An Automatic Approach for Generation of Fuzzy Membership Functions

Hossein Pazhoumand-Dar, Chiou Peng Lam(✉), and Martin Masek

School of Science, Edith Cowan University, Perth, Australia
{h.pazhoumanddar, c.lam, m.masek}@ecu.edu.au

Abstract. Eliciting representative membership functions is one of the fundamental steps in applications of fuzzy theory. This paper investigates an unsupervised approach that incorporates variable bandwidth mean-shift and robust statistics for generating fuzzy membership functions. The approach automatically learns the number of representative functions from the underlying data distribution. Given a specific membership function, the approach then works out the associated parameters of the specific membership function. Our evaluation of the proposed approach consists of comparisons with two other techniques in terms of (i) parameterising MFs for attributes with different distributions, and (ii) classification performance of a fuzzy rule set that was developed using the parameterised output of these techniques. This evaluation involved its application using the trapezoidal and the triangular membership functions. Results demonstrate that the generated membership functions can better separate the underlying distributions and classifiers constructed using the proposed method of generating membership function outperformed three other classifiers that used different approaches for parameterisation of the attributes.

Keywords: Fuzzy membership functions · Variable bandwidth mean-shift · Fuzzy logic · Activities of daily living · Abnormality detection · Robust statistics

1 Introduction

Eliciting representative membership functions (MFs) for data is one of the fundamental steps in applications of fuzzy theory as the success of many fuzzy approaches depends on the membership functions used. However, there are no simple rules, guidelines, or even consensus among the community on how to choose the number, type, and parameters of membership functions for any application or domain [1]. Several methods for the automatic generation of MFs have been proposed in the literature and the choice of function has been linked to the problem and the type of data available. However, in most of these techniques, the number of fuzzy sets has to be provided empirically. Furthermore, the range for membership functions generated by many existing techniques does not address the impact of outliers and noisy measurements in data.

In this paper, we propose a hybrid approach that incorporates variable bandwidth mean-shift (VBMS) and robust statistics for automatic generation of representative

J. van den Herik and J. Filipe (Eds.): ICAART 2016, LNAI 10162, pp. 247–264, 2017.
DOI: 10.1007/978-3-319-53354-4_14

MF(s) for an attribute. The analysis of the underlying data distribution is unsupervised as the proposed approach first determines the number of modes from the probability density function (PDF) and then uses this value as the number of clusters for a multimodal data distribution. The approach overcomes the problems associated with some of the existing approaches by

- determining the number of representative MFs for the attribute from the underlying data distribution automatically
- automatically handling noise and outliers in the attribute feature space.

The rest of this paper is organised as follow: Sect. 2 briefly reviews relevant literature on MF generation techniques. Section 3 describes the proposed approach. The experimental evaluation of our technique is presented in Sect. 4 followed by conclusions in Section 5.

2 Background

Many techniques have been proposed to generate fuzzy membership functions from an attribute. Three questions that have to be addressed are: (1) how many fuzzy sets should be defined to represent the attribute, (2) what type of membership function can represent the attribute better, and (3) how to determine the parameters associated with the adopted membership functions.

Typically, techniques have used manual partitioning of attributes, mostly based on expert knowledge, and adopted a pre-determined number of membership functions [2] to partition the data space for the attribute by (usually evenly-spaced) MFs. However, manual approaches suffer from the deficiency that they rely on subjective interpretations from human experts.

Given a labelled dataset, evolutionary methods can also be utilized to generate MFs. Moeinzadeh et al. [3] applied genetic algorithm (GA) and particle swarm optimization for the adjustment of MF parameters to increase degree of membership of data to their classes for classification problems. Authors in [4] applied GA for evolving parameters associated with MFs in a fuzzy logic controller for a helicopter. Initial guesses for the MFs are made by the expert and the GA adjusts the MF parameters to minimise the movement of a hovering helicopter. In the classification method proposed by Tang et al. [5], a fitness function quantifies how well the crisp values of attributes are classified into MFs and the GA process evolves over time by searching the best set of MF parameters which optimises result of the fitness function.

Takagi and Hayashi [6] also proposed the use of artificial neural networks (ANN) for the construction of membership functions. Once raw data are clustered into a specific number of groups, ANN is applied to the clustered data to determine the parameters associated with membership functions.

However, for situations where the training data is not labelled, MF generation techniques generally involve unsupervised clustering of data using a specific distance measure and then the parameters of detected clusters (mean, variance, etc.) are used to generate MFs. For example, techniques [7] have used the Fuzzy C-Means (FCM) clustering algorithm [8] to cluster a particular attribute into specific number of clusters. Cluster boundaries and the location of the centre were then used to determine the cluster membership function parameters. Doctor et al. [9] presented a fuzzy approach to model an occupant behaviour in a residential environment. They used a double clustering technique [10] combining FCM and agglomerative hierarchical clustering for extracting a predefined number of MFs from the user's recorded input/output data.

The disadvantage associated with most of these methods is that the number of fuzzy sets must be predefined. However, we usually do not know an optimal number of representative MFs for a particular attribute. In addition, outliers in data are included in range of MFs generated by many of these techniques. New robust techniques that can determine number of representative MFs automatically would address some of these limitations.

3 The Proposed Approach

For each attribute the approach automatically defines a number of associated MFs as linguistic variables. Let an attribute take a series of crisp numerical values x_n $(n = 1, \ldots, N)$ and these data points belong to an unknown probability density function (PDF) f. The two-step procedure of the proposed approach for generating MFs for the attribute is as follows

Step 1. use VBMS to find modes (local maxima) of f representing the attribute and cluster of data points associated with each mode
Step 2. use skewness adjusted boxplot (SAB) technique [11] to obtain the normal range of data for each cluster (where there are no outliers), and subsequently determine the parameters associated with a specific MF for the cluster.

