaces to capture the nonlinear relationship between the signals and locations. Furthermore, we perform a kernel canonical correlation analysis (KCCA) in the two spaces for feature extraction. This allows the pairwise similarity of samples that are measured in both spaces to be maximally correlated. We apply a Gaussian kernel to adapt the noisy characteristic of signal strength and use a Matérn kernel to sense the change in physical locations.

A major advantage of our proposed technique is that we can obtain higher accuracy while reducing the training cost by requiring only a fraction of the labeled samples as compared to previous methods. We demonstrate this result in a series of tests on WLAN location estimation using the data collected in a realistic environment.

# 2 Related Work

## 2.1 Location Estimation

Location-estimation systems are an integral component for probabilistic plan-recognition systems in AI [Liao *et al.*, 2004; Bui *et al.*, 2002]. In general, they operate in two phases: an *offline training phase* and an *online localization phase*. During the offline phase, a *radio map* is built which tabulates the signal strength received from the access points at selected locations in the area of interest. During the localization phase, the signal strength samples received from the access points are used to search the radio map to estimate a user's location. In building the radio maps, two techniques have been attempted: deterministic and probabilistic.

Deterministic techniques rely on radio propagation models which take into account such environmental factors as path-loss functions [Roos *et al.*, 2002a; Kaemarungsi and Krishnamurthy, 2004]. When this is possible to do where uncertainty is relatively low, signal distribution models can be derived with a relatively small number of samples using knowledge on propagation channels [Hashemi, 1993]. This also enables the nearest neighbor or triangulation techniques to be used [Bahl and Padmanabhan, 2000; Smailagic and Kogan, 2002; Bhasker *et al.*, 2004]. For example, the RADAR system by Microsoft Research [Bahl and Padmanabhan, 2000] uses nearest neighbor heuristics and triangulation methods to infer a user's location. Each signal strength sample is compared against the radio map and the coordinates of the best matches are averaged to give location estimation. The accuracy of RADAR is about three meters with fifty percent probability. In [Bhasker *et al.*, 2004], an online procedure based on feedback from users was employed to correct the location estimation of the system. A similar technique for sensor networks has been applied; for example, the LANDMARC system [Ni *et al.*, 2003] utilizes the concept of reference tags to alleviate the effects caused by the fluctuation in RF signal strength. The method first computes the distance between the signal-strength vectors received from the tracking tags and those from different reference-tags' respectively. It then uses $k$ nearest reference tags' coordinates to calculate the approximate coordinate of the tracking tag.

Probabilistic techniques confront the uncertainty in an indoor wireless environment, by constructing the signal-strength distributions over different locations in the radio map and use probabilistic inference methods for localization [Roos *et al.*, 2002b; Youssef *et al.*, 2003]. [Youssef *et al.*, 2003] used a joint clustering technique to group locations together in order to reduce the computational cost of the system. The method first determines a most likely cluster within which to search for the most probable location, then applies a maximum likelihood estimation (MLE) method to infer the most probable location within the cluster.

More recently, intense research efforts have been spent to seek additional knowledge in order to boost the accuracy of the location determination systems. For example, to handle the sequential characteristic of user traces, [Ladd *et al.*, 2002] suggest a sensor fusion model and shows a strong correlation between consecutive samples. The robotics-based location sensing system in [Ladd *et al.*, 2002] applies Bayesian inference to compute the conditional probabilities over locations based on received signal-strength samples from various access points. Then a postprocessing step, which utilizes the spatial constraints of a user's movement trajectories, refines the location estimation and rejects the estimates that show significant changes in the location space. Depending on whether the postprocessing step is used or not, the accuracy of this method is 83% or 77% within 1.5 meters. [Yin *et al.*, 2005; Ni *et al.*, 2003] study the dynamic features of signal strength and build a regression model to adapt the change of radio map with reference points. A radio map is constructed using data mining methods in an effort to reduce the samples collected for calibration[Chai and Yang, 2005].

In contrast to the above works, our approach builds a similarity based mapping function by making full use of the continuous location information in kernel-based transformation. In the next section, we first review kernels and kernel canonical correlation analysis.

