Practical Classification Methods for Indoor Positioning

Mahsa Honary^{*}, Lyudmila Mihaylova and Costas Xydeas

Lancaster University, InfoLab21, South Drive, Lancaster LA1 4WA, UK

Abstract: Location awareness is of primary importance in a wealth of applications such as transportation, mobile health systems, augmented reality and navigation. For example, in busy transportation areas (such as airports) providing clear, personalised notifications and directions, can reduce delays and improve the passenger journeys. Currently some applications provide easy access to information. These travel related applications can become context aware *via* the availability of accurate indoor/outdoor positioning. However, there are barriers that still have to overcome. One such barrier is the time required to set up and calibrate indoor positioning systems, another is the challenge of scalability with regard to the processing requirements of indoor positioning algorithms. This paper investigates the relationship between the calibration data and positioning system accuracy and analyses the performance of a k-Nearest Neighbour (k-NN) based positioning algorithm using real GSM data. Furthermore, the paper proposes a positioning scheme based on Gaussian Mixture Models (GMM). Experimental results show that the proposed GMM algorithm (without post-filtering) provides high levels of localization accuracy and successfully copes with the scalability problems that the conventional k-NN approach faces.

Keywords: Indoor positioning, Received Signal Strengths, asset tracking, informed traveller.

I. INTRODUCTION

Indoor positioning has been recently subject to an increased interest [1-5]. Many Location Prediction Engines (LPEs) which use Received Signal Strength Indications (RSSIs) for positioning rely on the a-priori collection of data for calibration. These calibration data sets are stored and later, using a classification methodology, are processed in real-time to determine location. An established classification technique is the k-nearest neighbour (k-NN) algorithm which has been proven to provide accurate results with a significant number of calibration data. However, the k-NN classification algorithm also has certain disadvantages. More specifically:

- *i*) k-NN does not associate probabilities to different clusters of data, which limits the effectiveness of post-filtering techniques, and
- k-NN processing time is intensive, especially when calibration sample sets are very large. This limits the applicability of k-NN based positioning algorithms on devices such as mobile phones and personal digital assistants (PDAs).

Apart from the k-NN approach, classification based positioning using Gaussian Mixture Models (GMM) is also considered in this paper and is extended to form of a novel GMM positioning method.

Thus this work investigates *i*) the relationship between the amount of calibration data and system positioning accuracy and *ii*) system scalability with respect to k-NN and a novel GMM based algorithm. The performance of the developed algorithms is validated based on real data and real world experiments, conducted in an office environment. In this framework, transceivers operating at 433 MHz frequency are attached to assets and to personnel. Also positioning algorithms are required to be computationally efficient so that they can operate in 433 MHz transceiver firmware. Hence, this work shows how well the two considered algorithms are suited for such real-time applications.

The accuracy of the developed algorithms is tested using Android mobile phones. This application enables the user to use the touch screen to select the correct room location. Locations are communicated using a wireless Local Access Network (LAN) (802.11.x) to a laptop where the RSSI data is collected and location stamped. An example is shown on Fig. (1). Collected data is used to build Probability Density Functions (PDF) prior to the system being operational (whereby the classification algorithm is used to identify the 433 MHz tag position).

The collection of 'training data' is an activity that is required by most Location Prediction Engines (LPEs) [5, 6]. Thus LPEs that use RSSI, usually need training data for both calibration and also to validate the system performance. However due to changing environments and consequently due to the changing transmission channel characteristics, the need for new calibration techniques often arises which in turn limits the usability of extensive time consuming data calibration collection algorithms.

In general, the amount of training data collected influences proportionally to the accuracy of the positioning system. However, large amounts of training data lead to significant computational complexity that in turn limits the algorithms' scalability. In this work experimentation has been carried out using 660 data samples collected per room in a given indoor office environment.

^{*}Address correspondence to this author at the Lancaster University, South Drive, Lancaster LA1 4WA, UK; Tel: +441524510388;

Fax: +44 1524 510493; E-mail: m.honary@lancaster.ac.uk

The remaining part of this paper is organized as follows. First in Section II, a brief description is given of the experimental set up. The key challenges to classification algorithms are outlined in Section III. Section IV describes the k-Nearest Neighbour approach. Next, the methodology underpinning the proposed novel GMM classification scheme is presented in Section V. Finally, experimental results are given in Section VI. Conclusions and suggestions for future work are provided in Section VII.

