

Автоматическая классификация типов мелодических контуров синтагм при проведении просодического анализа в рамках фоноскопического исследования



Ермоленко Т.В., кандидат технических наук

• просодические характеристики речи • мелодический контур • синтагма • интонационная конструкция • интонационный центр • частота основного тона • микропросодические явления • интонационные портреты акцентных единиц.

Важной частью экспертного криминалистического идентификационного исследования цифровых фонограмм является просодический анализ. Одна из его основных компонент — анализ интонационного типа синтагмы. Для классификации интонационных типов синтагм в работе используется система описания русской интонации Е.А. Брызгуновой, которая состоит из семи типов интонационных конструкций. Признаки классификации базируются на мелодическом, энергетическом и ритмическом контурах. Мелодический контур имеет сложную изрезанную структуру, вызванную резкими формантными переходами на границах звуков. Для повышения точности определения интонационного типа мелодический контур синтагмы подвергается обработке: удалению микропросодических явлений и сглаживанию. По полученной в результате обработки последовательности периодов основного тона формируются последовательности частот основного тона, энергий сигнала на периодах основного тона, временных расстояний от текущего периода до начала акцентного выделения. Эти последовательности нормируются, после чего используются для определения интонационного типа синтагмы с помощью DTW-алгоритма.

• prosodic speech features • melodic contour • syntagma, intonational structure • kernel of intonational structure • pitch, microprosody • description of intonational accentual units.

A prosodic analysis of speech signal is important part of expert criminalistics identification research of digital phonograms. An analysis of intonation type of syntagma is one of it's main components. The Bryzgunova's system of descriptions of Russian intonation consists

of seven intonational constructions types, which are used for intonation types of syntagma classification. A classification features are based on melodic, energetic and rhythmic contours. Melodic contour has a complex rugged structure, which is caused by the rapid formant transitions at the sounds boundaries. To improve the accuracy of intonation type classification a syntagma's melodic contour is processed as follows: mikroprosodic appearances are removed from it and then it is smoothed. After processing of pitch sequence a next sequences are formed: fundamental frequencies, signal energy at pitches, temporal distances from the current period prior to allocation of accentual. These sequences are normalized and then are used to determine the intonation type of syntagma using DTW-algorithm.

Постановка задачи

Просодика играет важную роль при восприятии речи человеком. Просодические характеристики речи зависят, с одной стороны, от интонации и стиля речи, с другой — от персональных характеристик говорящего, как постоянных, так и ситуативных. Поэтому в типовой структуре идентификационного лингвистического исследования эксперта-фоноскописта неотъемлемой частью является просодический анализ, основные компоненты которого: синтагматическое членение высказывания, анализ интонационного типа синтагмы и структурный анализ её мелодического контура.

Синтагма — ритмико-мелодическая единица слитной речи, грамматически оформленная и выражающая в пределах более сложного целого (предложения) относительно законченную мысль.

Синтагматическое членение высказывания — один из идентификационных признаков, исследуемых лингвистом при просодическом анализе. При нормальном синтагматическом членении говорящим границы синтагм ставятся, исходя из оптимального удовлетворения требований семантико-синтаксического, фонетического и физиологического характера. Первое из названных требований предписывает объединение внутри синтагмы семантически связанных слов и запрещает их разъединение на две синтагмы. Второе требование выражает стремление фонетической системы языка к определённым ритмическим конструкциям (например, по 2–3 слова), объединяемым в одну синтагму. И, наконец, третье предписывает формирование синтагмы с числом слов не более того количества, которое можно успеть произнести за один такт выдоха.

Интонационный тип синтагмы — ещё один важный параметр, исследование которого составляет часть лингвистического анализа. Основными интонационными типами синтагм принято считать: завершённость, незавершённость, вопрос и восклицание. Количество интонационных подтипов для основных интонационных типов может достигать нескольких десятков [1–4]. Всё это сильно усложняет исследование и делает невозможным автоматическое распознавание интонационного типа синтагмы. Поэтому не вызывает сомнений потребность в инструментарии, который позволит получать просодические параметры для классификации интонационного типа синтагмы и проводить структурный анализ её мелодического контура, тем самым облегчая работу лингвиста при проведении фоноскопических экспертиз.