## 2.2 (Kernel) Canonical Correlation Analysis

**Canonical Correlation Analysis**

Given two sets of variables, canonical correlation analysis (CCA) [Hotelling, 1936] attempts to find a basis for each set such that the correlation between the projections of the variables onto these basis vectors are mutually maximized. Mathematically, given $n$ instances $S = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \ldots, (\mathbf{x}_n, \mathbf{y}_n)\}$ of the pair $(\mathbf{x}, \mathbf{y})$, CCA finds directions (*canonical vectors*) $\mathbf{w_x}$ and $\mathbf{w_y}$ so that the transformed variables (*canonical variates*) $(\langle \mathbf{w_x}, \mathbf{x}_1\rangle, \langle \mathbf{w_x}, \mathbf{x}_2\rangle, \ldots, \langle \mathbf{w_x}, \mathbf{x}_n\rangle)$ and $(\langle \mathbf{w_y}, \mathbf{y}_1\rangle, \langle \mathbf{w_y}, \mathbf{y}_2\rangle, \ldots, \langle \mathbf{w_y}, \mathbf{y}_n\rangle)$, are maximally correlated. Now, define the total covariance matrix by

$$\mathbf{C} = \hat{\mathbb{E}}\left[\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}'\right] = \begin{bmatrix} \mathbf{C_{xx}} & \mathbf{C_{xy}} \\ \mathbf{C_{yx}} & \mathbf{C_{yy}} \end{bmatrix},$$

where $\hat{\mathbb{E}}[\cdot]$ is the empirical expectation operator, $\mathbf{C_{xx}}, \mathbf{C_{yy}}$ are the within-sets covariance matrices and $\mathbf{C_{xy}} = \mathbf{C}'_{\mathbf{yx}}$ is the between-sets covariance matrix. It can be shown that $\mathbf{w_x}$ can be obtained by solving the generalized eigenproblem $\mathbf{C_{xy}}\mathbf{C_{yy}}^{-1}\mathbf{C_{yx}}\mathbf{w_x} = \lambda^2 \mathbf{C_{xx}}\mathbf{w_x}$, and then subsequently obtain $\mathbf{w_y}$ as $\mathbf{w_y} = \frac{1}{\lambda}\mathbf{C_{yy}}^{-1}\mathbf{C_{yx}}\mathbf{w_x}$ [Hardoon *et al.*, 2004].

**Kernel Canonical Correlation Analysis**

A major limitation of CCA is that it can only exploit linear relationships between $\mathbf{x}$ and $\mathbf{y}$. As is now well-known, the use of kernels offers an efficient, nonlinear extension of CCA. Kernel canonical correlation analysis (KCCA) [Hardoon *et al.*, 2004] implicitly maps $\mathbf{x}$ and $\mathbf{y}$ to $\phi_x(\mathbf{x})$ and $\phi_y(\mathbf{y})$, and then performs traditional CCA in the two high-dimensional feature spaces. To control the flexibility of the projections, the norms of the associated weight vectors are penalized.

Using the dual representations for the projection directions $\mathbf{w}_{\phi_x(\mathbf{x})}$ and $\mathbf{w}_{\phi_y(\mathbf{y})}$:

$$\mathbf{w}_{\phi_x(\mathbf{x})} = \mathbf{S}'_{\phi_x(\mathbf{x})}\boldsymbol{\alpha}, \quad \text{and} \quad \mathbf{w}_{\phi_y(\mathbf{y})} = \mathbf{S}'_{\phi_y(\mathbf{y})}\boldsymbol{\beta}, \tag{1}$$

where $\mathbf{S}_{\phi_x(\mathbf{x})} = [\phi_x(\mathbf{x}_1), \ldots, \phi_x(\mathbf{x}_n)]'$ $\mathbf{S}_{\phi_y(\mathbf{y})} = [\phi_y(\mathbf{y}_1), \ldots, \phi_y(\mathbf{y}_n)]'$, and $\boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbb{R}^n$. Denote the corresponding kernel functions by $k_{\mathbf{x}}(\cdot,\cdot)$ and $k_{\mathbf{y}}(\cdot,\cdot)$, and the kernel matrices (defined on all $n$ instances) by $\mathbf{K}_{\mathbf{x}}$ and $\mathbf{K}_{\mathbf{y}}$. It can be shown that $\boldsymbol{\alpha}$ can be obtained by solving the generalized eigenproblem