II. EQUIPMENT SET UP FOR EXPERIMENTS

The experimentation equipment set-up is illustrated in Fig. (1). Tags with 433 MHz transmission frequency communicate an identification number every second. This communication number is received and decoded at 3 beacons called Monitoring Units (MUs) and placed appropriately in the building. An infrared connection from the MUs is been used to forward observed RSSIs at the MUs *via* a laptop to a central PC where the location calculations are completed. This laptop would not be required in the proposed final embodiment where the algorithms may be implemented on a lightweight embedded platform (such as the MU).



Fig. (1). Set up of equipment for experiments.

It is important to note that an existing LPE was not used in these experiments, enabling full control and customisation of the positioning algorithms.

Fig. (2) illustrates the locations of the MU in the building layout (shown as red circles). The locations of the MU are chosen primarily based on considerations of convenience. Finding the optimal location for the MU is an interesting research challenge, and is not covered in this paper.

The calibration of the equipment is performed by a random walk in each room of the building. The accuracy of the solution is calculated based on the difference between the results from the classification algorithms from a reference location that has been taken by manual entry on the Android phone touch screen.

III. KEY CHALLENGES

One of the challenges that classification algorithms face is how informative the RSSI data are and how to use the RSSIs for decision making, especially how to differentiate between different locations. The plot shown on Fig. (3) illustrates the probability density functions (PDFs) for RSSI observed from one tag at one MU for random paths within two adjacent rooms.

The transmission frequency of the tag is 433 MHz (with a wavelength of 0.69236 m). This allows radio waves to penetrate effectively through walls. Note that the majority of the location specific RSSI differences are based upon constructive/destructive interference caused by reflections. This is clearly depicted in Fig. (3), where the variations in RSSIs within a single room are seen to be significant.



Fig. (2). Locations of the MU in building layout.

In order to identify the appropriate RSSI data characteristics which can differentiate between the two rooms, the following strategy is adopted:

- i. identify whether or not the tag is moving, and
- ii. if the tag is moving, pass the observed RSSI at the MU through a filter that could remove RSSI transients from short-range movements.

This results in a more compact PDF from which locationrelated information content may be more readily ascertained. An illustration of this 'filtered' PDF is provided in Fig. (4). This filtering approach applies a thresholding operation (based on an empirically set threshold for the standard deviations observed from the 5 most recent samples) to identify if the tag is moving. If the tag is moving a moving average filter is used. This approach enabled the classification algorithms to extract sufficient information content and thus provide reliable performance. The localization results are discussed in details in Section VI.



Fig. (3). PDF of RSSIs observed at one MU in two adjacent rooms.



Fig. (4). PDFs of RSSI observed at one MU, from two adjacent rooms after filtering.

IV. CLASSIFICATION

A) k-NN Classification

The k-NN Classification has been used to benchmark the performance of the GMM Classification. In the k-NN approach we assume there are M clusters with centroids:

$$Z = \{z_1, z_2, \dots, z_M\}$$

where each training sample is assigned to one of the clusters.

The goal with the k-NN is to minimise the total mean squared error between the training samples and their representative cluster centroids. This can be summarised in the formula shown below:

$$argmin\sum_{i=1}^{N} [x_i - z_i]^2$$

b) GMM Based Classification

The classification process can be performed with a mixture of Gaussian PDFs. The likelihood for the RSSIs is represented as a Gaussian mixture. The GMM technique enables small sets of collected data to represent sufficiently well large data sets.

The GMM has several properties that make this model ideally suited for modelling signal strength measurements. The GMM represents well:

• **Continuous locations**: GMMs do not require a discretised representation of an environment, or the collection of calibrated data at pre-specified locations. They are able to predict signal strength measurements at arbitrary locations.

• Arbitrary likelihood models: GMMs are regression models and able to approximate an extremely wide range of non-linear signal propagation models.

• **Correct uncertainty handling**: In contrast to other regression models, GMMs provide uncertainty estimates for predictions at any set of locations. This uncertainty takes into account the local density of calibration data and the noise of the data points.

