

УДК 004.934.8'1

Я.Н. Имамвердиев, Л.В. Сухостат

МЕТОД ОБЪЕДИНЕНИЯ РЕШЕНИЙ КЛАССИФИКАТОРОВ ДЛЯ ЗАДАЧИ РАСПОЗНАВАНИЯ ДИКТОРА

Предлагается использование нечетких интегралов для объединения решений классификаторов систем распознавания диктора. В качестве набора признаков рассматриваются мгновенная частота и мгновенная амплитуда. Предлагаемый метод показывает значительно лучшие результаты по сравнению с применением единственного классификатора. Проводится сравнение предлагаемого метода с другими методами объединения решений классификаторов.

Введение

Верификация диктора – задача подтверждения личности заявленного пользователя на основе речевого образца. Некоторые из важных применений распознавания диктора включают верификацию клиентов для проведения банковских операций, доступ к банковским счетам через телефон, контроль использования кредитных карт, а также контроль безопасности в армии, военно-воздушных силах и на флоте.

Современные системы распознавания диктора по точности распознавания отстают от других биометрических систем, в том числе систем распознавания отпечатков пальцев. Между тем голос содержит богатую информацию о личности человека, и одним из путей повышения точности систем распознавания дикторов является объединение информации из разных источников в данных системах.

В биометрических системах существуют различные уровни объединения информации [1]:

- образца;
- признаков;
- значений соответствия;
- решений.

Согласно [2] для объединения значений соответствия наиболее часто применяется метод на основе преобразования данных. В этом методе для распознавания используются несколько классификаторов. Значения соответствия сначала нормализуются, и далее к ним применяются различные преобразования, например правило максимума, взвешенной суммы, логарифмического объединения и др.

Распознавание диктора включает два основных этапа: извлечение признаков и классификацию (рис. 1).

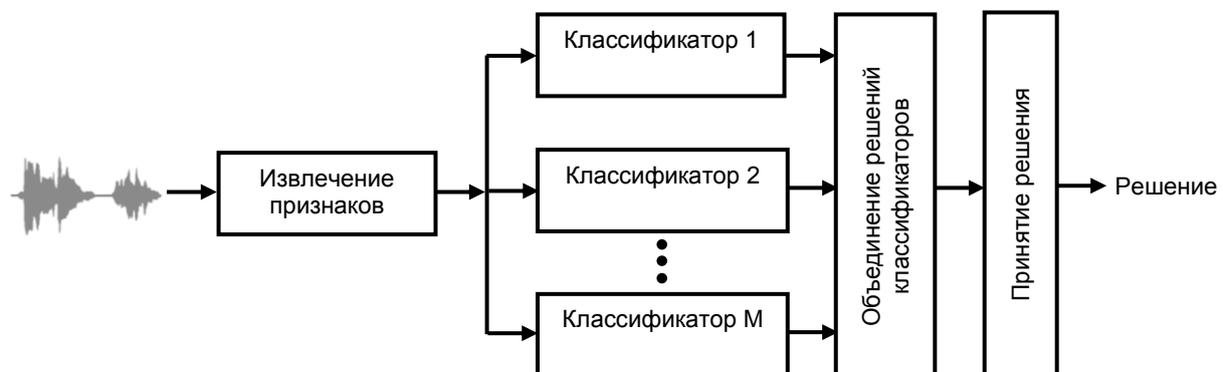


Рис. 1. Общая схема системы распознавания диктора

Цель выделения признаков заключается в преобразовании речевого сигнала к некоторому типу параметрического представления для дальнейшего анализа и обработки. Классификатор

использует эти характеристики, чтобы принять решение о принадлежности входных данных к одному из возможных классов, которые ассоциированы с отдельными дикторами из контрольной группы.

Классификатор состоит из $M + 1$ моделей дикторов (M моделей дикторов из контрольной группы и одной дополнительной модели «все остальные»), которые необходимы для принятия решения и строятся на этапе обучения. На этапе тестирования тестовый вектор признаков будет связан с каждой моделью диктора с указанием степени соответствия модели.

Большинство работ по биометрическим системам сосредоточено на методах объединения информации на уровне значений соответствия из-за скорости и эффективности.

