Fuzzy Feature Weighting Techniques for Vector Quantisation

Dat Tran, *Senior Member, IEEE*, Wanli Ma, Dharmendra Sharma and Phuoc Nguyen

***Abstract*— Vector quantization (VQ) is a simple but effective modelling technique in pattern recognition. VQ employs a clustering technique to convert a feature vector set in to a cluster center set to model the feature vector set. Some clustering techniques have been applied to improve VQ. However VQ is not always effective because data features are treated equally although their importance may not be the same. Some automated feature weighting techniques have been proposed to overcome this drawback. This paper reviews those weighting techniques and proposes a general scheme for selecting any pair of clustering and feature weighting techniques to form a fuzzy feature weighting-based VQ modelling technique. Besides the current techniques, a number of new feature weighting-based VQ techniques is proposed and their evaluations are also presented.**

I. Introduction

Vector quantization (VQ) is a popular modelling technique that has been used in pattern recognition, image processing, and data reduction methods [1]. VQ is used to convert a feature vector set into a small set of distinct vectors using a clustering technique. Advantages of this reduction are reduced storage and computation. The distinct vectors are called codevectors and the set of codevectors that best represents the training set is called the codebook. Since there is only a finite number of code vectors, the process of choosing the best representation of a given feature vector is equivalent to quantizing the vector and leads to a certain level of quantization error. This error decreases as the size of the codebook increases, however the storage required for a large codebook is non-trivial. The VQ codebook can be used as a model in pattern recognition. The key point of VQ modelling is to derive an optimal codebook which is commonly achieved by using a clustering technique.

Some clustering techniques have been applied to improve VQ [2]. However VQ is not always effective because all features are treated equally, but they may not have the same importance. Some automated feature weighting techniques have been proposed to overcome this drawback. This paper reviews these weighting techniques and proposes a general scheme for selecting any pair of clustering and feature weighting techniques to form a feature weighting-based VQ technique. Besides the current feature weighting-based VQ techniques, a number of new techniques is proposed and their experimental results are also presented.

II. Vector Quantization

A. Vector Quantization Modeling

VQ modeling can be summarized as follows. Given a training set of T feature vectors $X = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$, where each source vector $\mathbf{x}_t = (x_{t1}, x_{t2}, \ldots, x_{tM})$ is of M dimensions. Let $\lambda = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K\}$ represent the codebook of size K, where $\mathbf{c}_k = (c_{k1}, c_{k2}, \ldots, c_{kM})$, $k = 1, 2, \ldots, K$ are code vectors. Each code vector $\mathbf{c}_k$ is assigned to an encoding region R_k in the partition $\Omega = \{R_1, R_2, \ldots, R_K\}$. Then the source vector $\mathbf{x}_t$ can be represented by the encoding region R_k and expressed by

$$V(\mathbf{x}_t) = \mathbf{c}_k, \text{ if } \mathbf{x}_t \in R_k \tag{1}$$

B. K-Means Partition

Let $U = [u_{kt}]$ be a matrix whose elements are memberships of $\mathbf{x}_t$ in the nth cluster, $k = 1, \ldots, K$, $t = 1, \ldots, T$. A K-partition space for X is the set of matrices U such that [1]

$$u_{kt} \in \{0, 1\}\, \forall k, t, \sum_{k=1}^{K} u_{kt} = 1\, \forall t, \quad 0 < \sum_{t=1}^{T} u_{kt} < T\, \forall k \tag{2}$$

where $u_{kt} = u_k(\mathbf{x}_t)$ is 1 or 0, according to whether $\mathbf{x}_t$ is or is not in the kth cluster, $\sum_{k=1}^{K} u_{kt} = 1\ \forall t$ means each $\mathbf{x}_t$ is in exactly one of the K clusters, and $0 < \sum_{t=1}^{T} u_{kt} < T$ $\forall k$ means that no cluster is empty and no cluster is all of X because of $1 < K < T$.