$$(\mathbf{K}_{\mathbf{x}} + \kappa\mathbf{I})^{-1}\mathbf{K}_{\mathbf{y}}(\mathbf{K}_{\mathbf{y}} + \kappa\mathbf{I})^{-1}\mathbf{K}_{\mathbf{x}}\boldsymbol{\alpha} = \lambda^2\boldsymbol{\alpha}, \tag{2}$$

where $\mathbf{I}$ is the identity matrix and $\kappa$ is a user-defined regularization parameter. Subsequently, $\boldsymbol{\beta}$ can be obtained as

$$\boldsymbol{\beta} = \frac{1}{\lambda}(\mathbf{K}_{\mathbf{y}} + \kappa\mathbf{I})^{-1}\mathbf{K}_{\mathbf{x}}\boldsymbol{\alpha}. \tag{3}$$

Given any $\tilde{\mathbf{x}}$, its projection on $\mathbf{w}_{\phi_x(\mathbf{x})}$ is given by

$$P_{\mathbf{x}}(\tilde{\mathbf{x}}) = \phi_x(\tilde{\mathbf{x}})'\mathbf{w}_{\phi_x(\mathbf{x})} = \mathbf{k}'_{\tilde{\mathbf{x}}}\boldsymbol{\alpha}, \tag{4}$$

where $\mathbf{k}_{\tilde{\mathbf{x}}} = [k_{\mathbf{x}}(\tilde{\mathbf{x}}, \mathbf{x}_1), k_{\mathbf{x}}(\tilde{\mathbf{x}}, \mathbf{x}_2), \ldots, k_{\mathbf{x}}(\tilde{\mathbf{x}}, \mathbf{x}_n)]'$. Similarly, the projection of any $\tilde{\mathbf{y}}$ onto $\mathbf{w}_{\phi_y(\mathbf{y})}$ is $P_{\mathbf{y}}(\tilde{\mathbf{y}}) = \mathbf{k}'_{\tilde{\mathbf{y}}}\boldsymbol{\beta}$, where $\mathbf{k}_{\tilde{\mathbf{y}}} = [k_{\mathbf{y}}(\tilde{\mathbf{y}}, \mathbf{y}_1), k_{\mathbf{y}}(\tilde{\mathbf{y}}, \mathbf{y}_2), \ldots, k_{\mathbf{y}}(\tilde{\mathbf{y}}, \mathbf{y}_n)]'$.

The generalized eigenproblem in (2) can be solved by using the (complete) Cholesky decomposition. However, the kernel matrices $\mathbf{K}_{\mathbf{x}}$ and $\mathbf{K}_{\mathbf{y}}$ are of size $n$, and so obtaining the Cholesky decomposition can become computational expensive for large training sets. In that case, the incomplete Cholesky decomposition or the partial Gram-Schmidt orthogonalization can be used instead [Hardoon *et al.*, 2004].

## 3 Location Estimation using KCCA

### 3.1 Motivation

Consider the two-dimensional location estimation problem[1]. Its goal is to obtain a mapping between the space of signal strengths obtained at $p$ access points $\mathcal{S} = \{\mathbf{s} \equiv [s_1, s_2, \ldots, s_p]' \in \mathbb{R}^p\}$ and the physical space $\mathcal{P} = \{\boldsymbol{\ell} \equiv [x, y]' \in \mathbb{R}^2\}$. Traditional approaches typically learn the two mappings, $\mathcal{S} \to \{x\}$ and $\mathcal{S} \to \{y\}$, independently [Brunato and Battiti, 2004]. In this paper, we consider $x$ and $y$ together and emphasize the *correlation* between signal and physical spaces, observing that the pairwise similarity in the signal space should match the pairwise similarity in the physical space. For example, in Figure 1, signal $S_A$ should be more similar to $S_B$ than $S_C$, since $A$ is closer to $B$ in the physical space. Consequently, we consider both $x$ and $y$ together and use KCCA to learn the mapping between the two spaces.