VBMS proposed by Comaniciu et al. [12] is a nonparametric clustering technique which does not require the number of clusters to be defined. It takes multidimensional data with an unknown density f and estimates the density at each point by taking the average of locally-scaled kernels centered at each of the data points, and tries to map each data point to its corresponding mode. The output of this technique is locations of modes detected in f and the cluster of data associated with each mode. Usually the kernel K is taken to be a radially symmetric, nonnegative function centered at zero such that $K(x) = k(\|x^2\|)$. Details associated with this technique can be found in [13].

The SAB technique is a graphical tool (with a robust measure of skewness) used in robust statistics (RS) for the purpose of outlier detection [14]. Given a continuous unimodal data, SAB first calculates a robust measure of skewness (i.e., medcouple (MC) [15]) of the underlying data distribution. Then it outputs a normal range for the data which excludes possible outliers from the normal data. Details associated with its use in this approach can be found in [13].

When an attribute has a multimodal PDF and each mode may be associated with a different density distribution, one fixed global bandwidth is not optimal for estimating the location of modes in PDF, and thus local bandwidths should be computed [12]. Using VBMS, we determine a local bandwidth for each data point in a way that points corresponding to tails of the data distributions receive a bigger bandwidth than data points lying in large density region of distributions and hence the estimated density function for tails of the distributions is smoothed more. The output of Step 1 is the location of modes of f denoted as $m_{i(1 \leq i \leq N)}$ and the cluster of data associated with each mode.

In Step 2, we use the output from Step 1, the number of modes as the number of required MFs representing the attribute and for each cluster of data associated with a mode, we define a MF. We first use the SAB technique to determine the normal range (NR) for the cluster and we denoted this as $[l, h]$. The output of Step 2 for each attribute is a tuple (X, m_1, m_2, …, m_{nc}) as linguistic variables, where X stands for the attribute name and m_i stands for an MF defined over the universe of discourse for the attribute and nc stands for the number of modes identified in Step 1.

The proposed approach can be employed to determine the parameters associated with different forms of MFs to characterise the identified clusters. In this paper we demonstrate how the approach can be applied on two different types of MF, namely the triangular and trapezoidal membership functions. These two MFs are selected because of their simplicity of calculation and ability to represent skewed distributions.

3.1 Generating Triangular Membership Functions

As shown in Fig. 1, parameters of triangular MFs are defined by a triad (A, B, C), with point A representing the left foot of triangular MF, B is the location of the center, and C is the location of the right foot.

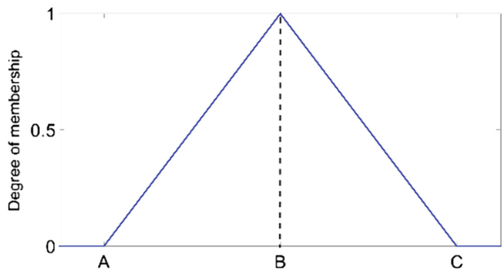
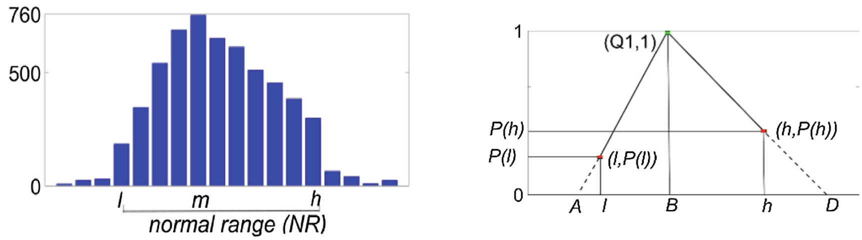


Fig. 1. An example of a triangular membership function defined by a triad (A, B, C).

To define a triangular MF for a detected cluster we use NR $[l, h]$ (l is the lower and h is the higher limit for the normal range, respectively) associated with the cluster, and the cluster mode, m, to determine its parameters (A, B, C).

This is illustrated using an example shown in Fig. 2. Figure 2(a) showed the histogram associated with the cluster, with the detected mode m and normal range $[l, h]$. A probability density distribution (PDF) is first obtained from this histogram. Please note that histograms in this paper are obtained using the plug-in rule technique with the bin size of the histogram equal to the bandwidth calculated from the plug-in rule [16].



(a) (b)

Fig. 2. (a) The histogram of a detected data cluster from Step 1. The vertical axis shows the number of observations. (b) The corresponding MF defined for the cluster.

Figure 2(b) shows the corresponding triangular MF defined for the cluster with m as the center point B for the triangular MF.

Next, using the generated PDF, we calculate the probability density of lower bound (l) and higher bound (h) of the cluster, denoted by $P(l)$ and $P(h)$ in Fig. 2(b), respectively. Then, we find the parameter A for the triangular MF by extrapolating the two points (m, 1) and (l, $P(l)$). In the same manner, we find the parameter C by extrapolating the two points (h, $P(h)$) and (m, 1). Now, triangular MF is defined using Eq. (1).

$$\mu^i(x) = \begin{cases} 0 & if\ x \leq A_i \\ \frac{x - A_i}{B_i - C_i} & if\ A_i < x \leq B_i \\ \frac{C_i - x}{C_i - B_i} & if\ B_i < x \leq C_i \\ 0 & if\ C_i \leq x \end{cases} \tag{1}$$

More details associated with employing the approach for generating triangular MFs can be found in [13].

3.2 Generating Trapezoidal Membership Functions

A trapezoidal MF is characterized by four parameters A, B, C, D (with $A < B \leq C < D$) as shown in Fig. 3. These determine the x coordinates of the four corners of the

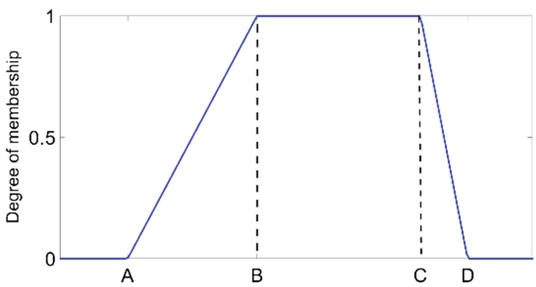


Fig. 3. An example of a trapezoidal membership function defined by a quad (A, B, C, D).

underlying trapezoidal defined over the attribute space. Specifically, points A and D specify the left and right feet. Parameters B and C specify the shoulders for the trapezoidal.