• **Consistent parameter estimation**: The parameters of GMMs can be learned from the calibration data *via* hyper-parameter estimation. These parameters include the spatial correlation between measurements and the measurement.

The GMM [9] approach uses an offline clustering process, followed by an online Maximum Likelihood Estimation (MLE) process. Various efficient implementations of the GMM algorithm have been proposed, including [10,11].

The offline process utilizes the maximized likelihood function and applies the GMM algorithm recursively until a converged mean, covariance and weight for each class of the Calibration data is satisfied. This process can be summarized by the following 5 steps.

- i. Obtain a model of the signal properties at various locations
- ii. Generation of the PDF for each room.
- iii. Get the initial values for mean and covariance.
- Apply a mixture of Gaussian models for the calibration data given the number of classes and their parameters.
- v. Maximize the likelihood function and apply the GMM algorithm recursively until we get the final and converged mean, covariance and respective weight for each class.

In the GMM applied in this paper, each location x can be viewed as arising from a super population G which is a mixture of a finite number, k, of populations $G_1, ..., G_K$ in proportions $\pi_1, ..., \pi_K$, respectively where

$$\sum_{i=1}^{K} \pi_i = 1 \text{ and } \pi \ge 0$$

The PDF of an observation x (of dimensionality d) in the finite mixture is

$$p(x;\theta) = \sum_{i=1}^{K} \pi_i \cdot \frac{1}{(2\pi)^{d/2} \left|\sum_i\right|^{1/2}} exp^{-\frac{1}{2}(x-\mu_i)^T(\Sigma_i) - 1(x-\mu_i)}$$

where π_i is the mixing parameter, $p_i(x;\theta)$ is the PDF corresponding to G_i , and θ is the vector of all unknown parameters associated with the parametric forms adopted by k component densities. The vector θ consists of elements of the mean vectors μ_i and the elements of the covariance matrices \sum_i for i = 1,..., K. The vector $\theta = (\pi^T, \theta^T)^T$ of all unknown parameters belongs to some parameter and is estimated using the following EM algorithm, which forms the closed form negative log-likelihood function:

$$E = -\mathrm{In}(L) = -\sum_{j=1}^{n} \mathrm{In}p(x_j) = -\sum_{j=1}^{n} (\sum_{i=1}^{K} \pi_i p(x_j | i))$$

This can be regarded as the error function, and maximising the likelihood L is therefore equivalent to minimising E.

The EM algorithm aims to find a θ that maximises $p_i(x; \theta)$. For this purpose we introduce the function:

$$Q(\theta'|\theta) = E(\log p(x|\theta')y,\theta)$$

which is the expected value of log $p_i(x; \theta)$ according to the conditional distribution of x, given y and parameter θ . The expectation is assumed to exist for all pairs (θ', θ) .

The EM algorithm steps can be summarised by the E-step and the M-step, which can be summarised by the following equations:

E-step:

$$p_{i,k} = \frac{a_k^{(p)} \theta(x_i | \mu_k^{(p)}, \sum_{k=1}^{(p)} \beta(x_i | \mu$$

M-step:

$$a_{k}^{(p+1)} = \frac{\sum_{i=1}^{n} \rho_{i,k}}{n}$$

$$\mu_{k}^{(p+1)} = \frac{\sum_{i=1}^{n} \rho_{i,k} x_{i}}{\sum_{i=1}^{n} \rho_{i,k}}$$

$$\sum_{k} e^{(p+1)} = \frac{\sum_{i=1}^{n} \rho_{i,k} (x_{i} - \mu_{k}^{(p+1)}) (x_{i} - \mu_{k}^{(p+1)})}{\sum_{i=1}^{n} \rho_{i,k}}$$

where *p* is the iteration number, $P_{i,k}$ is the posterior probability function; that give the probability of Y=k, where k=1,...,K.

 $))^{t}$

These steps are repeated until convergence is achieved.

C) Novel GMM Classification Method

The online classification process uses the optimal parameters calculated during the offline process to identify the probability to which the real time data samples belong to the GMM clusters. The probabilities are calculated through an expectation maximization function, as in [5], through the EM function shown above. The greatest probability indicates the cluster for which it is most-likely that the sample belongs to.



Fig. (5). RSSI from 3 rooms observed at 3 MU.