Цель работы — разработка методики, позволяющей автоматически определять интонационный тип синтагмы путём представления её основных интонационных параметров в численном виде. Для достижения цели поставлены и решены следующие задачи:

- сделать обзор методов анализа структуры интонационного контура, используемых современными системами идентификации говорящего;
- разработать и программно реализовать методику формирования просодических параметров, характеризующих интонационный тип синтагмы, в численном виде;
- осуществить проверку эффективности классификации на речевом материале, содержащем записи нескольких дикторов.

Обзор современных методов анализа структуры интонационного контура

Основными компонентами просодики являются интонация и ударение. Физически интонация и ударение реализуются совокупностью акустических средств (просодических характеристик речи), к числу которых относятся:

- мелодика — движение частоты основного тона (F0);
- ритмика — текущее изменение длительности звуков и пауз (T);
- энергетика — текущее изменение силы (амплитуды) звука (A).

Главным компонентом интонации является мелодика, которая описывается мелодическим контуром. *Мелодический контур* — характерная для речи картина изменения основного тона (ОТ), освобождённая от сегментных и позиционных влияний.

В литературе о русской интонации долгое время не было чёткого, обоснованного и последовательного разграничения и рассмотрения фактов языка и явлений речи. И всё же идея наличия интонационных моделей справедливо завоёвывала всё большее признание и распространение. В 60-х годах двадцатого столетия Е.А. Брызгунова предложила первую цельную систему описания русской интонации, отражающую наиболее частотные интонационные конструкции (ИК). Согласно работе Е.А. Брызгуновой [3], система состоит из семи ИК, определяемых следующими признаками синтагмы: три тональных признака — восходящий, нисходящий и ровный тоны (/, \, =), интенсивность на акценте (+), глоттализация в виде гортанной смычки (q). У каждой синтагмы есть свой «интонационный центр», который приходится на тот слог, на котором реализуется значимая перемена в высоте тона.

Обобщим вышесказанное в виде таблицы характеристик и различительных признаков всех семи типов ИК (см. табл. 1).

Таблица 1. Различительные признаки ИК, включающих предцентр и постцентр

Типы ИК	Типы предложений	Направление тона в центре ИК	Уровень тона в центре	Уровень тона в постцентре	Признаки для 2-х типов ИК
ИК-1	Повествовательное предложение	Нисходящее (= \)	Ниже предцентра	Ниже предцентра	См. ИК-2
ИК-2	Специальный вопрос, повеление	Нисходящее (= \+)	В пределах предцентра или незначительно ниже	Ниже предцентра	Усиление словесного ударения на гласном центра в отличие от ИК-1
ИК-3	Общий вопрос, незавершённость	Восходящее (/)	Выше предцентра	Ниже предцентра	См. ИК-7
ИК-4	Сопоставительный вопрос	Нисходяще-восходящее (V)	Ниже предцентра	Выше предцентра	
ИК-5	Восклицание	1-й центр — восходящее, 2-й центр — нисходящее (/=\)	Выше предцентра	Ниже предцентра	Увеличение длительности центров по сравнению с ИК-2
ИК-6	Оценочное восклицание	Восходящее (/=\)	Выше предцентра	Выше предцентра	
ИК-7	Экспрессивная оценка	Восходящее (/q)	Выше предцентра	Ниже предцентра	Смычка голосовых связок на гласном центра в отличие от ИК-3

Наиболее влиятельные подходы к способу анализа структуры интонационного контура изложены в [2, 5, 6]. В качестве ключевого элемента мелодического контура рассматривается центр. Именно он является единственным обязательным элементом контура, формирует его «лицо» и играет решаю-

щую роль при восприятии типа высказывания. Центр чаще всего реализуется в конце высказывания и, таким образом, совпадает с мелодическим завершением. В качестве основных параметров, различающих типы ИК, используются: интервал, диапазон, регистр (или тональный уровень), крутизна или скорость изменения тона, форма (вогнутый-выпуклый). С каждой из этих характеристик мелодического контура могут быть сопоставлены определённые интонационные значения. Таким образом, можно говорить о функциональной нагрузке признаков мелодического контура.