В контексте распознавания диктора решения классификаторов, полученные на основе значений соответствия от различных моделей для каждого диктора, объединяются. Эти модели могут быть обучены с помощью различных речевых данных, признаков или методов моделирования. В конечном счете желательно, чтобы ошибки одной модели были исправлены другими и наоборот. Если все модели имеют согласованные ошибки, т. е. все они делают одну и ту же ошибку, то ни одна комбинация ее не исправит. Тем не менее, до тех пор пока существует некоторая степень несоответствия между ошибками, производительность может быть улучшена при правильном сочетании информационных данных.

Одной из отличительных особенностей нечеткого интеграла является то, что он способен представлять определенные взаимодействия между критериями. Теория нечетких интегралов была применена к распознаванию образов [3–5], обработке изображений [6–8] и объединению информации [9].

В данной работе предлагается метод объединения значений соответствия, полученных от классификаторов на основе нечетких интегралов для повышения точности распознавания диктора. В качестве речевых признаков рассматриваются мгновенные частота и амплитуда.

1. Извлечение речевых признаков для задачи распознавания диктора

Многие исследования были посвящены разработке различных схем извлечения характерных для диктора акустических признаков из речевых высказываний. Вульф в [10] среди наиболее существенных параметров выделяет частоту основного тона, спектральные признаки гласных, оценку голосового источника, продолжительность слова и др. Киннунен и соавторы [11] сделали обзор и обобщили основные особенности речи, которые были использованы для системы распознавания диктора. Наряду с классическими и ведущими признаками были приведены некоторые недавно полученные наборы параметров.

Идеальные признаки для систем распознавания диктора [10, 12] должны иметь большую вариабельность между дикторами и небольшую изменчивость внутри дикторов, быть достаточно устойчивыми к фоновому шуму и искажениям, часто использоваться в обычной речи, легко измеряться, быть стабильными во времени, не зависеть от здоровья (настроения) говорящего и быть трудно имитируемыми.

С точки зрения физической интерпретации [13] признаки, как правило, делятся на пять групп: спектральные, спектрально-временные, признаки голосового источника, просодические и признаки высокого уровня.

Среди спектральных признаков можно выделить кепстральные коэффициенты линейного предсказания (Linear Prediction Cepstral Coefficients, LPCC) [14], кепстральные коэффициенты по шкале мел (Mel-Frequency Cepstral Coefficients, MFCC) и др. [15].

MFCC-признаки являются наиболее известными и популярными спектральными признаками. Помимо кепстральных коэффициентов также рассматриваются их первые и вторые производные (дельта-признаки). В отличие от признаков высокого уровня, требующих более сложной предварительной обработки [16], их легче вычислить и получить хорошие результаты [17]. Помимо кепстральных особенностей, речь имеет и источник возбуждения, который содержит полезные свойства для распознавания диктора. Кроме того, в реальных ситуациях существуют большие различия между этапами разработки и практического применения системы распознавания диктора. Вследствие этого кепстральные признаки недостаточны, чтобы

обеспечить удовлетворительную и надежную точность распознавания. Они также не учитывают нестационарность и нелинейность человеческой речи.

В последние годы при описании и анализе свойств речи был использован подход на основе АМ-ФМ (Amplitude-Modulation Frequency-Modulation)-моделирования [18]. Монокомпонент АМ-ФМ-сигнала описывается уравнением

$$x(n) = A(n) \cos[\Theta_n], \quad (1)$$

где $A(n)$ обозначает мгновенную амплитуду монокомпонентного сигнала, а Θ_n – мгновенную фазу. При этом многокомпонентный сигнал сначала разлагается на монокомпоненты, каждый его компонент описывается мгновенной огибающей и мгновенной частотой. Данный подход показывает значительные улучшения показателей распознавания диктора.

2. Объединение решений классификаторов для задачи распознавания диктора

Сочетание различных источников информации было изучено в таких областях, как объединение данных [5], достижение консенсуса [19], теория принятия решений [20], комбинация решений нескольких экспертов [20] и т. д. Под комбинацией данных здесь понимается их объединение.

Выбор методов объединения данных может быть сделан в зависимости от типа информации, которая будет объединена. Например, если выходы модели есть вероятности, то могут быть применены такие методы, как линейное или логарифмическое объединение [19]. Если выходы модели на самом деле есть метки класса, то используются такие методы, как голосование [21] или ранжирование [20].

Правило максимума. Данный метод выбирает то значение соответствия, которое является наибольшим среди участвующих признаков. Математически он обозначается следующим образом:

$$P_{\max}(x) = \max(p_i(x)), \quad i = 1, \dots, n,$$

где $p_i(x)$ – вероятностный выход i -й модели; n – число моделей.