The K-means VQ (KMVQ) technique is based on minimization of the sum-of-squared-errors function as follows

$$J_{KM}(U, \lambda; X) = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt} d_{kt}^2 \tag{3}$$

where λ is a set of prototypes, in the simplest case, it is the set of cluster centers $\lambda = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K\}$, and d_{kt} is the Euclidean norm of $(\mathbf{x}_t - \mathbf{c}_k)$. Minizing $J_{KM}(U, W, \lambda; X)$ over the variables U and λ yields the following equations

$$\mathbf{c}_k = \sum_{t=1}^{T} u_{kt}\mathbf{x}_t \Big/ \sum_{t=1}^{T} u_{kt} \qquad 1 \le k \le K \tag{4}$$

$$u_{kt} = \begin{cases} 1 : d_{kt} < d_{jt} & j = 1, \ldots, K,\ j \neq k \\ 0 : \text{otherwise} \end{cases} \tag{5}$$

C. Fuzzy C-Means Partition

Let $U = [u_{kt}]$ be a matrix whose elements are fuzzy memberships of $\mathbf{x}_t$ in the kth cluster, $k = 1, \ldots, K$, $t = 1, \ldots, T$. A K-partition space for X is the set of matrices U such that [2]

$$u_{kt} \in [0, 1]\ \forall k, t, \sum_{k=1}^{K} u_{kt} = 1\ \forall t, \quad 0 < \sum_{t=1}^{T} u_{kt} < T\ \forall k \tag{6}$$

Dat Tran, Wanli Ma, Dharmendra Sharma, and Phuoc Nguyen is with Faculty of Information Sciences & Engineering, University of Canberra, ACT 2601, Australia (phone: +61 2 6201 2394; email: dat.tran@canberra.edu.au).

978-1-4244-8126-2/10/$26.00 ©2010 IEEE

where $u_{kt} \in [0,1]\ \forall k,t$ and $\sum_{k=1}^{K} u_{kt} = 1\ \forall t$ mean it is possible for each $\mathbf{x}_t$ to have an arbitrary distribution of fuzzy membership among the N fuzzy clusters, and $0 < \sum_{t=1}^{T} u_{kt} < T\ \forall k$ means that no cluster is empty and no cluster is all of X because of $1 < K < T$.

The fuzzy c-means VQ (FCMVQ) technique is based on minimization of the sum-of-squared-errors function as follows [2]

$$J_{FCM}(U,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt}^{\gamma} d_{kt}^{2} \tag{7}$$

where λ is a set of prototypes, $\gamma > 1$ is a weighting exponent on each fuzzy membership u_{it} and controls the degree of fuzziness, in the simplest case, it is the set of cluster centers $\lambda = \{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_K\}$, and d_{kt} is the Euclidean norm of $(\mathbf{x}_t - \mathbf{c}_k)$. The basic idea of the FCM method is to minimize $J_{FCM}(U,\lambda;X)$ over the variables U and λ on the assumption that matrix U, which is part of the optimal pairs for $J_{FCM}(U,\lambda;X)$, identifies the good partition of the data. The FCMVQ algorithm is summarized as follows.

1) Given a training data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$, where $\mathbf{x}_t = (x_{t1}, x_{t2}, \ldots, x_{tK})$, $t = 1, 2, \ldots, T$.
2) Initialize the membership values $u_{kt}, 1 \le k \le K, 1 \le t \le T$, at random
3) Given $\epsilon > 0$ (small real number).
4) Set $i = 0$ and $J_{FCM}^{(i)}(U,\lambda;X) = 0$. Iteration:
   a) Compute cluster centers

$$\mathbf{c}_k = \sum_{t=1}^{T} u_{kt}^{\gamma}\mathbf{x}_t \Big/ \sum_{t=1}^{T} u_{kt}^{\gamma} \tag{8}$$

   b) Compute d_{kt} and $J_{FCM}^{(i+1)}(U,\lambda;X)$

$$d_{kt} = \|\mathbf{c}_k - \mathbf{x}_t\|_2 \tag{9}$$

   c) Update membership values

$$u_{kt} = \frac{1}{\sum_{n=1}^{K}\left(d_{kt}^{2}/d_{nt}^{2}\right)^{1/(\gamma-1)}} \tag{10}$$

5) If

$$\frac{|J_{FCM}^{(i+1)}(U,\lambda;X) - J_{FCM}^{(i)}(U,\lambda;X)|}{J_{FCM}^{(i+1)}(U,\lambda;X)} > \epsilon \tag{11}$$

then set $J_{FCM}^{(i)}(U,\lambda;X) = J_{FCM}^{(i+1)}(U,\lambda;X)$, $i = i+1$ and go to step (a).