To obtain the parameters (A, B, C, D) associated with the trapezoidal MF for a cluster, *NR* $[l, h]$ associated with the cluster, and the first and third quartiles for the cluster are used. Figure 4(a) shows the histogram for an example cluster obtained from Step 1. In this example Q1 and Q3 denote the location of the first and third quartiles, respectively, and the normal range for the cluster is shown as $[l, h]$. The following operations are performed to define the respective trapezoidal MF, shown in Fig. 4(b), to represent the cluster.

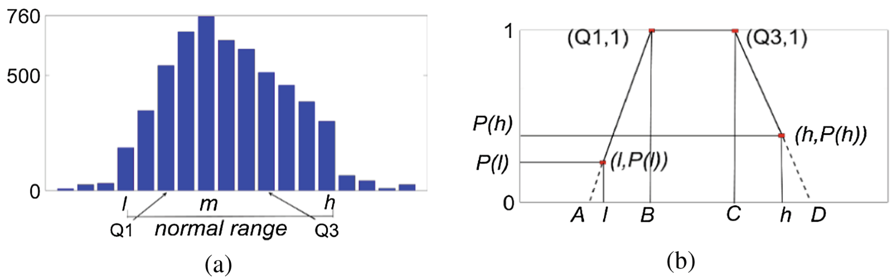


Fig. 4. (a) An example of a histogram of a data cluster obtained from Step 1. The vertical axis shows the number of data points in each bin. (b) The corresponding trapezoidal MF defined for the cluster.

1. Calculate the probability density of lower bound (l) and higher bound (h) of the cluster, denoted by $P(l)$ and $P(h)$, respectively (see Fig. 4(b)).
2. Find the parameter A for the trapezoidal MF by extrapolating the two points ($Q1$, 1) and (l, $P(l)$).
3. Set B and C as the location of Q1 and Q3, respectively.
4. Find the parameter D by extrapolating the two points (h, $P(h)$) and ($Q3$, 1).

Define the trapezoidal MF using Eq. (2).

$$\mu^i(x) = \begin{cases} 0 & \text{if } x \leq \mathrm{A} \\ \frac{x-\mathrm{A}}{\mathrm{B}-C} & \text{if } A < x < \mathrm{B} \\ 1 & \text{if } B \leq \mathrm{x} \leq \mathrm{C} \\ \frac{\mathrm{C}-x}{\mathrm{C}-\mathrm{B}} & \text{if } \mathrm{C} < x \leq \mathrm{D} \\ 0 & \text{if } \mathrm{C} \leq x \end{cases} \tag{2}$$

Here, data between Q1 and Q3 are assigned with full membership as they represent the middle 50% of cluster values. Accordingly, in the extrapolating performed in operations 2 and 4, ($Q1$, 1), and ($Q3$, 1) represent the coordinates for the location of shoulders.

3.3 Impact of the Shape of Cluster on Support of MFs

Depending on the shape of data distribution for a cluster obtained from Step 1, the support for the generated triangular MF can be greater than the support for the trapezoidal MFs. Figure 5(a) shows an example data distribution with different colours indicating the range for clusters obtained from Step 1 of the proposed approach. Figure 5(b) shows the characteristics of the triangular and trapezoidal MFs generated for the cluster shown in red color. As can be observed, since this cluster is located in the middle of the data distribution, the tail end of this cluster has been truncated. When the extrapolations are performed to obtain parameters of the triangular and trapezoidal MFs, it is clearly evident that the range for the triangular MF is greater than that for the trapezoidal MF. This means that generated triangular MFs generally offer more perturbation of normal data without labelling them as abnormal.

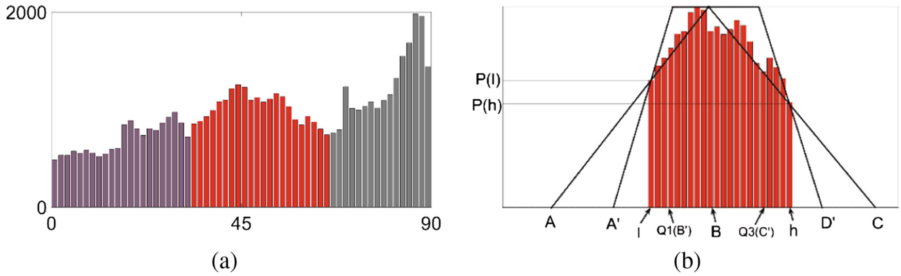


Fig. 5. An example data distribution - different colours indicate range for clusters obtained using Step 1. (b) The characteristics of the triangular and trapezoidal MFs generated for the cluster shown in red color. (Color figure online)

4 Experimental Results

Our evaluation consists of comparison between the proposed approach and two other techniques in terms of (i) parameterising MFs for attributes with different distributions, and (ii) classification performance of a fuzzy rule set that was developed using the parameterised output of each of the 3 techniques.

4.1 Dataset

We evaluated the effectiveness of the proposed approach using attributes associated with a dataset for classification activities of daily living (ADLs), as previously used in [7]. This dataset is collected via multiple Kinect cameras, each installed in a different area of a single monitored house. The system used for the gathering of data consisted of Windows 8.1 notebook PCs, with one notebook per Kinect device. Custom data collection code was written in C# under the Microsoft.Net framework. Data analysis was subsequently performed in MATLAB™.

Data was collected from this house for a period of five weeks, during which a single occupant undertook activities typical of a retired elderly person. From each Kinect, observations for activities undertaken are taken at one-second intervals and ones in which a person is detected are stored. The entire dataset consisted of more than two million observations. The attributes we extracted from this dataset were the occupant's Centre of Gravity pixel location (Xc, Yc), Aspect Ratio (AR) of the 3D axis-aligned bounding box, and Orientation (O).