### 3.2 The LE-KCCA Algorithm

**Offline Training Phase**

In the training phase, the following steps are taken:

[1]Extension to the three-dimensional (or even higher-dimensional) case is straight-forward

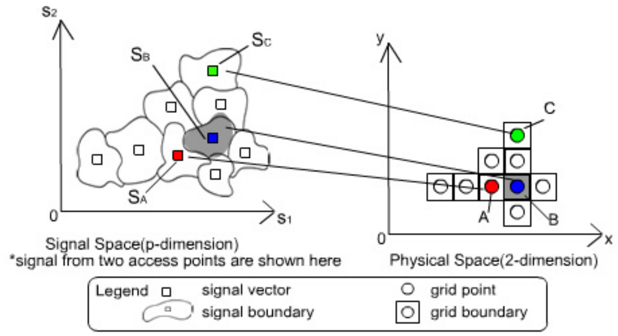


Figure 1: Correlation between the signal and physical spaces.

1. Signal strengths are collected at various grid locations.
2. KCCA, with appropriate choices for the two kernels (Section 3.3), is used to learn the relationship between the signal and physical spaces. In particular, $\lambda_i$'s and $\boldsymbol{\alpha}_i$'s are obtained from the generalized eigenproblem in (2), and the corresponding $\boldsymbol{\beta}_i$'s from (3).
3. For each training pair $(\mathbf{s}_i, \boldsymbol{\ell}_i)$, its projections

$$P(\mathbf{s}_i) = [P_1(\mathbf{s}_i), P_2(\mathbf{s}_i), \ldots, P_T(\mathbf{s}_i)]' \tag{5}$$

on the $T$ canonical vectors are obtained from (4).

**Online Localization Phase**

In the localization phase, the location of a new signal strength vector $\tilde{\mathbf{s}}$ is estimated as follows:

1. Use (4) to project $\tilde{\mathbf{s}}$ onto the canonical vectors and obtain $P(\tilde{\mathbf{s}}) = [P_1(\tilde{\mathbf{s}}), P_2(\tilde{\mathbf{s}}), \ldots, P_T(\tilde{\mathbf{s}})]'$.
2. Among the projections (5) from the training samples, find the $M$ neighbors closest to $P(\tilde{\mathbf{s}})$. In this paper, the weighted Euclidean distance $d_i = \sum_{j=1}^{T} \lambda_j (P_j(\tilde{\mathbf{s}}) - P_j(\mathbf{s}_i))^2$ is employed.
3. Interpolate these neighbors' physical locations to predict the physical location of $\tilde{\mathbf{s}}$. In this paper, we simply output the median (or mean for continuous location estimation) of the $(x, y)$ coordinates of these $M$ neighbors.

### 3.3 Choice of Kernels

Signal-strength values are often noisy. The empirical distribution (*radio map*) is difficult to obtain, especially when the training samples are scarce. As a first approximation, the Gaussian distribution has been used in characterizing the nonlinearity of the signal strengths [Roos *et al.*, 2002b]. Hence, in this paper, we also use the Gaussian kernel $G(\mathbf{x}_1, \mathbf{x}_2)$:

$$\exp(-w_G^2 \|\mathbf{x}_1 - \mathbf{x}_2\|^2) \tag{6}$$

for the $\mathcal{S}$ space. Here, $w_G$ is a user-defined parameter that reflects the smoothness of the radio map.