D. Fuzzy Entropy Partition

Define $U = [u_{kt}]$ and fuzzy c-partition space as in fuzzy c-means partition. The fuzzy entropy technique is based on minimisation of the following function [3]:

$$J_{FE}(U,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt} d_{kt}^{2} + \delta \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt}\log u_{kt} \tag{12}$$

where $U = \{u_{kt}\}$ is a fuzzy c-partition of X, $\delta > 0$ controls the degree of fuzzy entropy, λ and d_{kt} are defined as in (7). The basic idea of the FE technique is to minimize $J_{FE}(U,\lambda;X)$ over the variables U and λ.

The fuzzy entropy VQ (FEVQ) algorithm is summarized as follows.

1) Given a training data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\}$, where $\mathbf{x}_t = (x_{t1}, x_{t2}, \ldots, x_{tK})$, $t = 1, 2, \ldots, T$.
2) Initialize the membership values $u_{kt}, 1 \le t \le T, 1 \le k \le K$, at random
3) Given $\epsilon > 0$ (small real number).
4) Set $i = 0$ and $J_{FE}^{(i)}(U,\lambda;X) = 0$. Iteration:
   a) Compute cluster centers

$$\mathbf{c}_k = \sum_{t=1}^{T} u_{kt}\mathbf{x}_t \Big/ \sum_{t=1}^{T} u_{kt} \tag{13}$$

   b) Compute d_{kt} and $J_{FE}^{(i+1)}$

$$d_{kt} = \|\mathbf{c}_k - \mathbf{x}_t\|_2 \tag{14}$$

   c) Update membership values

$$u_{kt} = \frac{e^{-d_{kt}^{2}/\delta}}{\sum_{n=1}^{K} e^{-d_{nt}^{2}/\delta}} \tag{15}$$

5) If

$$\frac{|J_{FE}^{(i+1)}(U,\lambda;X) - J_{FE}^{(i)}(U,\lambda;X)|}{J_{FE}^{(i+1)}(U,\lambda;X)} > \epsilon \tag{16}$$

then set $J_{FE}^{(i)}(U,\lambda;X) = J_{FE}^{(i+1)}(U,\lambda;X)$, $i = i+1$ and go to step (a).

III. Fuzzy Feature Weighting

There are currently 2 automated feature weighting techniques found in the literature. The first technique employs a weight vector $W = \{w_m\}$, $m = 1, \ldots, M$ for all feature vectors in the M dimensional feature space. Each feature m is assigned a weight w_m, $m = 1, \ldots, M$ [4]. The second technique employs a weight matrix $W = \{w_{mk}\}$, $m = 1, \ldots, M, k = 1, \ldots, K$, where K is the number of clusters. This means that every cluster k has its own weight vector $w_{mk}, m = 1, \ldots, M$ [5], [6], [7]. Weight values were estimated using either FCM-based estimation technique [4], [5] or FE-based estimation technique [6]. These two feature weighting techniques were used in FCM and K-Means clustering.

It is noticed that feature weighting techniques are applied to feature levels, that is they are independent of clustering techniques (or partition techniques in VQ). Therefore a weighting estimation technique can be applied to any clustering techniques provided that these clustering techniques employ an objective function-based optimisation method. For example, an FCM feature weighting has been applied to FCM clustering as seen in [5] and to K-means clustering in [4] and [6].

From this notice, we present all combinations of the two fuzzy feature weighting techniques and the three clustering techniques presented in the previous section. Moreover there are also two kinds of weights which are weight vector and weight matrix as mentioned above, so there are totally 2 * 3 * 2 = 12 combinations found. Four of these 12 combinations have been used in [4], [5] and [6]. The other eight combinations are proposed in this paper.

A. FCM Feature Weighting for KMVQ Using Weight Vector

Let $W = [w_1, w_2, \ldots, w_M]$ be the weight vector for M dimensions and α be a parameter weight for w_m. The objective function is defined as follows

$$J_{FCM-KM-v}(U, \lambda; X) = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt} \sum_{m=1}^{M} w_m^{\alpha} d_{ktm}^2 \quad (17)$$

where $\alpha > 1$, d_{ktm} is the m-th component distance of the distance d_{kt} between $\mathbf{c}_k$ and $\mathbf{x}_t$

$$d_{ktm}^2 = (c_{km} - x_{tm})^2 \quad (18)$$

and weight values satisfy the following conditions:

$$0 \leq w_m \leq 1 \quad \forall m, \qquad \sum_{m=1}^{M} w_m = 1 \quad (19)$$

The well-known Langrange multiplier method is used to minimise the objective function. We have

$$w_m = \frac{1}{\sum_{n=1}^{M} \left(D_m^2 / D_n^2\right)^{1/(\alpha-1)}} \quad (20)$$

where

$$D_m^2 = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt} d_{ktm}^2 \quad (21)$$

Cluster centers and membership functions are calculated in (4) and (5), respectively for K-means partition above.