Эти основные мелодические параметры описываются набором признаков, которые получают в численном виде [7]:

- 1) начальная частота — значение первого отсчёта (в Гц) в начальной точке ядерного фрагмента контура;
- 2) конечная частота — значение последнего отсчёта (в Гц) в конечной точке ядерного фрагмента контура;
- 3) максимальная частота — максимальное значение частоты ОТ (в Гц) в пределах ядерного фрагмента контура;
- 4) минимальная частота — минимальное значение частоты ОТ (в Гц) в пределах ядерного фрагмента контура;
- 5) средняя частота — среднее значение частоты ОТ (ЧОТ) (в Гц) в пределах ядерного фрагмента контура;
- 6) время максимума — координата максимального значения ЧОТ в процентах от общей длительности ядерного фрагмента;
- 7) время минимума — координата минимального значения ЧОТ в процентах от общей длительности ядерного фрагмента;
- 8) время половинной частоты — координата значения половинной частоты (от интервала между максимумом и минимумом) в процентах от общей длительности ядерного фрагмента; частично соответствует описательной категории выпуклость-вогнутость;
- 9) интервал — разница между максимальным и минимальным значением частоты ОТ (в Гц и в полутонах);
- 10) скорость изменения тона — средняя скорость убывания или возрастания тона на выделенном участке контура в Гц/мсек., соответствует описательной категории «крутизна».

В основе используемой нами методики определения интонационного типа лежит модель интонационных портретов акцентных единиц (ПАЕ-модель) в виде мелодического, ритмического и энергетического оформления синтагмы. Эта модель была предложена Б.М. Лобановым более 20 лет назад [8] и с тех пор успешно использовалась во многих моделях синтеза речи по тексту [9–11], в том числе многоязычных [12]. Этот подход позволяет получать тип ИК синтагмы автоматически, определяя просодические характеристики речи диктора, тем самым значительно облегчая задачу лингвисту при проведении просодического анализа в идентификационном исследовании.

Формирование просодических параметров, характеризующих интонационный тип синтагмы

Ниже будут изложены основные идеи предлагаемой методики. Просодические параметры, используемые для определения интонационного типа, базируются на мелодическом, энергетическом и ритмическом контурах синтагмы. Для получения мелодической кривой в данной работе использовался выделитель ОТ, изложенный в [13]. Этот метод позволяет получить достаточно точную оценку ЧОТ, которая плавно меняется во времени в соответствии с изменениями голоса. Другим достоинством этого метода является его устойчивая работа с сигналом, содержащим шум.

Чтобы повысить эффективность классификации типа ИК, необходимо получить крупные и плавные изменения (подъёмы, падения и более сложные конфигурации) ЧОТ, реализующиеся в пределах слогов, слов и синтагм. Для этого осуществляется обработка ме-



лодического контура, состоящая в удалении микропросодических явлений и сглаживании.

Мелодический, энергетический и ритмический контуры представляют собой временные ряды, размерность которых зависит от темпа произнесения фраз. Поэтому классификация типа ИК по полученным контурам проводилась с помощью метода нелинейного растяжения-сжатия оси времени — Dynamic time warping (DTW) [14, 15].

Результатом работы выделителя ОТ по сигналу $S = \{s_j\}_{j=0}^{N_s-1}$ является временной ряд квазипериодов $P = \{p_i\}_{i=1}^N$, по которым определяются соответствующие им ЧОТ $F = \{F_i\}_{i=1}^N$. Прежде чем анализировать и описывать мелодический контур, выделяя общие тенденции изменения ЧОТ, необходимо удалить его искажения, вызванные микропросодическими явлениями.

Этап 1 — обработка мелодического контура

Мелодический контур имеет сложную структуру. Соседние периоды ОТ, как правило, отличаются по величине друг от друга, и эти различия передают разную информацию. В поперечном графике изменения ЧОТ по времени, построенном с достаточной точностью, наблюдаются постоянные флюктуации частоты, т.е. быстрые и сравнительно небольшие по диапазону колебания значений параметра на фоне более медленного и плавного его изменения. Кроме того, в определённых местах речевой цепи наблюдаются кратковременные, но значительные по величине отклонения от изменяющейся по определённому закону кривой ЧОТ. Эти возмущения, или пертурбации, связаны обычно с участками резких формантных переходов на границах звуков (в отличие от флюктуаций, которые затрагивают и квазистационарные участки звуков) и в определённой степени отражают сегментный состав отрезка речи, на котором реализуется тот или иной интонационный рисунок. Вариации рисунка мелодического контура, вызванные особенностями сегментного состава фразы, фиксируются аппаратурой, но человеческим ухом не воспринимаются. Такие перцептивно и лингвистически незначимые вариации получили название микроинтонации или микропросодики. Так, например, известно, что наличием ОТ характеризуются только сегменты, произносимые с участием голоса. Таким образом, значительная часть звуков, т.е. все глухие согласные, вообще не имеют собственных мелодических характеристик, но оказывают влияние на соседние сегменты. Так, гласный, следующий за глухим согласным (особенно за глухим щелевым), характеризуется более высокой частотой ОТ, чем соответствующий гласный, следующий за звонким согласным. При этом частотный пик гласного в первом случае приходится на начальный участок (артикуляции), а во втором — на средний. К микропросодическим явлениям относятся также значительные колебания ЧОТ на протяжении дрожащего сонанта, обусловленные резкой переменной в соотношении подсвязочного и надсвязочного давлений.