On the other hand, measurements of the physical locations are relatively clean. As in [Schwaighofer *et al.*, 2004], we adopt the Matérn kernel $M(\mathbf{x}_1, \mathbf{x}_2)$ (Figure 2):

$$\frac{2(\sqrt{\nu}w\|\mathbf{x}_1 - \mathbf{x}_2\|)^{\nu}}{\Gamma(\nu)} K_{\nu}(2\sqrt{\nu}w_M\|\mathbf{x}_1 - \mathbf{x}_2\|), \tag{7}$$

where $\nu$ is a smoothness parameter, $\Gamma(\nu)$ is the Gamma function and $K_{\nu}(r)$ is the modified Bessel function of the second kind. It can be shown that when $\nu \to \infty$, the Matérn kernel degenerates to the Gaussian kernel, and to the exponential kernel $\exp(-\sqrt{2}w\|\mathbf{x}_1 - \mathbf{x}_2\|)$ when $\nu = 0.5$. Note that both the Gaussian and Matérn kernels are isotropic and so are invariant to the location of origin and to arbitrary rotation.

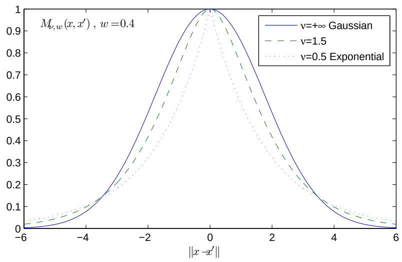


Figure 2: The Matérn kernel.

# 4 Experiments

In this Section, experiments are performed to measure the accuracy of LE-KCCA, and also test whether this is useful in reducing calibration effort. For comparison, we also run: (1) support vector machine (SVM) [Brunato and Battiti, 2004]; (2) maximum likelihood estimation (MLE) [Youssef *et al.*, 2003]; and (3) RADAR [Bahl and Padmanabhan, 2000].

## 4.1 Setup

Experiments are performed based on the office layout in the Department of Computer Science , The Hong Kong University of Science and Technology (Figure 3). There are four hallways with a total of 99 grids, where each grid measures 1.5m $\times$ 1.5m. There are only three access points in Figure 3, though a total of eight access points (including some from other floors) are detected. Each sample is thus an 8-dimensional signal strength vector, with the measurements averaged in one second. 100 such samples are collected at the center of each grid, with a total of 9,900 samples obtained. We randomly use 65% of the 9,900 samples for training and the rest for testing. To reduce statistical variability, results here are based on averages over 10 repetitions.

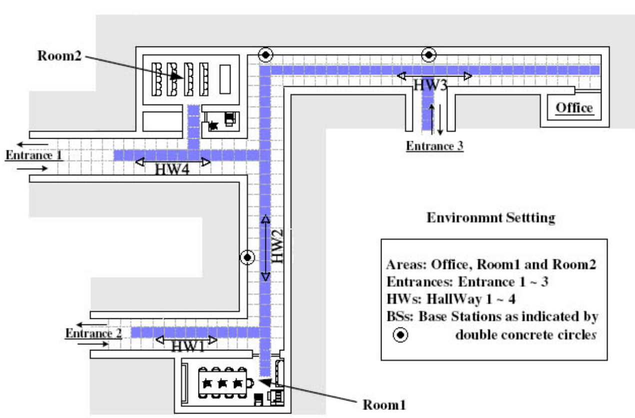


Figure 3: Office layout used in the experiments.

We set $w_G$ in (6) to 0.1, $w_M$ in (7) to 0.05, and $M$ (the number of neighbors used in the localization phase) to 17. Preliminary studies show that the accuracy only changes slightly (within 1%) when $M$ varies in the range 7 to 31. Moreover, $\nu$ in (7) is set to 0.5, and thus the Matérn kernel degenerates to the exponential kernel.

## 4.2 Location Estimation Accuracy

Figure 4 plots the average testing accuracies (and the corresponding standard deviations) at different acceptable error distances. At an error distance of 3.0m, the accuracy of LE-KCCA is 91.6% while those of SVM, MLE and RADAR are 87.8%, 86.1% and 78.8% respectively. Thus, by utilizing the pairwise distance similarities in physical locations, LE-KCCA obtains better performance than the other methods.