B. FCM Feature Weighting for KMVQ Using Weight Matrix

Let $W = \{w_{km}\}$, $k = 1, \ldots, K, m = 1, \ldots, M$, where K is the number of clusters. The objective function is defined as follows

$$J_{FCM-KM-m}(U, W, \lambda; X) = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt} \sum_{m=1}^{M} w_{km}^{\alpha} d_{ktm}^2 \quad (22)$$

where α and d_{ktm} are previously defined. Weight matrix is calculated as follows

$$w_{km} = \frac{1}{\sum_{n=1}^{M} \left(D_{km}^2 / D_{kn}^2\right)^{1/(\alpha-1)}} \quad (23)$$

where

$$D_{km}^2 = \sum_{t=1}^{T} u_{kt} d_{ktm}^2 \quad (24)$$

Cluster centers and membership functions are calculated in (4) and (5), respectively for K-means partition above.

C. FCM Feature Weighting for FCMVQ Using Weight Vector

The objective function is defined as

$$J_{FCM-FCM-v}(U, W, \lambda; X) = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt}^{\gamma} \sum_{m=1}^{M} w_m^{\alpha} d_{ktm}^2 \quad (25)$$

Use (20), (8) and (10) to calculate weight values, cluster centers and membership functions, respectively.

D. FCM Feature Weighting for FCMVQ Using Weight Matrix

The objective function is defined as

$$J_{FCM-FCM-m}(U, W, \lambda; X) = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt}^{\gamma} \sum_{m=1}^{M} w_{km}^{\alpha} d_{ktm}^2 \quad (26)$$

Use (23), (8) and (10) to calculate weight values, cluster centers and membership functions, respectively.

E. FCM Feature Weighting for FEVQ Using Weight Vector

The objective function is defined as

$$J_{FCM-FE-v}(U, W, \lambda; X) = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt} \sum_{m=1}^{M} w_m^{\alpha} d_{ktm}^2 + \delta \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt} \log u_{kt} \quad (27)$$

Use (20), (13) and (15) to calculate weight values, cluster centers and membership functions, respectively.

F. FCM Feature Weighting for FEVQ Using Weight Matrix

The objective function is defined as

$$J_{FCM-FE-m}(U, W, \lambda; X) = \sum_{k=1}^{K} \sum_{t=1}^{T} u_{kt}^{\gamma} \sum_{m=1}^{M} w_{km}^{\alpha} d_{ktm}^2 \quad (28)$$

Use (23), (13) and (15) to calculate weight values, cluster centers and membership functions, respectively.

G. FE Feature Weighting for KMVQ Using Weight Vector

$$J_{FE-KM-v}(U,W,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt} \sum_{m=1}^{M} w_m d_{ktm}^2 + \beta \sum_{m=1}^{M} w_m \log w_m \quad (29)$$

The weight vector is calculated as follows

$$w_m = \frac{e^{-D_m^2/\beta}}{\sum_{n=1}^{M} e^{-D_n^2/\beta}} \quad (30)$$

where D_m is calculated using (21)

H. FE Feature Weighting for KMVQ Using Weight Matrix

$$J_{FE-KM-m}(U,W,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt} \sum_{m=1}^{M} w_{km} d_{ktm}^2 + \beta \sum_{k=1}^{K}\sum_{m=1}^{M} w_{km} \log w_{km} \quad (31)$$

The weight matrix is calculated as follows

$$w_{km} = \frac{e^{-D_{km}^2/\beta}}{\sum_{n=1}^{M} e^{-D_{kn}^2/\beta}} \quad (32)$$

where D_{km} is calculated using (24)

I. FE Feature Weighting for FCMVQ Using Weight Vector

$$J_{FE-FCM-v}(U,W,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt}^{\gamma} \sum_{m=1}^{M} w_m d_{ktm}^2 + \beta \sum_{m=1}^{M} w_m \log w_m \quad (33)$$

Use (30), (8) and (10) to calculate weight values, cluster centers and membership functions, respectively.