Помимо сегментного состава фразы на рисунок мелодического контура влияют и другие факторы. Поскольку мелодический контур является отражением биомеханической активности голосовых связок, получаемая кривая колебаний обладает определённой степенью диспериодичности: соседние периоды практически никогда не бывают полностью идентичны друг другу. Некоторые участки речевой последовательности могут характеризоваться значительными отклонениями. Обычно это начало и конец процесса фоциации.



Некоторые изменения ЧОТ на различных участках во фразе могут вызываться также чисто психофизиологическими, не подконтрольными говорящему факторами.

Как свидетельствуют экспериментальные данные, всю эту информацию слушатель не учитывает (отбрасывает) при восприятии высказывания. Человеческое ухо играет роль своеобразного фильтра, пропуская лишь ту информацию, которая необходима для правильной интерпретации высказывания.

Итак, основываясь на результатах экспериментов в области восприятия интонации, большинство исследователей пошло по пути отделения макропросодии (лингвистически релевантных изменений ЧОТ) от микропросодии (контекстно обусловленных вариаций, относящихся к уровню речепроизводства) и исключения микропросодических характеристик из рассмотрения. Стандартный подход к решению этой проблемы — использование сглаживания мелодического контура. Однако часто скачки ЧОТ достаточно сильные, и, чтобы их удалить, требуется использовать процедуры сглаживания с большим окном сглаживания в 100 и более мс. Это приводит к искажению мелодического контура в силу сильного усреднения и влечёт за собой потерю существенной информации. Поэтому удалять микропросодические явления целесообразно по следующему алгоритму.

1. Определяется начало микропросодического явления, т.е. текущий период под номером i считается началом микропросодического явления, если его ЧОТ более чем на один полутон отличается от ЧОТ предыдущего.

В данной работе используется шкала полутонов натурального строя, в которой отношение частот двух тонов, образующих полутон, равно 15:16. Номер полутона N , которому соответствует частота F , вычисляется по формуле:

$$N \approx 34 + 12 \left(\log_2 F - \log_2 \left(\frac{55}{2} \left(1 + \sqrt[12]{2} \right) \right) \right)$$

Следовательно, с помощью функции

$$I(F_1, F_2) = \left| 12 \log \frac{F_1}{F_2} \right|$$

можно определить интервал между частотами в полутонах. Таким образом, если $I(F_i, F_{i-1}) < 1 \wedge I(F_i, F_{i+1}) > 1$, то i -й период считаем началом микропросодического явления.

2. Определяется длина микропросодического явления (в периодах): микропросодическое явление продолжается до тех пор, пока разница ЧОТ последовательно идущих друг за другом периодов не будет меньше одного полутона: если

$$I(F_{i+L+1}, F_{i+L}) > 1 \wedge \forall j = 1, \dots, L-1 \ I(F_{i+j+1}, F_{i+j}) > 1,$$

то длина микропросодического явления равна L .

3. Проводится процедура сглаживания ЧОТ в рамках микропросодического явления, которая заключается в соединении прямой линией в мелодическом контуре частоты периода, который предшествует данному микропросодическому явлению, и частоты периода, следующего за микропросодическим явлением:

$$F_n = F_{i-1} + n \frac{F_{i+L+1} - F_{i-1}}{L}, \quad n = 1, \dots, L$$

После удаления микропросодических явлений мелодический контур сглаживается. Целью сглаживания является устранение ложных скачков траектории ЧОТ, чтобы отразить

основные тенденции изменения ЧОТ в мелодическом контуре. Поскольку сильные флуктуации, связанные с микропросодикой, уже были удалены, то процедура сглаживания не требует больших окон обработки. В данной работе использовалось сглаживание по трём точкам.