The dataset for each location was partitioned into a training set and an unseen test set. The training set for each location consisted of nearly one million observations of behaviour patterns associated to typical (or normal) ADLs of the occupant. The test set holds some sequences of normal behaviour (i.e. typical ADLs) and abnormal events (e.g. occupant lying on the floor of the kitchen).

4.2 Comparison of Techniques for Parameterizing Attributes with Different Characteristics

Attributes with different data distribution were used to compare the parameterisation results between the proposed approach (VBMS–RS) and two other techniques: (i) using MS (instead of VBMS) in Step 1 of the proposed approach followed by the procedure of robust statistics in Step 2 (MS-RS), and (ii) using the Fuzzy-C-Means (FCM) clustering algorithm [17] to generate a fixed number of membership functions over the domain of a particular attribute without the use of robust statistics. For each particular attribute, we empirically set this number for FCM according to the number of modes in the attribute probability density function, as discussed in the following sections. In each case, comparisons are made through the clusters and MFs produced by each of the 3 techniques.

Attribute with Separated Distributions. One example with well separated distributions is for the attribute *Xc* associated with the living room dataset, as shown in Fig. 6. The reason is that, as shown in Fig. 7, the living room was occupied mainly for sitting at a computer desk (the left distribution) and using the sofa for watching TV (the distribution to the right) and as a result, values for *Xc* are mostly concentrated around two separate regions in feature space of *Xc* (i.e., 150 and 325), respectively.

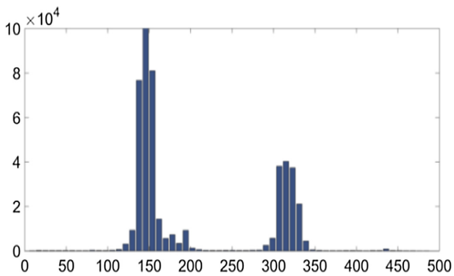


Fig. 6. A bimodal distribution for the *Xc* attribute associated with the living room dataset. Note that the base distributions are well separated.

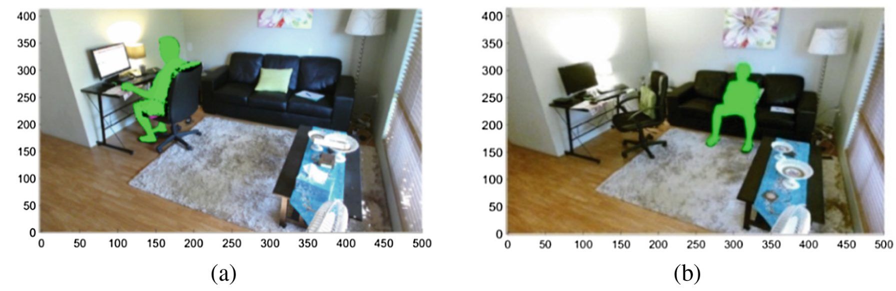


Fig. 7. (a) Sitting at a computer desk, and (b) watching TV while sitting on a sofa in the living room. The body of the occupant is masked by its binary silhouette obtained from the Kinect SDK and the numbers in the vertical and horizontal axis indicate pixel location.

Figure 8(a) illustrates the results of using VBMS–RS for parameterising distributions of *Xc* from the living room. Each underlying distribution of data associated with a detected mode is shown with a different colour.

VBMS–RS could separate correctly this attribute feature space into two main underlying distributions. The distribution to the right in Fig. 8(a) is in the shape of

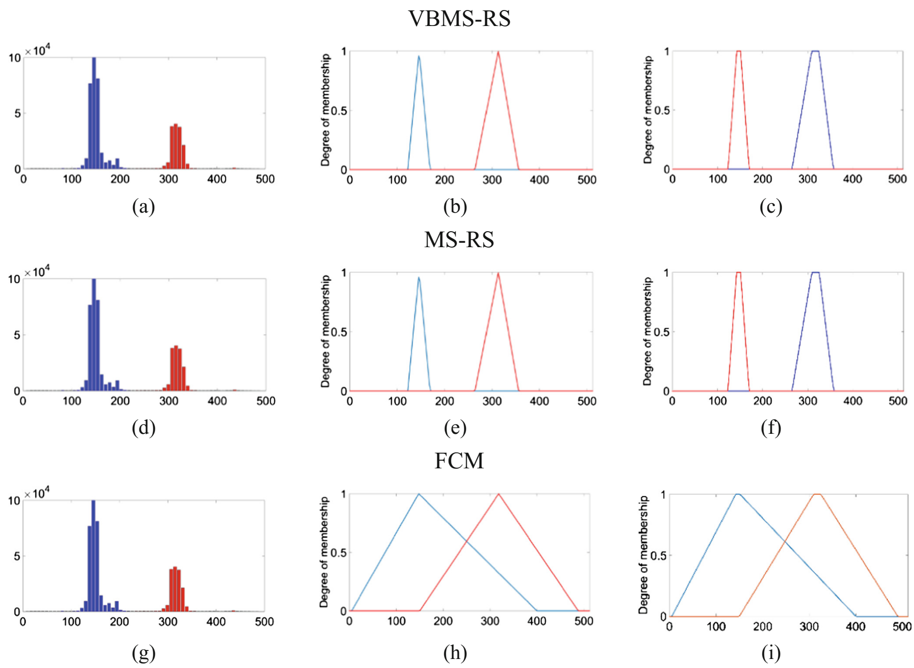


Fig. 8. Results for different techniques for parameterising the two base distributions shown in Fig. 6. The different colours in each of (a), (D), and (G) show range for clusters obtained using different techniques. (B), (E), and (H) show the respective triangular Mfs, resulted from the output of the 3 techniques, respectively. (C), (F) and (I) show the corresponding trapezoidal Mfs resulted from the output of the 3 techniques, respectively. (Color figure online)

reverse-J (skewed to the left), and the corresponding triangular and trapezoidal MFs defined by VBMS–RS (parts (b) and (c)) represents only the range for the normal data points associated with this distribution. Note that since the normal range obtained for both clusters in Fig. 8(a) is small, the shoulder of trapezoidal MFs in Fig. 8(c) is small and thus both trapezoidal MFs nearly have the same shape and cover the same area in the feature space.