To illustrate the approach we consider data from 3 rooms, which are shown using green, blue and red circles. As illustrated on Fig. (5), each transmission from the tag is received by 3 MUs and displayed in a three dimensional space. Please note that for simplicity only 3 MUs are considered in this example. In general, given N MUs, data is clustered in an N-dimensional space.

In the conventional GMM approach, each room is associated with a Gaussian distribution, as illustrated in Fig. (6). Classification of an online sample can then be calculated from a distance calculation in order to determine the most likely cluster which the online sample belongs to.



Fig. (6). Clusters formed from conventional GMM.

The novel online GMM algorithm proposed in this paper extends the offline algorithm from a different viewpoint. Generally, the conventional GMM offline algorithm attempts to classify clusters for each room in which data is collected. For instance, in one of the implementations, if there are 15 rooms (N=15), the data is split to 15 classes. The proposed scheme allows for an initial "over segmentation" of the data in to M classes where as a rule of thumb $M \ge 2N$. The value of M is environment dependent.

Thus, the collected data are classified into these M clusters (without specifying that clusters should be formed from data, within specific rooms). Different clusters may contain data which are collected from different rooms. Fig.

(7) illustrates that the data points are clustered purely on their physical characteristics. In this example M=5.



Fig. (7). M Clusters are formed in the novel GMM approach.

Note that within some clusters, almost all the samples may belong to one specific room. In such cases, if in the online process, real-time data is situated in that cluster a high probability may be given to the algorithm's ability to estimate the correct room. However, some clusters may consist of data collected from many different rooms.

In Fig. (8), it is shown that each Gaussian distribution has a degree of membership to possible clusters (in our case room locations). If we consider the cluster labelled '1', if an online sample falls into this cluster we are able to state that there is:

5/6, 83.3%, probability that is belongs to Green;

1/6, 16.7%, probability that is belongs to Blue;

0/6, 0%, probability that is belongs to Red.



Fig. (8). Considering Clusters '1'.

If the probabilities for the degree of membership do not satisfy a threshold condition, the GMM process is repeated on this cluster, and smaller clusters are formed.

This is an iterative process and is applied until the majority (in our case 90%) of the data samples within a cluster belong to one room. An example of this formation of sub clusters is shown on cluster '1' in Fig. (9).



Fig. (9). GMM iteration on cluster 1.

V. EXPERIMENTAL RESULTS AND OBSERVATIONS

A) Investigation of Calibration vs Accuracy for k-NN Approach (with PDF Filtering)

First, system calibration is completed by the 'calibrator' walking randomly within a specified location (such as a particular room) with the 433MHz tag around his/her wrist. Next, the calibration measurements are processed and used by the classification algorithms. The dependence of the number of calibration samples with respect to the system room location accuracy is shown in Fig. (10). The accuracy in Fig. (10) is given as a percentage of the correct room been identified.



Fig. (10). Dependence of k-NN accuracy based upon calibration samples taken for each location.

It is evident from Fig. (10) that with more than 600 calibration samples the k-NN algorithm is able to correctly identify in which room of the building the tag is situated. An important factor determining the system performance is the delay (or latency) introduced in the decision making process. The impact of the delay on the algorithm performance, however, is application dependent. In many applications, such as positioning of luggage in an airport, we expect that a delay of several seconds will not impact the decision making process. Such a delay, however, would be significant for applications such as passenger navigation aids.

Fig. (11) shows this 'latency' graphically, and provides an example where for the first 2-3 seconds after entering a 'new' room the 'old' room is still identified.



Fig. (11). Illustration of latency in room identification as user enters another room space.

The result shown on Fig. (12) is obtained by identifying the latency, in particular, based on the transitions from one room to another and the time taken for the classification algorithms to identify the 'new' room. Fig. (11) indicates that in 3 samples the correct room is identified 90% of the time.



Fig. (12). Time taken for to identify room transition for k-NN classification.

The Cumulative Distribution Function (CDF) shown on Fig. (13) characterises the accuracy of the k-NN algorithm obtained by using 660 calibration samples per room. This CDF is generated by observing the true distance of the tag from the location estimate which the classification algorithm provided.