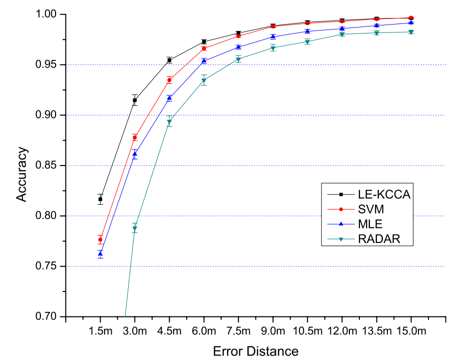


Figure 4: Testing accuracies at different error distances.

## 4.3 Reduction in Calibration Effort

In the first experiment, we use all the 99 grid locations but only a random subset of the signal samples available at each location for training. Figure 5 shows the testing accuracies at error distances of 1.5m or 3.0m. As can be seen, by using only 10-15 random training samples at each location, LE-KCCA can already outperform the other methods that use a full set.

In the second experiment, we select a subset of the available grid locations, and then use all the training signal samples at those selected locations for training. Evaluation is performed on both the unseen signal samples at the selected grid locations and also at the unseen locations. Again, Figure 6 shows that LE-KCCA yields better performance. Thus, LE-KCCA, by maximizing the correlation between signal and physical spaces, can reduce the calibration effort dramatically. With the physical location similarity as feedback, samples from the same locations move closer together, while those from different locations are pushed away under the feature-space mapping built by LE-KCCA.

We believe that LE-KCCA performs well. To validate this claim, we conduct another three experiments for comparison. The first one treats all the grid locations separately, which can be achieved by using a large value for $w_M$ in the Matérn kernel (in this experiment, we set $w_M = 50$). This is denoted as Isolated-KCCA in the sequel. Another extreme case is to modify the LE-KCCA algorithm so that the signal space is correlated with $x$ and $y$ *separately* (denoted 1D-KCCA). A third experiment is to modify both Gaussian and Matérn kernels to linear ones (Linear CCA). Experiments are performed in the same setting as for Figure 5. As can be seen in Figure 7, all three variants lead to degraded performance as expected.

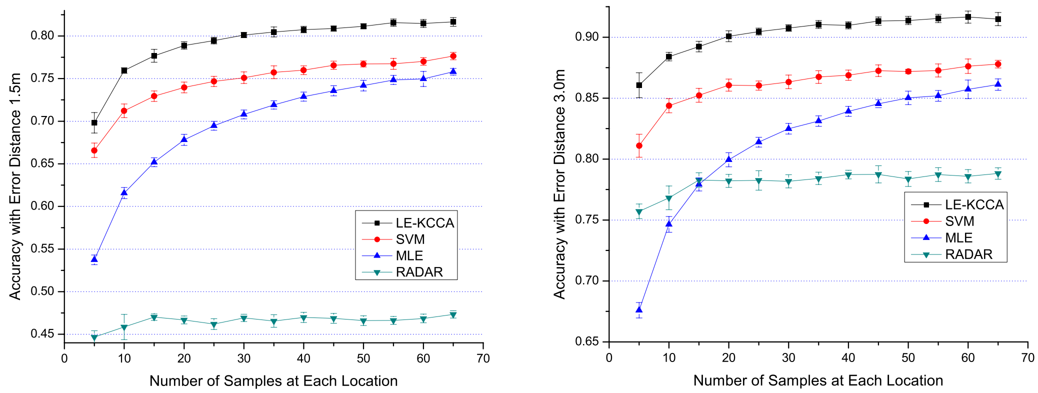


Figure 5: Effect on varying the number of training samples at each grid location.