To further evaluate VBMS-RS, VBMS was replaced with MS in Step 1 of the proposed approach and the experiment was repeated. By comparing the results, it was observed that where the distributions in the attribute feature space are separated distinctly, both methods work equally well. However, MS–RS requires an empirical input, the bandwidth parameter, whereas for VBMS in the proposed approach, the initial bandwidth is derived from the data automatically [13].

In the comparison using the FCM technique, the number of membership functions was empirically set to 2 (as this is obvious from a visual examination of the data). As shown with blue and red colours, Fig. 8(g) demonstrates that this technique correctly separated the attribute into two distributions in the attribute feature space. As a result, MFs in Fig. 8(h) and (i) were generated to represent the two distributions detected in the attribute feature space via triangular and trapezoidal shapes, respectively. Since this technique does not use robust statistics, the resulting parameterization of the MFs is not the same as the proposed approach. More specifically, MFs generated by this technique have a wider support and hence represent a wider area outside the normal range for the two main distributions in Fig. 6. As a result, the MFs generated by this technique will also encompass many rare observations (outliers) around the main distributions. For example, triangular MFs generated by FCM give membership degrees 0.17 and 0.83 to the outlier point $Xc = 380$ so that the sum of memberships of this point becomes one. This is in contrast to triangular MFs generated by VBMS-RS which give zero membership to this outlier point.

Attribute with a Unimodal Distribution. One example of the attributes that have unimodal skewed distribution is the *AR* attribute from the dining room, as illustrated in images on the left hand side of Fig. 9 (i.e., (a), (d), and (g)). The overall distribution shown in those images illustrates the skewed distribution for *AR*. Different colors in each of the images indicate the distributions related to the clusters that have been obtained using different techniques. Figure 9(b), (e), and (h) illustrate triangular MFs generated using the 3 techniques. Figure 9(c), (f), and (i) show results of generating trapezoidal MFs for the distribution of the *AR* attribute using the 3 techniques. As shown in Fig. 9(a), VBMS-RS correctly associated all data points with the only mode in the distribution. However, as shown in Fig. 9(d), MS-RS has broken the distribution into two clusters. This difference is mainly because in VBMS, points that correspond to the tails of the underlying density will get a broader neighbourhood and a smaller importance. Therefore, they will be included to main structures and hence, tail of distributions will not be broken into pieces. This is unlike MS, where it assigns a fixed global bandwidth to all data points and hence all points receive the same importance when estimating the PDF of data.

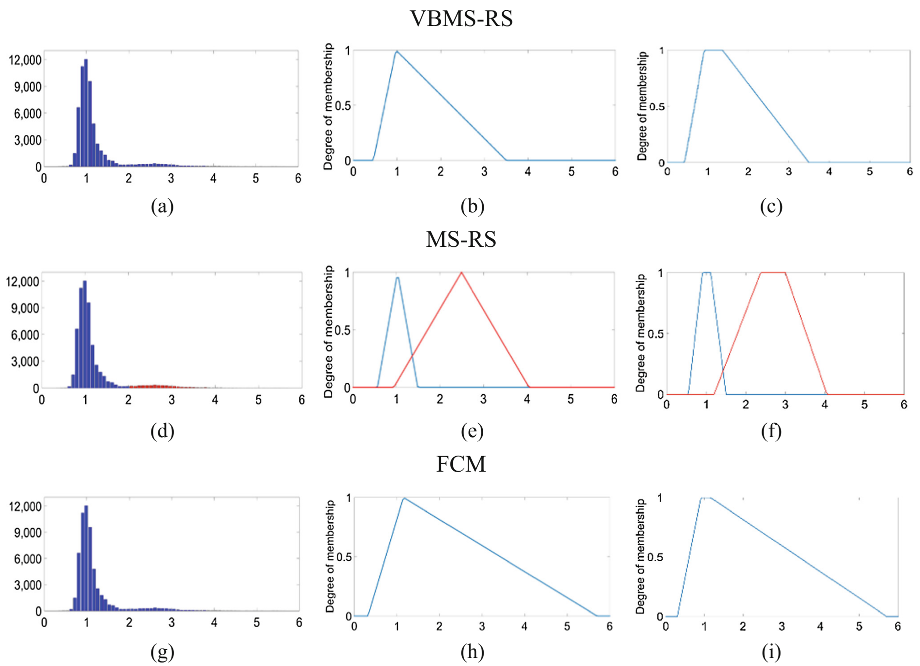


Fig. 9. Using different techniques for parameterising distribution of *AR* attribute for the dining room dataset. (a), (d), and (g) show the range for clusters obtained using the 3 different techniques. (b), (e), and (h) show the respective triangular MFs resulted from the output of 3 techniques. (c), (f) and (i) show the corresponding trapezoidal MFs resulted from the output of 3 techniques. (Color figure online)

As the distribution is unimodal, input value for the number of clusters in FCM was set to 1. From Fig. 9(g) we can see that although this technique has grouped all data points in the distribution into the stipulated one cluster, the supports of the generated MFs in Fig. 9(h) and (i) are much broader than corresponding MFs generated by VBMS-RS which might lead to non-specific responses for classification of the attribute values (i.e., every point is considered to be in the set). Also, when the application of generating MFs is for classification of outliers, the generated MFs by this technique represents many rare observations (outliers) located between $AR = 4$ and $AR = 6$, and thus will be not able to correctly classify a new abnormal observation within that range. However, both triangular and trapezoidal MFs from the proposed approach does not represent any data point for outside the normal range [0.5, 3.5] and therefore, VBMS-RS method can obtain better classification results for normal points and better accuracy for handling outlier observations.