J. FE Feature Weighting for FCMVQ Using Weight Matrix

$$J_{FE-FCM-m}(U,W,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt}^{\gamma} \sum_{m=1}^{M} w_{km} d_{ktm}^2 + \beta \sum_{k=1}^{K}\sum_{m=1}^{M} w_{km} \log w_{km} \quad (34)$$

Use (32), (8) and (10) to calculate weight values, cluster centers and membership functions, respectively.

K. FE Feature Weighting for FEVQ Using Weight Vector

$$J_{FE-FE-v}(U,W,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt} \sum_{m=1}^{M} w_m d_{ktm}^2 + \delta \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt} \log u_{kt} + \beta \sum_{m=1}^{M} w_m \log w_m \quad (35)$$

Use (30), (13) and (15) to calculate weight values, cluster centers and membership functions, respectively.

L. FE Feature Weighting for FEVQ Using Weight Matrix

$$J_{FE-FE-m}(U,W,\lambda;X) = \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt} \sum_{m=1}^{M} w_{km} d_{ktm}^2 + \delta \sum_{k=1}^{K}\sum_{t=1}^{T} u_{kt} \log u_{kt} + \beta \sum_{k=1}^{K}\sum_{m=1}^{M} w_{km} \log w_{km} \quad (36)$$

Use (32), (13) and (15) to calculate weight values, cluster centers and membership functions, respectively.

IV. Evaluation

We present speaker characteristic classification results as an evaluation of the proposed feature weighting-based VQ models.

Classifying speaker characteristics is an important task in Dialog Systems, Speech Synthesis, Forensics, Language Learning, Assessment Systems, and Speaker Recognition Systems. In Human-Computer Interaction applications, the interaction between users and computers taking place at the speech-driven user interface. For example, Spoken Dialogs Systems provide services in domains of finance, travel, scheduling, tutoring, or weather. The systems need to gather automatically information from the user in order to provide timely and relevant services. Most telephone-based services today use spoken dialog systems to either route calls to the appropriate agent or even handle the complete service by an automatic system. In Human-Centered applications, the computers stay in the background attempting to anticipate and serve peoples needs. One example is Smart Room Environments in which computers watch and interpret peoples actions and interactions in order to support communication goals. We particularly focus on Australian accent classification in this paper. Although the accent is only spoken by a minority of the population, it has a great deal of cultural credibility. It is disproportionately used in advertisements and by newsreaders. According to linguists, three main varieties of spoken English in Australia are Broad (spoken by 34% of the population), General (55%) and Cultivated (11%). They are part of a continuum, reflecting variations in accent. Although some men use the pronunciation, the majority of Australians that speak with the accent are women. Broad Australian English is usually spoken by men, probably because this accent is associated with Australian masculinity. It is used to identify Australian characters in non-Australian media programs and is familiar to English speakers. The majority of Australians speak with the General Australian accent. Cultivated Australian English has some similarities to British Received Pronunciation, and is often mistaken for it. In the past, the cultivated accent had the kind of cultural credibility that the broad accent has today. For example, until 30 years ago newsreaders on the government funded ABC had to speak with the cultivated accent.

We developed a classification system that can classify persons based on their gender, age and accent simultaneously.

Voice features are extracted as feature vectors and are used to train speaker group models with different feature weighting-based VQ techniques which are FE-KM-v, FE-KM-m, FE-FCM-v, FE-FCM-m, FE-FE-v, and FE-FE-m Fusion of classification results from those groups is then performed to obtain results for each gender, age and accent.

A. ANDOSL Database

The Australian National Database of Spoken Language (ANDOSL) corpus [12] comprises carefully balanced material for Australian speakers, both Australian-born and overseas-born migrants. The aim was to represent as many significant speaker groups within the Australian population as possible. Current holdings are divided into those from native speakers of Australian English (born and fully educated in Australia) and those from non-native speakers of Australian English (first generation migrants having a non-English native language). A subset used for speaker verification experiments in this paper consists of 108 native speakers. There are 36 speakers of General Australian English, 36 speakers of Broad Australian English and 36 speakers of Cultivated Australian English in this subset. Each of the three groups comprises 6 speakers of each gender in each of three age ranges (18-30, 31-45 and 46+). So there are total of 18 groups of 6 speakers labeled as ijk, where i denotes f (female) or m (male), j denotes y (young) or m (medium) or e (elder), and k denotes g (general) or b (broad) or c (cultivated). For example, the group fyg contains 6 female young general Australian English speakers. Each speaker contributed in a single session, 200 phonetically rich sentences. All waveforms were sampled at 20 kHz and 16 bits per sample.