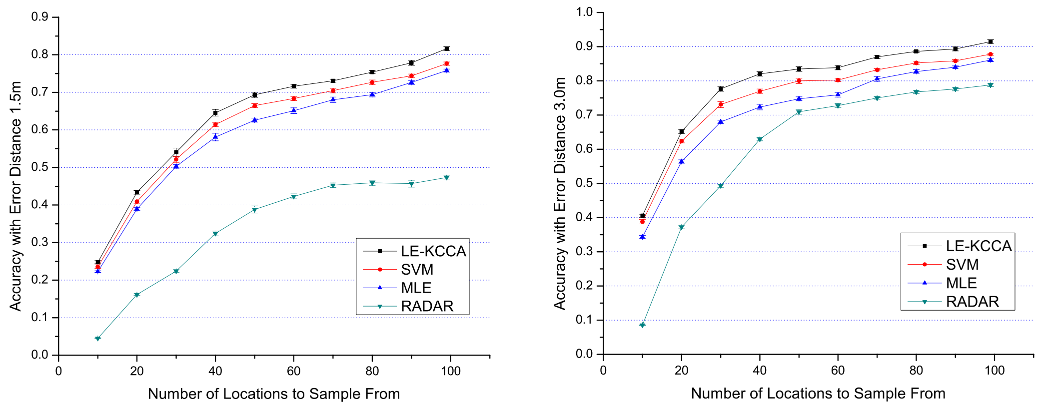


Figure 6: Effect on varying the number of grid locations in training.

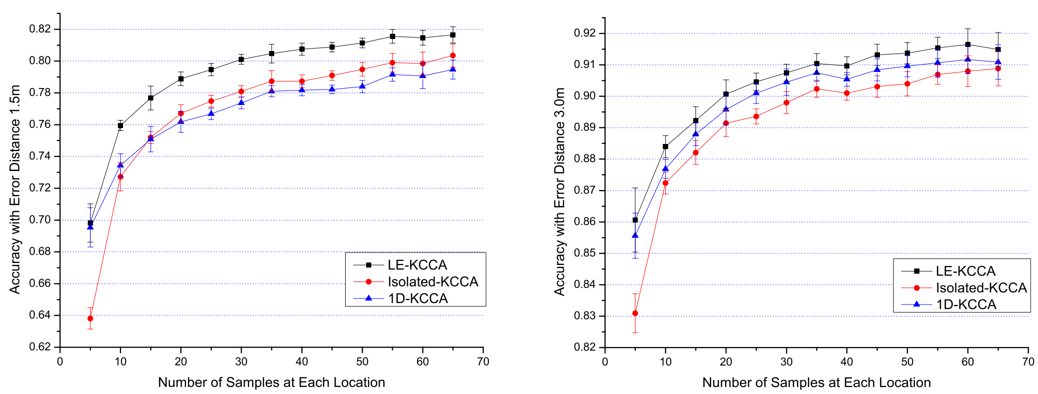


Figure 7: The effect when the information in both physical dimensions ($x$ and $y$) are not used together. Isolated-KCCA considers each grid location as independent, while 1D-KCCA considers $x$ and $y$ as independent. Linear CCA has even lower accuracy and the curve is not shown in the figure.

Linear CCA has an accuracy that is lower than 60% on every test set, and so its results are not shown in the figure.

## 5 Conclusions and Future Work

This paper focuses on using KCCA to improve the accuracy of location estimation in wireless LANs. We have found that kernel transformation and CCA allow the construction of an accurate mapping between the physical space and the signal space. One advantage is the higher accuracy obtained in localization with much less calibration effort. We use a Gaussian kernel for the signal space to adapt to the noisy characteristic of radio-propagation channels, and a Matérn kernel for the physical space. The advantage of KCCA is shown through extensive experimental tests in a realistic environment. In this paper, we constructed the mapping without considering the additional information of user motion profiles. In the future, we plan to take the user-motion profiles into account, which may lead to higher accuracy with even less calibration effort. Likewise, we wish to experiment more with other environmental and contextual factors to boost the performance.

### Acknowledgment

We would like to thank Hong Kong RGC for supporting this work under grant HKUST6187 / 04E. We would also like to thank Jie Yin and Xiaoyong Chai for their effort in collecting the data and discussing the research problems.

## References

[Bahl and Padmanabhan, 2000] P. Bahl and V.N. Padmanabhan. RADAR: An in-building RF-based user location and tracking system. In *Proceedings of the Conference on Computer Communications*, volume 2, pages 775–784, 2000.