Attribute with Multimodal Distribution. An example of an attribute with multimodal distribution is *Xc* from the kitchen dataset. From the ground truth in examining the video data for this attribute there were three distinct places for *Xc* where the occupant performed most of the activities in the kitchen. As a result, PDF for this attribute has 3 modes, each associated with a particular distribution and the 3 distributions overlap.

Results of parameterising this attribute using the different techniques are shown in Fig. 10. Input value for the number of clusters to be created by FCM was set to 3. It is clear from the results in Fig. 10 that, VBMS-RS partitions the feature space into the right number of membership functions whereas using other techniques were unable to separate the mixed distributions correctly. The difference between results for VBMS-RS and MS-RS is due to the fact that, using VBMS, the data points lying in large density regions will get a narrower neighbourhood since the kernel bandwidth is smaller, but are given a larger importance. So when base distributions are mixed in the attribute feature space, VBMS can better separate those structures than MS. This finding is consistent with Comaniciu et al. [12].

As observed in Fig. 10, the range of triangular MFs generated by 3 techniques, in general, is greater than their respective trapezoidal MFs. For instance, the wider data distribution associated with the right-hand-side cluster (ranging from pixel location 220 to 500) in Fig. 10(a) and (d) causes the respective trapezoidal MFs in Fig. 10(c) and

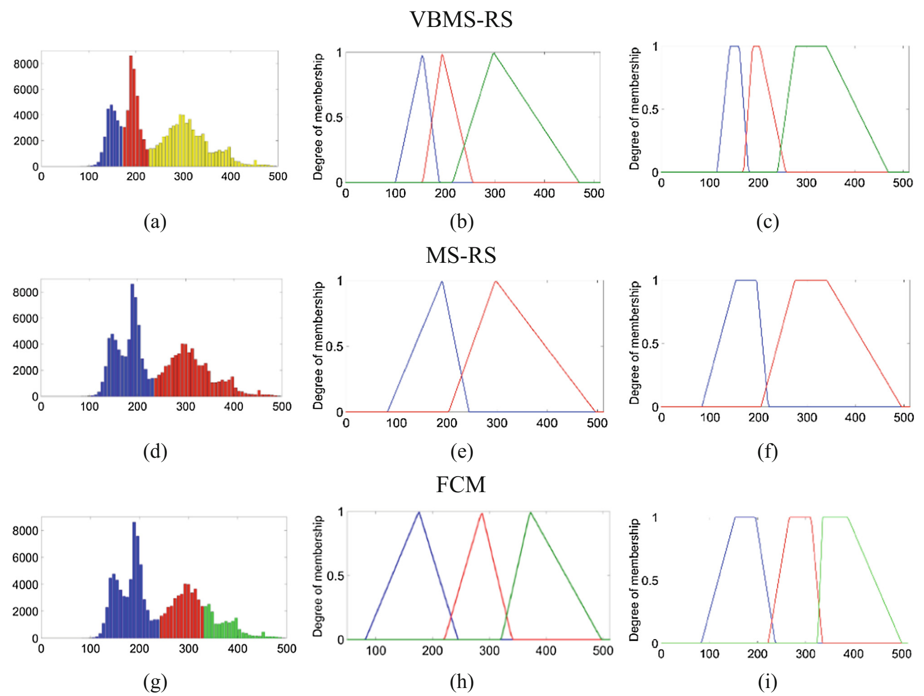


Fig. 10. Results for using the 3 different techniques for parameterising distribution of *Xc* associated with the kitchen dataset. (a), (d), and (g) show the range for clusters obtained using the 3 different techniques. (b), (e), and (h) show the respective triangular MFs resulted from the output of 3 techniques. (c), (f) and (i) show the corresponding trapezoidal MFs resulted from the output of 3 techniques. (Color figure online)

(f) to have a relatively wider shoulder, which in turn has resulted in those MFs to receive steeper descending foots upon performing extrapolation in Step 2 of the approach. Accordingly, they cover less area when compared to their respective triangulate MFs in Fig. 10(b) and (e). As already pointed out in Sect. 3.3, at the stage of classification, the wider support of triangular MFs allows more variations for normal data for each cluster.

From Fig. 10(g) FCM has partitioned the attribute feature space to be represented by three MFs. However, the parameters for these three MFs are different to those of the results from VBMS-RS. The reason is that this technique aims to minimise the distance of data points from their respective cluster centres. As a result, the locations of centre of clusters are not always corresponding to the modes in distribution of data. Furthermore, as seen in Fig. 10(h) and (i), distributions with their modes located on pixel locations 150 and 200, respectively, are represented by the same MF. Hence, MFs generated by this technique are not accurately representing data distributions in this attribute feature space.

4.3 Results on Classification Accuracy Using MFs Produced by Different Techniques

The characteristics of MFs generated by a particular technique have a direct impact on performance of the corresponding fuzzy rule set for classification purposes. In other words, a better technique to estimate the base distributions for attributes can lead to more representative MFs and hence a better classification accuracy of the corresponding fuzzy rule set. To investigate this, we conducted experiments in which we applied the output of the 3 different MF generation techniques, including the proposed approach, to obtain a fuzzy rule set for the application of detecting abnormal activities in ADLs. As we had data from 5 rooms and each room was associated with 4 attributes with different number of modes in their corresponding PDF, we empirically set the number of clusters for FCM to a specific number (i.e., 3) to suite across all situations, a technique used typically by existing fuzzy approaches [18]. To obtain the classifier, we extracted the attributes (described in Sect. 4.1) from the training dataset associated with each location and developed the fuzzy system using the approach from [7]. A brief description of this approach is described below:

The unsupervised ADLs monitoring approach proposed by [7] uses a set of attributes derived from Kinect camera observations and consists of two phases: training and classification (monitoring).

During the training phase, the system learns "normal" behaviour patterns of the occupant as a set of fuzzy rules. More specifically, for each monitored location, epochs of activity are first determined and for each epoch, normal behaviour patterns are then learnt by finding frequent occurrences of attributes via the use of a fuzzy association rule mining algorithm [8]. Hence, for each monitored location a fuzzy rule set is obtained. The antecedent part of each rule in a fuzzy rule set represents a combination of fuzzy linguistic values describing a frequent behaviour of the occupant along with their expected epoch of activity. The normal duration of the frequent behaviour is specified in the consequent part of the rule. Also, for each monitor location, the duration of infrequent behaviours is estimated.