Линейное объединение. Метод линейного объединения является широко используемым методом объединения данных, удобным благодаря своей простоте. Он оценивается как взвешенная сумма выходов для каждой модели:

$$P_{\text{linear}}(x) = \sum_{i=1}^n \alpha_i p_i(x),$$

где $P_{\text{linear}}(x)$ – линейное объединение; α_i – веса; $p_i(x)$ – вероятностный выход i -й модели; n – число моделей. Параметры α_i , как правило, выбираются между 0 и 1, их сумма принимает равной единице.

Линейное объединение на выходе дает распределение вероятностей, и веса α_i предоставляют грубую меру для распределения i -й модели. Тем не менее следует отметить, что распределение вероятностей на выходе сумматора, а именно $P_{\text{linear}}(x)$, может быть мультимодальным. Есть вероятность, что это усложнит процесс принятия стратегии. Метод линейного объединения был оценен по нескольким сценариям в рамках распознавания диктора. Они включают в себя сочетание VQ кодовых книг, обученных на кепстральных и дельтакепстральных признаках [22], а также другие подходы [23–26].

Логарифмическое объединение. Данный метод является альтернативой линейному объединению, но в отличие от линейного при логарифмическом объединении выходное распределение унимодально [19]. Логарифмическое объединение состоит из взвешенного произведения выходов моделей:

$$P_{\text{linear}}(x) = \prod_{i=1}^n p_i^{\alpha_i}(x).$$

Следует отметить, что, согласно этой формуле, если для любой модели вероятность равна нулю, то и сочетание вероятностей также будет равно нулю. Метод логарифмического объединения был оценен для верификации диктора [23, 25] и обеспечивает такую же производительность, как и линейный метод объединения.

Нечеткие интегралы. Исходя из концепции нечетких множеств, предложенных Л. Заде [27], М. Сугено ввел понятия нечеткой меры и нечеткого интеграла [28]. Нечеткая мера представляет собой набор функций с монотонностью, но не всегда аддитивностью, и нечеткий интеграл является функционалом с монотонностью, которая используется для агрегирования информации из различных источников по отношению к нечеткой мере.

Пусть Y – произвольное множество и \tilde{B} – поле Бореля из Y . Функция множества g , определенная на \tilde{B} , является нечеткой мерой, если она удовлетворяет следующим трем аксиомам:

- граничные условия: $g(\emptyset) = 0$, $g(Y) = 1$;
- монотонность: $g(A) \leq g(B)$, если $A \subset B$ и $A, B \in \tilde{B}$;
- непрерывность: $\lim_{i \rightarrow \infty} g(A_i) \leq g(\lim_{i \rightarrow \infty} A_i)$, если $A_i \in \tilde{B}$ и $\{A_i\}$ монотонно (возрастающая последовательность измеримых множеств).

Нечеткая мера g_λ , также предложенная М. Сугено, удовлетворяет еще одному условию, известному как правило λ ($\lambda > -1$):

$$g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B),$$

где $A, B \subset Y$ и $A \cap B = \emptyset$.

Следует отметить, что, когда $\lambda = 0$, нечеткая мера g_λ становится вероятностной мерой. В общем, значение константы λ может быть определено из свойств нечеткой меры g_λ следующим образом.

Пусть $Y = \{y_1, y_2, \dots, y_m\}$. Если нечеткая плотность нечеткой меры g_λ определяется как функция $g: y_i \in Y \rightarrow [0, 1]$, такая, что $g_i = g_\lambda(\{y_i\})$, $i = 1, 2, \dots, m$, то нечеткая мера g_λ конечного множества может быть получена в виде [29]

$$g_\lambda(Y) = \sum_{i=1}^n g_i + \lambda \sum_{i_1=1}^{m-1} \sum_{i_2=i_1+1}^n g_{i_1} g_{i_2} + \dots + \lambda^{m-1} g_1 g_2 \dots g_m. \quad (2)$$

При условии, что $\lambda \neq 0$, (2) можно переписать в виде

$$g_\lambda(Y) = \frac{1}{\lambda} \left[\prod_{i=1}^m (1 + \lambda g_i) \right].$$

Учитывая граничное условие $g(Y) = 1$, постоянная λ может быть определена путем решения следующего уравнения:

$$\lambda + 1 = \prod_{i=1}^n (1 + \lambda g_i). \quad (3)$$

Пусть (Y, \tilde{B}, g) – пространство с нечеткой мерой и $f: Y \rightarrow [0, 1]$ есть \tilde{B} -мерная функция. Нечеткий интеграл из $A \subset Y$ от функции f по отношению к нечеткой мере g определяется как

$$\int_A f(y) \circ g(\cdot) = \sup_{\alpha \in [0, 1]} [\min(\alpha, g(f_\alpha))], \quad (4)$$

где $f_\alpha = \{y: f(y) \geq \alpha\}$.