[Bhasker *et al.*, 2004] E.S. Bhasker, S.W. Brown, and W.G. Griswold. Employing user feedback for fast, accurate, low-maintenance geolocationing. In *IEEE International Conference on Pervasive Computing and Communications*, pages 111–120, Orlando, Florida, USA, March 2004.

[Brunato and Battiti, 2004] M. Brunato and R. Battiti. Statistical learning theory for location fingerprinting in wireless LANs. *Computer Networks*, 2004.

[Bui *et al.*, 2002] H.H. Bui, S. Venkatesh, and G. West. Policy recognition in the abstract hidden Markov model. *Journal of Artificial Intelligence Research*, 17:451–499, 2002.

[Chai and Yang, 2005] X. Chai and Q. Yang. Reducing calibration effort for location estimation using unlabeled samples. In *Proceedings of the Third IEEE International Conference on Pervasive Computing and Communications*, Kauai Island, Hawaii, USA, 2005.

[Gentile and Berndt, 2004] C. Gentile and L.K. Berndt. Robust location using system dynamics and motion constraints. In *Proceedings of the IEEE International Conference on Communications*, pages 1360–1364, June 2004.

[Hardoon *et al.*, 2004] D.R. Hardoon, S. Szedmak, and J. Shawe-Taylor. Canonical correlation analysis; an overview with application to learning methods. *Neural Computation*, 16:2639–2664, 2004.

[Hashemi, 1993] Homayoun Hashemi. The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7):943–968, 1993.

[Hotelling, 1936] H. Hotelling. Relations between two sets of variates. *Biometrika*, 28:312–377, 1936.

[Kaemarungsi and Krishnamurthy, 2004] K. Kaemarungsi and P. Krishnamurthy. Modeling of indoor positioning systems based on location fingerprinting. In *Proceedings of the IEEE International Conference on Computer Communications*, volume 23, pages 1013–1023, 2004.

[Ladd *et al.*, 2002] A. Ladd, K. Bekris, G. Marceau, A. Rudys, L. Kavraki, and D. Wallach. Robotics-based location sensing using wireless ethernet. In *Proceedings of the Eighth ACM International Conference on Mobile Computing and Networking*, pages 227–238, Atlanta, Georgia, USA, September 2002.

[Liao *et al.*, 2004] L. Liao, D. Fox, and H. Kautz. Learning and inferring transportation routines. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence*, pages 348–353, San Jose, CA, USA, July 2004.

[Ni *et al.*, 2003] L.M. Ni, Y. Liu, Y.C. Lau, and A.P. Patil. LANDMARC: Indoor location sensing using active RFID. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, pages 407–416, Dallas, TX, USA, March 2003.

[Roos *et al.*, 2002a] T. Roos, P. Myllymäki, and H. Tirri. A statistical modeling approach to location estimation. *IEEE Transactions on Mobile Computing*, 1(1):59– 69, 2002.

[Roos *et al.*, 2002b] T. Roos, P. Myllymaki, H. Tirri, P. Misikangas, and J. Sievanen. A probabilistic approach to WLAN user location estimation. *International Journal of Wireless Information Networks*, 9(3):155–164, 2002.

[Schwaighofer *et al.*, 2004] A. Schwaighofer, M. Grigoras, V. Tresp, and C. Hoffmann. GPPS: A Gaussian process positioning system for cellular networks. In S. Thrun, L. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems 16*, Cambridge, MA, 2004. MIT Press.

[Smailagic and Kogan, 2002] A. Smailagic and D. Kogan. Location sensing and privacy in a context-aware computing environment. *IEEE Wireless Communications*, 9(5):10–17, October 2002.

[Yin *et al.*, 2005] J. Yin, Q. Yang, and L. Ni. Adaptive temporal radio maps for indoor location estimation. In *Proceedings of the 3rd Annual IEEE International Conference on Pervasive Computing and Communications*, Kauai Island, Hawaii, USA, March 2005.

[Youssef *et al.*, 2003] M. Youssef, A. Agrawala, and U. Shankar. WLAN location determination via clustering and probability distributions. In *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, pages 143–150, Fort Worth,Texas, USA, March 2003.