B. Speech Processing

In speaker characteristics feature research, prosodic approaches attempt to capture speaker-specific variation in intonation, timing, and loudness. Because such features are supra-segmental (are not properties of single speech segments but extend over syllables and longer regions), they can provide complementary information to systems based on frame-level or phonetic features. One of the most studied features is speech fundamental frequency (or as perceived, pitch), which reflects vocal fold vibration rate and is affected by various physical properties of the speakers vocal folds, including their size, mass, and stiffness. Distributions of frame-level pitch values have been used in a number of studies. Although they convey useful information about a speakers distribution of pitch values, such statistics do not capture dynamic information about pitch contours and are thus not viewed as high-level here. Speech processing was performed using open source openSMILE feature extraction [10]. There are 16 low-level descriptors chosen including ZCR, RMS energy, pitch frequency, HNR, and MFCC 1-12 in full accordance to HTK-based computation [11]. To each of these, the delta coefficients are additionally computed. Next the 12 functionals including mean, standard deviation, kurtosis, skewness, minimum and maximum value, relative position, and range as well as two linear regression coefficients with their mean square error (MSE) are applied on a chunk basis. Thus, the total feature vector per chunk contains 16 * 2 * 12 = 384 features.

C. Experimental Results

TABLE I
GENDER CLASSIFICATION RESULTS (IN %) WITH DIFFERENT FUZZY FEATURE WEIGHTING-BASED VQ. NUMBER OF TEST UTTERANCES = 21600. NUMBER OF CLUSTERS = 8

Technique	Gender	
	Female	Male
KM	99.90	99.83
FE-KM-v	99.96	99.87
FE-KM-m	99.96	99.9 6
FCM	99.94	99.91
FE-FCM-v	99.95	99.95
FE-FCM-m	99.98	99.97
FE	99.91	99.94
FE-FE-v	99.97	99.98
FE-FE-m	99.98	100.0

TABLE II
AGE CLASSIFICATION RESULTS (IN %) WITH DIFFERENT FUZZY FEATURE WEIGHTING-BASED VQ. NUMBER OF TEST UTTERANCES = 21600. NUMBER OF CLUSTERS = 8

Technique	Age		
	Young	Middle	Elderly
KM	97.26	97.17	97.03
FE-KM-v	97.73	97.59	97.17
FE-KM-m	97.73	98.42	98.33
FCM	97.50	98.10	98.15
FE-FCM-v	97.73	98.19	97.69
FE-FCM-m	97.73	98.10	98.15
FE	97.27	97.18	97.04
FE-FE-v	97.69	97.73	97.18
FE-FE-m	97.55	98.38	98.33

TABLE III
ACCENT CLASSIFICATION RESULTS (IN %) WITH DIFFERENT FUZZY FEATURE WEIGHTING-BASED VQ. NUMBER OF TEST UTTERANCES = 21600. NUMBER OF CLUSTERS = 8

Technique	Accent		
	Broad	General	Cultivated
KM	97.63	97.31	96.34
FE-KM-v	98.19	98.10	96.97
FE-KM-m	98.28	98.05	97.17
FCM	98.15	98.15	96.76
FE-FCM-v	98.33	98.43	96.67
FE-FCM-m	98.39	98.78	97.27
FE	97.64	97.27	96.34
FE-FE-v	98.19	98.15	96.26
FE-FE-m	98.24	98.16	97.18

Experimental results for FCM feature weighting-based VQ can be found in the literature [4], [5], [6], [7]. In this paper, results for FE feature weighting-based VQ are presented. Results for others have similar improvements for fuzzy feature weighting VQ comparing with non-feature weighting ones.

Tables I, II and III present classification results for each gender, age and accent classification using 32 cluster centers. Table I shows very good result for gender classification for all techniques. Table II shows an improvement for fuzzy feature weighting-based VQ (FE-KM-v, FE-KM-m, FE-FCM-v, FE-FCM-m, FE-FE-v and FE-FE-m) comparing with non feature weighting VQ (KM, FCM and FE). Similar result is found for accent classification in Table III. The lowest classification rate for Cultivated is found comparing with the other two accents Broad and General.