The monitoring phase takes the fuzzy rule set obtained from the training phase as input, and for each location, it classifies the current behaviour of the occupant as abnormal if it is not in the set of frequent behaviours. For more detail, we refer the reader to [7].

Table 1 shows the total number of rules obtained from the output of each technique. Variations in performing activities typically create a number of base distributions for an attribute. Hence, each particular activity can be typically represented by a specific combination of base distributions over different attributes. Since MFs generated by VBMS-RS for an attribute represent the normal range for base distributions, different versions of a particular activity are usually represented by the same combination of fuzzy attributes, and thus one fuzzy rule. When each frequent activity is modelled by one fuzzy rule, the total number of fuzzy rules for a location becomes considerably less than situations where multiple rules are developed to represent versions of the same activity. For example, from the generated rules by VBMS-RS, it was observed that, for each of the four activity epochs detected for the living room, two fuzzy rules were generated to represent frequent activities of sitting behind the computer desk, and sitting on the sofa, respectively, forming eight (out of nine) rules for the living room rule set. Also, for the afternoon epoch, a rule represents the activity of sleeping on the sofa as it was occasionally carried out by the occupant during that period.

Table 1. The number of fuzzy rules obtained from the output of different MF generation techniques.

Dataset Technique		Kitchen	Living room	Dining room	Bedroom	Overall
FCM	Triangular	30	28	22	5	85
	Trapezoidal	33	29	23	5	90
MS-RS	Triangular	24	8	6	2	40
	Trapezoidal	26	8	6	2	42
VBMS-RS	Triangular	15	9	6	2	32
	Trapezoidal	16	9	6	2	33

From Table 1 it can be seen that using the output of other MF generation techniques (e.g., FCM and MS-RS) resulted in a higher number of rules. This is because MFs obtained from the output of those techniques do not necessarily represent data distributions over space of attributes well, and therefore, the values of an attribute for different versions of an activity might be represented by different MFs defined over the attribute. This results in different fuzzy rules with a different combination of fuzzy attributes to be generated for modelling slightly different versions of the same activity, hence a higher number of rules.

It is also observed that when using FCM and MS-RS techniques, the number of rules obtained by trapezoidal MFs is slightly more than those obtained from using triangular MFs. As already mentioned, MFs generated by these techniques do not necessarily represent base distributions well and, as the support of trapezoidal MFs are less in comparison with triangular MFs, more combinations of MFs are required to represent variations of attributes during activities. However, since triangular or trapezoidal MFs generated by VBMS-RS represent base distributions well, variation of attributes during activities have been represented by almost the equal number of rules.

Table 2 quantitatively compares classification accuracy for fuzzy rules obtained using the output of the 3 different MF generation techniques with different MFs. More specifically, 40 sequences of different scenarios for normal and abnormal behaviour in the unseen test set (20 sequences for each category of normal and abnormal behaviour, respectively) were used to evaluate the accuracy of the fuzzy rule set obtained using the output of a particular technique with a particular type of MF (i.e., triangular and trapezoidal) and the resulting classification accuracy is reported in Table 2.

Table 2. Results of using the output of different MF generation techniques to obtain a fuzzy rule set for the application of detecting abnormal activities in ADLs.

Method		Normal behaviour	Abnormal behaviour	Overall accuracy
FCM (3 clusters)	Triangular	70.0%	85.0%	77.5%
	Trapezoidal	60.0%	85.0%	72.5%
MS-RS	Triangular	90.0%	80.0%	85.0%
	Trapezoidal	85.0%	80.0%	82.5%
VBMS	Triangular	100.0%	35.0%	67.5%
	Trapezoidal	100.0%	40.0%	70.0%
VBMS-RS	Triangular	100.0%	85.0%	92.5%
	Trapezoidal	95.0%	85.0%	90.0%

From Table 2 it can be observed that when we use MS-RS to obtain triangular MFs for fuzzy rules, 6 of the test sequences, mostly representing an abnormal behaviour, were classified incorrectly. This is mainly because MS couldn't distinctly separate overlapped distributions in feature space of attributes. Therefore, for some attributes two or more behaviour patterns belonging to different overlapped distributions were represented by the same MF and hence represented by the same fuzzy rule. For example, distributions of *AR* for crouching on the kitchen floor (to pick up an object) and bending down (to manipulate objects inside the kitchen cabinet), while belonging to different main distributions in the attribute feature space, considered as belonging to the same cluster, and hence, the corresponding fuzzy rule set was not able to label a sequence for spending a long time sitting on the kitchen floor as abnormal behaviour. Using MS-RS to obtain trapezoidal MFs results in misclassification of 7 test sequences, resulting in a classification accuracy of 82.5%. Specifically, as the support of trapezoidal MFs is less in comparison with their respective triangular ones, one another test sequence for normal behavior that was slightly different from the training samples was misclassified as being abnormal.

Classification that results from using FCM to generate triangular MFs produced accuracy of 77.5%. This is mostly because the test sequences involving normal behaviour patterns that were slightly different from their corresponding training patterns were misclassified by this classifier as abnormal. This was mainly because FCM broke main distributions for some attributes into pieces, and consequently, for a particular activity, when most of training values belonged to a particular part of the distribution and the values for test sequences fell into another part of the distribution, the corresponding fuzzy rule for the activity could not be able to trigger and hence less accuracy of the classifier. Using FCM to generate trapezoidal MFs produced an accuracy of 72.5%.