Нечеткий интеграл в (4) называется интегралом Сугено. Когда $Y = \{y_1, y_2, \dots, y_n\}$ есть конечное множество и $0 \leq f(y_1) \leq f(y_2) \leq \dots \leq f(y_n) \leq 1$ (если нет, то элементы из Y переупорядочивают, чтобы сохранить эту связь), интеграл Сугено можно вычислить согласно

$$\int_A f(y) \circ g(\cdot) = \max_{i=1}^m [\min(f(y_i), g(A_i))],$$

где $A_i = \{y_i, y_{i+1}, \dots, y_m\}$ и $g(A_i)$ можно рекурсивно рассчитать с точки зрения нечеткой меры g_λ как

$$g(A_i) = g_i + g(A_{i-1}) + \lambda g_i g(A_{i-1}), \quad 1 \leq i \leq m. \quad (5)$$

Т. Муруфуши и М. Сугено в [29] предложили так называемый интеграл Шоке. Данный интеграл от f по отношению к нечеткой мере g определяется следующим образом:

$$\int_A f(y) dg(\cdot) = \sum_{i=1}^m [f(y_i) - f(y_{i-1})]g(A_i),$$

где $f(y_0) = 0$.

Интегралы Шоке и Сугено [29] являются идемпотентными, непрерывными, монотонно неубывающими операторами; это свойство подразумевает, что нечеткие интегралы всегда ограничены между минимумом и максимумом. Данные интегралы существенно различны по своей природе, так как первый основывается на линейных операторах, а второй – на нелинейных (минимума и максимума).

Интересным свойством нечеткого интеграла Шоке является то, что если g служит вероятностной мерой, интеграл Шоке эквивалентен классическому интегралу Лебега и вычисляет ожидание f относительно g по обычной вероятностной схеме.

Интеграл Шоке больше подходит для количественной агрегации (где числа имеют реальный смысл), в то время как интеграл Сугено – для порядковой (где только порядок имеет смысл).

3. Метод объединения значений соответствия на основе нечетких интегралов

Блок-схема предлагаемого метода включает выделение признаков для каждого диктора, классификацию и объединение полученных значений соответствия (рис. 2). Нечеткий интеграл используется для объединения результатов классификаторов с целью получения конечного решения. Значения, необходимые для объединения, получаются с учетом близости вектора признаков к соответствующим эталонным значениям.

В работе в качестве признаков рассматриваются мгновенная частота и мгновенная амплитуда [30]: x_1, x_2 . Для каждого речевого признака определяются нечеткие меры $g(A_1)$ и $g(A_2)$. На основе формулы (3) вычисляется λ . Далее из формул (5) находятся нечеткие меры для всевозможных комбинаций биометрических характеристик. Полученные нечеткие меры обозначаются через $g(A_1)$ и $g(A_2)$. С помощью функции принадлежности фазифицируются значения соответствия, полученные при сравнении рассматриваемых речевых признаков, что позволяет вычислить нечеткий интеграл.

С помощью метода нечетких интегралов значения соответствия каждого классификатора объединяются в одно значение, а затем принимается решение о пользователе или самозванце, учитывая пороговое значение.

В приложениях реального времени дикторозависимые пороги должны быть оценены априори. Порог определяется как линейная комбинация средних, дисперсий или стандартное отклонение от соотношения пользователь/самозванец [15]. Другие подходы основаны на кривых FAR (False Acceptance Rate) и FRR (False Rejection Rate) [31].