V. Conclusions

We have reviewed current feature weighting techniques and have proposed a general scheme for selecting a pair of clustering and feature weighting techniques to apply to VQ. Besides the current feature weighting-based VQ techniques, a number of new techniques has been proposed and experimental results for fuzzy entropy technique have also been presented. The Australian speech database consisting of 108 speakers and 200 utterances for each speaker was used for evaluation. An improvement for fuzzy feature weighting-based VQ has been found via gender, age and accent classification using 21600 utterances from 108 speakers.

References

[1] R.O. Duda and P.E. Hart, *Pattern classification and scene analysis*, John Wiley & Sons, New York, 1973.

[2] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New York and London, 1981.

[3] D. Tran and W. Wagner, "Fuzzy entropy clustering", in Proceedings of FUZZ-IEEE Conference, 2000, vol. 1, pp. 152-157.

[4] J.Z. Huang, M. K. Ng, H. Rong, and Z. Li, "Automated Variable Weighting in k-means Type Clustering", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2005, vol. 27, no. 5, pp. 657-668.

[5] H. Frigui and O. Nasraoui. Unsupervised learning of prototypes and attribute weights. Pattern Recognition, 37(3):943952, 2004.

[6] L. Jing, M. K. Ng, and J. Zhexue Huang, "An Entropy Weighting k-Means Algorithm for SubspaceClustering of High-Dimensional Sparse Data", IEEETransactions on knowledge and data engineering, 19, pp. 1026-1041, 2007.

[7] L. Jing, M.K. Ng, J. Xu, and J.Z. Huang, "Subspace Clustering of Text Documents with Feature Weighting k-Means Algorithm," Proc. Ninth Pacific-Asia Conf. Knowledge Discovery and Data Mining, pp. 802-812, 2005.

[8] D. Tran and T. Pham, "Modeling Methods for Cell Phase Classification", Book chapter in the book Advanced Computational Methods for Biocomputing and Bioimaging, Editors: T.D. Pham, H. Yan, D. I. Crane, Nova Science Publishers, New York, USA, ISBN: 1-60021-278-6, 2007, chapter 7, pp. 143-166.

[9] D. Tran, W. Ma, D. Sharma and T. Nguyen, "Fuzzy Vector Quantization for Network Intrusion Detection", IEEE International Conference on Granular Computing, pp. 566-570, Silicon Valley, 2-4 November 2007, USA

[10] F. Eyben, M. Wollmer, B. Schuller (2009): Speech and Music Interpretation by Large-Space Extraction, http://sourceforge.net/projects/openSMILE

[11] P.C. Woodland, M.J.F. Gales, D. Pye & S.J. Young (1997).Broadcast news transcription using HTK. Proc. ICASSP97, pp. 719722, Munich

[12] J. B. Millar, J. P. Vonwiller, J. M. Harrington, andP. J. Dermody, "The Australian National Database of Spoken Language", in Proc. Int. Conf. Acoust., Speech, Signal Processing (ICASSP94), 1, pp. 97-100, 1994.

[13] T. Schultz, "Speaker characteristics," in Speaker Classification I. Springer Berlin / Heidelberg, 2007, pp. 47-74.

[14] N. Minematsu, M. Sekiguchi, and K. Hirose. Automatic estimation of ones age with his/her speech based upon acoustic modeling techniques of speakers. In Proc. IEEE Intl Conference on Acoustic Signal and Speech Processing, pp. 137140, 2002.

[15] I. Shafran, M. Riley, and M. Mohri. 2003. Voice signatures. In Proc. IEEE Automatic Speech Recognition and Understanding Workshop.

[16] F. Metze, J. Ajmera, R. Englert, U. Bub, F. Burkhardt,J. Stegmann, C. Mller, R. Huber, B. Andrassy, J.G. Bauer,and B. Littel, "Comparison of Four Approaches to Age and Gender Recognition for Telephone Applications," in ICASSP2007 Proceedings, IEEE International Conference on Acoustics, Speech and Signal Processing, Honolulu, Hawaii, USA,2007, vol. 4, pp. 1089 1092.

[17] E. Shriberg, "Higher-Level Features in Speaker Recognition," in Speaker Classification I. Springer Berlin / Heidelberg, 2007, pp. 241-259.

[18] S. Schtz, "Acoustic analysis of adult speaker age," in Speaker Classification I. Springer Berlin / Heidelberg, 2007, pp. 88-107.