We also evaluated the classification accuracy of the fuzzy rules obtained by applying the proposed approach without robust statistics and results are shown Table 2 denoted by VBMS. When using triangular MFs, we observed that many test sequences for abnormal behaviour have been labelled as normal. In those sequences, the values of attributes were well outside of the normal range for the main distributions in the feature space of attributes. However, since the range of generated triangular MFs was wider than the range of main distributions, they included many outlier observations, and hence, outlier observations in each of those test sequences triggered a corresponding rule for a normal behaviour in the rule base to be fired and resulted in the test sequence being labelled normal. The generated trapezoidal MFs cover less area (less outliers) in attributes space and hence caused slightly less test sequences for abnormal behaviour to be labelled as normal.

From the last two rows of Table 2 we see that the rule set obtained from the results of VBMS-RS with triangular MFs could classify 37 test sequences correctly and hence an accuracy of 92.5%. We observed that for almost all attributes, using the combination of VBMS and robust statistics yields in the resulting triangular MFs representing only the normal range for the base distributions in the attributes. Therefore, while outlier observations for abnormal behaviours were classified correctly, attribute values during most of sequences for normal behaviour were within the bounds associated with the generated MFs, and hence, those sequences triggered a rule corresponding to a normal behaviour to fire. However, it is observed that using trapezoidal MFs caused one test sequence for normal behaviour to be misclassified as abnormal, resulting in an overall accuracy of 90% for the classifier.

Note that although, in average, using trapezoidal MFs resulted in slightly less classification accuracy, we observed that they assign higher degrees of membership to normal data points during the classification stage. Specifically, trapezoidal MFs assign the maximum membership degree to those data points that fall between the two shoulders in the trapezoidal shape whereas triangular MFs give the maximum membership value just to those data points correspond to the centroid of the triangular shape.

5 Conclusion

In this paper, we presented an unsupervised approach that incorporates variable bandwidth mean-shift and robust statistics for generating fuzzy membership functions. The approach automatically learns the number of representative functions from the underlying data distribution and then works out the associated parameters of a given membership function. We examined the proposed approach using the trapezoidal and the triangular membership functions and compared its performance against two other techniques. Results in Sect. 4.2 demonstrated that, from perspective of partitioning an attribute, the generated membership functions generated by VBMS-RS can better separate the underlying distributions. As a better technique to estimate the base distributions for attributes can lead to more representative MFs and hence a better classification accuracy of the corresponding fuzzy rule set, we examined classifiers constructed using the proposed method of generating membership function and 3 other methods for both the trapezoidal and the triangular membership functions. Results were examined from the perspectives of number of fuzzy rules and classification accuracy associated with each classifier. Results in Tables 1 and 2 showed that classifiers associated with VBMS-RS outperformed three other classifiers that used different approaches for parameterisation of the attributes.

References

1. Medasani, S., Kim, J., Krishnapuram, R.: An overview of membership function generation techniques for pattern recognition. Int. J. Approx. Reason. **19**(3), 391–417 (1998)
2. Seki, H.: Fuzzy inference based non-daily behavior pattern detection for elderly people monitoring system. In: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC 2009, pp. 6187–6192. IEEE (2009)
3. Moeinzadeh, H., et al.: Improving classification accuracy using evolutionary fuzzy transformation. In: Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers. ACM (2009)
4. Amaral, T.G., Crisóstomo, M.M.: Automatic helicopter motion control using fuzzy logic. In: 2001 The 10th IEEE International Conference on Fuzzy Systems. IEEE (2001)
5. Tang, K., Man, K., Chan, C.: Fuzzy control of water pressure using genetic algorithm. In: Proceedings of the Safety, Reliability and Applications of Emerging Intelligent Control Technologies (2014)
6. Takagi, H., Hayashi, I.: NN-driven fuzzy reasoning. Int. J. Approx. Reason. **5**(3), 191–212 (1991)
7. Pazhoumand-Dar, H., Lam, C.P., Masek, M.: A novel fuzzy based home occupant monitoring system using kinect cameras. In: IEEE 27th International Conference on Tools with Artificial Intelligence, Vietri sul Mare, Italy (2015)
8. Kuok, C.M., Fu, A., Wong, M.H.: Mining fuzzy association rules in databases. ACM Sigmod Rec. **27**(1), 41–46 (1998)
9. Doctor, F., Iqbal, R., Naguib, R.N.: A fuzzy ambient intelligent agents approach for monitoring disease progression of dementia patients. J. Ambient Intell. Humaniz. Comput. **5**(1), 147–158 (2014)
10. Castellano, G., Fanelli, A., Mencar, C.: Generation of interpretable fuzzy granules by a double-clustering technique. Arch. Control Sci. **12**(4), 397–410 (2002)
11. Hubert, M., Vandervieren, E.: An adjusted boxplot for skewed distributions. Comput. Stat. Data Anal. **52**(12), 5186–5201 (2008)
12. Comaniciu, D., Ramesh, V., Meer, P.: The variable bandwidth mean shift and data-driven scale selection. In: Proceedings of the Eighth IEEE International Conference on Computer Vision, ICCV 2001. IEEE (2001)

13. Pazhoumand-Dar, H., Lam, C.P., Masek, M.: Automatic generation of fuzzy membership functions using adaptive mean-shift and robust statistics. In: Proceedings of the 8th International Conference on Agents and Artificial Intelligence, Italy (2016)
14. Rousseeuw, P.J., Hubert, M.: Robust statistics for outlier detection. Wiley Interdisc. Rev. Data Mining Knowl. Discov. **1**(1), 73–79 (2011)
15. Brys, G., Hubert, M., Struyf, A.: A robust measure of skewness. J. Comput. Graphical Stat. **13**(4), 996–1017 (2004)
16. Sheather, S.J., Jones, M.C.: A reliable data-based bandwidth selection method for kernel density estimation. J. R. Stat. Soc. Ser. B (Methodol.) **53**, 683–690 (1991)
17. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: The fuzzy c-means clustering algorithm. Comput. Geosci. **10**(2), 191–203 (1984)
18. Tajbakhsh, A., Rahmati, M., Mirzaei, A.: Intrusion detection using fuzzy association rules. Appl. Soft Comput. **9**(2), 462–469 (2009)