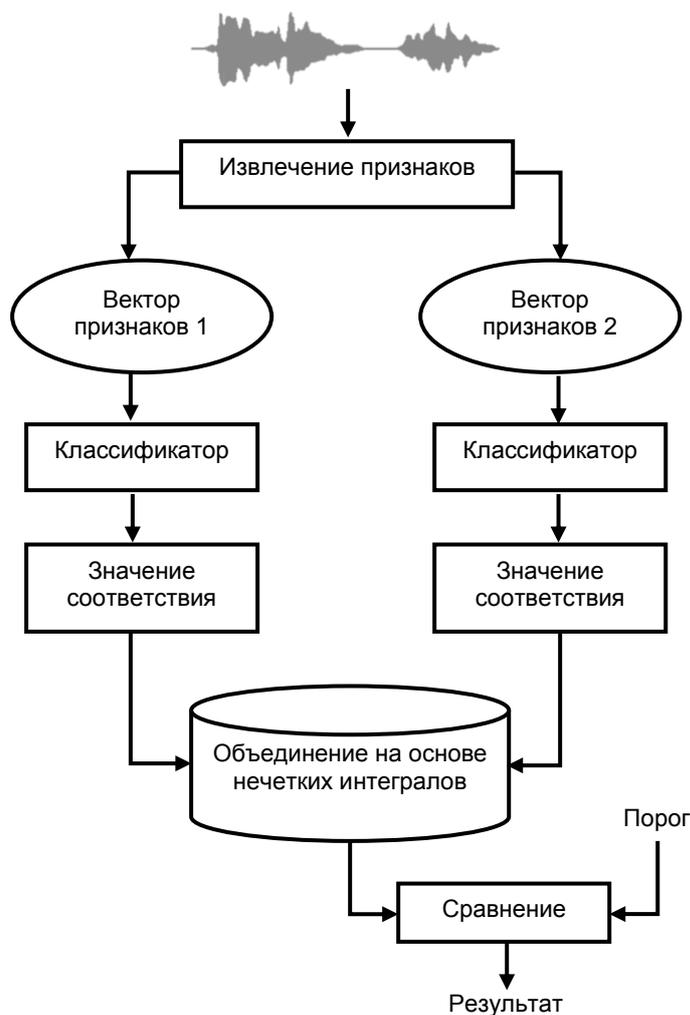


Рис. 2. Блок-схема предлагаемого метода объединения решений классификаторов

Метод на основе нечетких интегралов имеет высокую производительность и прост в применении. Основным преимуществом предлагаемого метода является то, что он может принять во внимание значимость каждого классификатора и взаимодействие между ними.

4. Результаты экспериментов

Для проведения экспериментов были взяты речевые образцы, полученные из пяти гласных английского языка (/a/, /e/, /i/, /o/ и /u/) от различных пользователей. Использовался набор из 50 образцов каждой из пяти гласных для 10 дикторов с целью идентификации на непересекающихся наборах гласных.

Гласные определяют значительную часть информации о речевом тракте человека и поэтому существенны при распознавании диктора. В качестве классификатора был рассмотрен метод SVM (Support Vector Machine) с радиальной базисной функцией ядра. С помощью процедуры кросс-валидации данные разделяются на k частей (блоков). В нашем случае $k = 10$.

Эксперименты проводились с применением четырех методов объединения: правила максимума, взвешенной суммы, логарифмического объединения и предлагаемого подхода на основе нечетких интегралов. Веса α_1 и α_2 принимаются равными 0,4 и 0,6 соответственно [32].

Для каждой гласной в табл. 1 показаны полученные значения соответствия (нормализованные в диапазоне [0, 1]) в результате классификации для мгновенных частот и амплитуд ($g(x_1)$ и $g(x_2)$).

Таблица 1
Значения соответствия для гласных /a/, /e/, /i/, /o/ и /u/

Гласные	Мгновенная амплитуда	Мгновенная частота
/e/	0,4420	0,5835
/i/	0,3703	0,5758
/o/	0,5735	0,6796
/a/	0,4677	0,6816
/u/	0,3844	0,5896

На этапе классификации каждый набор речевых данных разделен на обучающий и тестовый (в отношении 90 к 10 соответственно).

В табл. 2 показана эффективность предлагаемого подхода в сравнении с правилом максимума, методами взвешенной суммы и логарифмического объединения для значений мгновенной частоты и мгновенной амплитуды из табл. 1. При распознавании гласной /e/ значения, полученные в результате логарифмического объединения и на основе нечеткого интеграла, совпадают и уступают правилу максимума.