This result implies that system performance is comparable, if not better than many existing systems operating at higher communication frequencies (e.g. 802.11.x positioning solutions), and also with higher RSSI reporting frequencies (often a fraction of a second) [7, 8].

The following accuracy measures are used for the k-NN based system: the Circular Error Probability (CEP) (50%) of 2m, CEP (67%) of 2.7 m, CEP (95%) of 4.8 m. Hence, it has

been observed that 50% of the location estimates will be within 2m of the actual location. Only 1 in 3 of the estimated locations will be further than 2.7m away from the actual location; and only 1 in 20 estimated locations will have an error in excess of 4.8m.



Fig. (13). CDF illustrating accuracy of k-NN Classification Algorithm.

B) Comparison Between k-NN and GMM Classification Approaches

The complexity of the k-NN algorithm is directly proportional to the number of calibration samples multiplied by the number of rooms. Hence, for example, with 15 rooms and 500 samples in each room, 15x500 samples are required to be compared every time a decision is made. Thus for the k-NN approach the number of samples and/or the number of rooms increase the system complexity.

On the other hand, in the conventional GMM classification algorithm each room is approximated by a Gaussian distribution and therefore only 15 distributions need to be compared, which results in a much more efficient process. In the proposed GMM implementation more distributions need to be compared (due to the sub-clusters). Again, since the numbers of these sub-clusters are also relatively low a significant saving in complexity has been made.

Based on the processing time taken from Matlab, one can conclude that the conventional GMM algorithm provides on average a 68-fold improvement in terms of speed over the k-NN approach. The proposed GMM process provides on average a 44 times improvement in speed over the k-NN approach. This significant enhancement with respect to the processing speed is particularly important for cases where the LPE has to be implemented on portable devices (such as a mobile phone) where the battery life and processor capability are of primary concern. However, since the GMM approach can be considered as a form of lossy compression and the problem is that by reducing the complexity of calculation localized information may be lost which in turn degrades system localization accuracy.

Table 1 highlights the relative localization accuracy observed using these different classification methods.

Fig. (14) provides a CDF based comparison for the k-NN (blue line), modified GMM (green dashed line) and conventional GMM (red line) systems. Table 2 presents a comparison of the computational accuracy of the k-NN approach compared with the conventional GMM and the

proposed GMM algorithm. In our experiments, the k-NN classification provided the best results where k=5.

 Table 1.
 Classifier
 Size and
 Processing
 Time for
 Each

 Classification
 Method

| Classification Method | Classifier Size | Processing Time (ms) |
|-----------------------|-----------------|----------------------|
| k-NN (k=5) | 9900 | 1.428 |
| ConvGMM | 15 | 0.021 |
| Novel-GMM | 26 | 0.036 |

 Table 2.
 Accuracies
 Observed
 with
 Each
 Classification

 Methods

| | Circular Error Probability (CEP) | | |
|----------------|----------------------------------|-------|--------|
| Classification | P50 | P67 | P95 |
| k-NN | 2.00m | 2.62m | 4.70m |
| ConvGMM | 4.56m | 5.64m | 11.82m |
| Proposed-GMM | 2.07m | 2.71m | 4.94m |



Fig. (14). CDF comparing accuracy of classification algorithms.

Fig. (14) shows that the modified GMM (green line) performs significantly better than the conventional GMM approach (red line).

VI. DISCUSSION, CONCLUSIONS AND FUTURE WORK

Existing indoor positioning systems that employ classification algorithms are primarily based on the 802.11.x standard. Such systems [7, 8] operate at a much higher frequency band (2.4GHz) than the 433 MHz transmitters considered in this paper.

Hence, the transmitters used in this research have a much greater wavelength, and therefore their respective RSSIs are not absorbed to the same extent as in 802.11.x, and this increases the challenge in providing accurate positioning estimates.

While accepting this adverse effect, experiments have shown that both the k-NN approach and the proposed GMM algorithm can provide equally accurate results using tags transmitting at 433 MHz. Furthermore, these results have illustrated the importance of localized data which can associate the dependencies between sets of readings (from multiple antennas) within localized regions. Due to the spread of RSSI readings associated with a single room it can be concluded that classification techniques which attempt to associate a distribution to calibration readings, such as conventional GMMs, do not perform well with positioning systems operating at 433 MHz.