Таблица 2

Результаты объединения значений классификаторов

Гласные	Правило максимума	Метод взвешенной суммы	Метод логарифмического объединения	Предлагаемый метод
/e/	0,5835	0,5269	0,5815	0,5815
/i/	0,5758	0,4936	0,5389	0,5699
/o/	0,6796	0,6372	0,6860	0,7256
/a/	0,6816	0,5960	0,6330	0,7285
/u/	0,5896	0,5075	0,5523	0,5908

Экспериментальное тестирование показало, что предложенный подход дает лучшие результаты на исследуемом наборе речевых данных по сравнению с другими методами.

Заключение

Результаты экспериментов показывают, что метод на основе нечетких интегралов для объединения значений соответствия, полученных в результате классификации мгновенной частоты и мгновенной амплитуды, эффективен для задачи распознавания личности по голосу.

Значимость метода состоит не только в объединении результатов классификаторов, но и в рассмотрении каждого речевого признака в отдельности. Проведенный анализ показывает, что применение нечетких интегралов эффективно для объединения значений соответствия и существенно улучшает проверку идентичности диктора.

Дальнейшие исследования будут направлены на совершенствование предлагаемого метода, а также разработку новых методов для повышения точности распознавания диктора.

Работа выполнена при финансовой поддержке Фонда развития науки при Президенте Азербайджанской Республики – грант № EIF-RITN-MQM-2/IKT-2-2013-7(13)-29/18/1.

Список литературы

1. Ross, A.A. Handbook of Multibiometrics / A.A. Ross, K. Nandakumar, A.K. Jain. – London : Springer, 2006. – 198 p.
2. Solomonoff, A. Advances in channel compensation for SVM speaker recognition / A. Solomonoff, W. Campbell, I. Boardman // Proc. of ICASSP. – Philadelphia, PA, 2005. – P. 629–632.
3. Gader, P.D. Fusion of handwritten word classifiers / P.D. Gader, M.A. Mohamed, J.M. Keller // Pattern Recognition Letters. – 1996. – № 17. – P. 577–584.

4. Michel, G. The representation of importance and interaction of features by fuzzy measure / G. Michel // *Pattern Recognition Letters*. – 1996. – № 17. – P. 567–575.
5. Kuncheva, L.I. Decision templates for multiple classifier fusion: an experimental comparison / L.I. Kuncheva, J.C. Bezdek, R.P.W. Duin // *Pattern Recognition*. – 2001. – № 34. – P. 299–314.
6. Mirhosseini, A.R. Human face image recognition: an evidence aggregation approach / A.R. Mirhosseini, H. Yan // *Computer Vision and Image Understanding*. – 1998. – № 71. – P. 213–230.
7. Pham, T.D. Color image segmentation using fuzzy integral and mountain clustering / T.D. Pham, H. Yan // *Fuzzy sets and systems*. – 1999. – № 107. – P. 121–130.
8. Kwak, K.-C. Face recognition using fuzzy integral and wavelet decomposition method / K.-C. Kwak, W. Pedrycz // *IEEE Transactions on Systems, Man, and Cybernetics*. – 2004. – № 34. – P. 1666–1675.
9. Auephanwiriyaikul, S. Generalized Choquet fuzzy integral fusion / S. Auephanwiriyaikul, M.K. James, P.D. Gader // *Information Fusion*. – 2002. – № 3. – P. 69–85.
10. Wolf, J.J. Efficient acoustic parameters for speaker recognition / J.J. Wolf // *J. Acoustical Society of America*. – 1982. – Vol. 51, № 6. – P. 2044–2056.
11. Kinnunen, T. An overview of text-independent speaker recognition: from features to super-vectors / T. Kinnunen, H. Li // *Speech Communication*. – 2010. – Vol. 52, № 1. – P. 12–40.
12. Rose, P. Forensic speaker identification. Taylor & Francis forensic science series / P. Rose. – N.Y. : Taylor & Francis, 2002. – 380 p.
13. Kinnunen, T. Spectral features for automatic text-independent speaker recognition. Licentiate thesis / T. Kinnunen. – Finland : University of Joensuu, 2003.
14. Маркел, Дж. Линейное предсказание речи / Дж. Маркел, А.Х. Грей. – М. : Связь, 1980. – 308 с.
15. Furui, S. Cepstral analysis techniques for automatic speaker verification / S. Furui // *IEEE tran. acoust., speech, signal processing*. – 1981. – Vol. 27. – P. 254–272.
16. Reynolds, D. Channel robust speaker verification via feature mapping / D. Reynolds // *Proc. of ICASSP*. – Hong Kong, 2003. – Vol. 2. – P. 53–56.
17. Doddington, G. Speaker recognition based on idiolectal differences between speakers / G. Doddington // *Proc. of Eurospeech*. – Aalborg, Denmark, 2001. – Vol. 4. – P. 2521–2524.
18. Hemant, A.P. Forensic Speaker Recognition / A.P. Hemant, Amy Neustein. – Heidelberg : Springer, 2012. – 540 p.
19. Benediktsson, J.A. Consensus theoretic classification methods / J.A. Benediktsson, P.H. Swain // *IEEE Trans. Systems Man Cybernet*. – 1992. – № 22. – P. 688–704.
20. Ho, T.K. Decision combination in multiple classifier systems / T.K. Ho, J.J. Hull, S.N. Srihari // *IEEE Trans. Pattern Anal. Machine Intelligence*. – 1994. – № 16. – P. 66–75.
21. Xu, L. Methods of combining multiple classifiers and their applications to hand-written character recognition / L. Xu, A. Krzyzak, C.Y. Suen // *IEEE Trans. Systems Man Cybernet*. – 1992. – № 23. – P. 418–435.
22. Soong, F.K. On the use of instantaneous and transitional spectral information in speaker recognition / F.K. Soong, A.E. Rosenberg // *IEEE Trans. Acoust. Speech, Signal Process*. – 1988. – ASSP-36. – P. 871–879.
23. Farrell, K.R. Text-dependent speaker verification using data fusion / K.R. Farrell // *IEEE Intern. Conf. on Acoustic, Speech and Signal Processing*. – Detroit, Michigan, USA, 1995. – P. 349–352.
24. Sub-word speaker verification using data fusion methods / K.R. Farrell [et al.] // *IEEE Workshop on Neural Networks for Signal Processing*. – Amelia Island, Florida, 1997. – P. 531–540.
25. Farrell, K.R. An analysis of data fusion methods for speaker verification / K.R. Farrell, R.P. Ramachandran, R.J. Mammone // *IEEE Intern. Conf. on Acoustic, Speech and Signal Processing*. – Washington, USA, 1998. – P. 1129–1132.
26. Schalkwyk, J. Speaker verification with low storage requirements / J. Schalkwyk, N. Jain, E. Barnard // *IEEE Intern. Conf. on Acoustic, Speech and Signal Processing*. – Georgia, USA, 1996. – P. 693–696.
27. Zadeh, L.A. Fuzzy sets / L.A. Zadeh // *Information and Control*. – 1965. – № 8. – P. 338–353.
28. Gupta, M.M. Fuzzy measures and fuzzy integrals / M.M. Gupta, G.N. Saridis, B.R. Gaines. – N.Y. : Elsevier, 1977. – 510 p.

29. Murofushi, T. A theory of fuzzy measures. Representation, the Choquet integral and null sets / T. Murofushi, M. Sugeno // J. Math. Anal. Appl. – 1991. – Vol. 159, № 2. – P. 532–549.
30. Maragos, P. On amplitude and frequency demodulation using energy operators / P. Maragos, J.F. Kaiser, T.F. Quatieri // IEEE Trans. on Signal Processing. – 1993. – Vol. 41, № 4. – P. 1532–1550.
31. Zhang, W.D. A priori threshold determination for phrase-prompted speaker verification / W.D. Zhang [et al.] // Proc. Eurospeech'99. – Budapest, Hungary, 1999. – P. 1203–1206.
32. Hamid, L.A. Quality based Speaker Verification Systems using Fuzzy Inference Fusion Scheme / L.A. Hamid, D.A. Ramli // Proc. of the Intern. Conf. on Communications, Signal Processing and Computers. – Interlaken, Switzerland, 2014. – P. 96–103.

Поступила 27.10.2014

*Институт информационных технологий
Национальной академии наук Азербайджана,
Баку, ул. Б. Вахабзаде, 9
e-mail: yadigar@lan.ab.az,
lsuhostat@hotmail.com*

Y.N. Imamverdiyev, L.V. Sukhostat

MERGING CLASSIFIER DECISIONS FOR SPEAKER RECOGNITION

The paper proposes using fuzzy integrals for merging classifier decisions in speaker recognition systems. Instantaneous frequency and instantaneous amplitude are considered as the set of features. The approach shows significantly better results than a single classifier. A comparison of the proposed approach with the other methods for merging classifier decisions is provided.