Importantly, the proposed GMM approach also offers two important benefits, (i) it provides a probability to each room, which may be used within post-filtering activities; and (ii) the online GMM approach is extremely quick in comparison with the k-NN approach.

The fact that the novel GMM implementation introduced in this paper has been able to provide accuracy that is significantly better than that of the conventional GMM can be largely attributed to the additional localised information taken into account *via* the sub clusters.

Future work has been considered through the more intelligent use of building layouts. The building layout is currently used for the identification of:

- Room polygons
- Intra-floor joins (i.e. doors between rooms)
- Inter-floor joins (i.e. stairs and lifts between floors)

In this current embodiment, room polygons are currently used solely for display purposes within a Graphical User Interface.

For Intra-floor joins and Inter-floor joins post filtering can be used to improve the performance given by the positioning algorithms. The locations of these joins identify practical physical transitions for a user's location. Filtering techniques may use Markov or Bayesian models to utilise this information to correct errors arising from jumps to locations which are considered unpractical.

In order to reduce the calibration process, two propagation models specific to the 433MHz frequency have also been investigated. One possible solution consists in estimating the path loss based on the Obstructed Line of Sight whereas another solution identifies the Most Dominant Path. Both solutions are favourable than the conventional ray-tracing and ray-launching approaches since dependencies upon exact building dimensions, and processing time can be optimized. Preliminary results from these propagation models have been observed. However the models require significant refinements which have not been completed to date. From these propagation models, additional further work can be formulated that allows MU locations to be optimally estimated for maximum localization system performance.

ACKNOWLEDGEMENTS

The authors acknowledge the support of HW Communication Ltd and EPSRC for the support under grant number CAN/05/105.

CONFLICT OF INTEREST

Neither of the authors have a conflict of interest to declare in relation to the production of this manuscript.

REFERENCES

- K. Pahlavan, F. Akgul, Y. Ye, T. Morgan, F. Alizadeh-Shabodz, M. Heidari and C. Steger, "Taking Positioning Indoors. Wi-Fi Localization and GNSS", *Inside GNSS*, pp. 40-47, May, 2010.
- [2] H. Liu, H. Darabi, P. Banerjee and J. Liu, "Survey of wireless indoor positioning techniques and systems", *IEEE Trans. Syst. Man Cybern. C Appl. Rev.*, vol. 37, no. 6, pp. 1067-1080, 2007.
- J. Zhou and J. Shi, "RFID localization algorithms and applications - a review", J. Intell. Manuf., vol. 20, no. 6, pp. 695-707, 2009.
- [4] V. Algeier, B. Demissie, W. Koch and R. Thomä, "State space initiation for blind mobile terminal position tracking", *EURASIP J. Adv. Signal Process.*, vol. 2008, p. 394219, 2008.
- [5] M. A. Youssef, A. Agrawala, S. Noh, "A probabilistic clusteringbased indoor location determination system", *Technical Report CS*-

Received: May 27, 2012

Revised: June 20, 2012

Accepted: June 24, 2012

© Honary et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/ 3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.

TR-4350 and UMIACS-TR-20002-30, University of Maryland, 2002.

- [6] P. Bahl and V. N. Padmanabhan, "RADAR: an in-building RFbased user location and tracking system" *Proc. IEEE INFOCOM* 2000, vol. 2, pp. 775-784, 2000.
- [7] M. Youssef, A. Agrawala and S. Udaya, "WLAN location determination via clustering and probability distributions". Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom), Fort Worth, Texas, 2003.
- [8] K. Prasithsangaree and P. K. Chrysanthis, "On Indoor Position Location with Wireless LANs", Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2002). Lisbon, Portugal, 2000.
- [9] I.D. Dinov, "Expectation Maximization and Mixture Modeling Tutorial", California Digital Library, Statistics Online Computational Resource, Paper EM MM, Dec. 2008.
- [10] A. Tashk, A. Sayadiyan, P. Mahale, M. Nazari, "Pattern Classification using SVM with GMM Data Selection Training Method", *Proceedings* of the IEEE International Conference on Signal Processing and Communications, 2007.
- [11] M. Shi, A. Bermak, "An Efficient Digital VLSI Implementation of Gaussian Mixture Models-Based Classifier", *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 14, no. 9, pp. 962-974, 2006.