Полученный таким образом сглаженный мелодический контур обрабатывался далее с целью выделения акустических признаков сигнала, необходимых для получения ПАЕ и классификации их интонационных типов.

Этап 2 — формирование портретов акцентных единиц

После процедур выделения ОТ и обработки мелодического контура для синтагмы получаем временной ряд квазипериодов и соответствующих им ЧОТ, а также отсчётов сигнала:

$$P = \{p_i\}_{i=1}^N, \quad F = \{F_i\}_{i=1}^N, \quad S = \{s_j\}_{j=0}^{N_s-1},$$

где N — количество квазипериодов, которое содержится в синтагме, N_s — длина

$$\text{сигнала, причём } \sum_{i=1}^N p_i = N_s.$$

На основе полученных последовательностей P, F, S строим в виде временного ряда энергетический контур синтагмы:

$$A = \{A_i\}_{i=1}^N, \quad A_i = \frac{n_i + p_i - 1}{n_i} \sum_{k=1}^{i-1} s_k^2, \quad n_i = \sum_{k=1}^{i-1} p_k, \quad i = 2, \dots, N, \quad n_1 = 0,$$

где A_i — энергия сигнала на i -ом периоде, n_i — начало i -го периода.

При этом необходимо учесть микропросодикой, как это было сделано при получении последовательности периодов ОТ, а именно сгладить энергетический контур в рамках микропросодического явления длиной L периодов:

$$A_n = A_{i-1} + n \frac{A_{i+L+1} - A_{i-1}}{L}, \quad n = 1, \dots, L,$$

где i — номер периода, соответствующий началу просодического явления.

Ритмический контур синтагмы:

$$T = \{T_i\}_{i=1}^N, \quad T_i = n_{\min} - n_i,$$

где n_{\min} — начало периода, соответствующего максимальной ЧОТ, в последовательности P (как правило, этот период минимальный).

Величина T_i характеризует временное расстояние от i -го периода до начала акцентного выделения. Она может принимать отрицательные значения для периодов, следующих за ядерным слогом.

Для того чтобы контуры отображали динамику изменения соответствующих характеристик независимо от интенсивности сигнала и его продолжительности, необходимо выполнить нормировку.

$$FO = \{FO_i\}_{i=1}^N, \quad FO_i = \frac{F_i}{\max_j F_j},$$

$$AO = \{AO_i\}_{i=1}^N, \quad AO_i = \frac{A_i}{\max_j A_j},$$

$$TO = \{TO_i\}_{i=1}^N, TO_i = \frac{T_i}{\max_j T_j}.$$

Итак, последовательность:

- $F0$ представляет собой F0-ПАЕ;
- $A0$ представляет собой A0-ПАЕ;
- TO представляет собой T0-ПАЕ.

Общий портрет АЕ представляет собой временной ряд, каждый элемент которого — вектор:

$$PAE = \{PAE_i = (F0_i, A0_i, TO_i)\}_{i=1}^N.$$

Для определения интонационного типа синтагмы и оценки эффективности просодических параметров в данной работе предлагается использовать метод DTW, позволяющий сравнивать временные ряды.

Исследование эффективности классификации

С целью оценки различительного потенциала параметров ПАЕ было проведено численное исследование описанных выше характеристик для семи типов ИК, представляющих собой синтагмы. Речевой материал составили записи десяти различных дикторов, мужчин и женщин с разными голосовыми данными. Каждый диктор наговаривал обучающий и контрольный наборы, состоящие из семи типов ИК. Речевые сигналы для обеспечения максимального приближения их характеристик к исходным аналоговым сигналам записывались в формате WAV PCM с частотой дискретизации 22050 Гц, глубиной битности 16 бит. Запись осуществлялась в монорежиме с помощью программы Audacity 1.3.12-beta. Сегментация речевых сигналов на синтагмы проводилась вручную.

Для проверки дикторнезависимости и эффективности классификации типов ИК на речевом материале обучающего набора каждого k -го диктора был сформирован словарь эталонов интонационных типов W_k .

$$W_k = \{PAE_j\}_{j=1}^7, k = 1, \dots, 10,$$

где PAE_j — последовательность ПАЕ, полученная для j -го типа ИК.

Материал для тестирования содержал все контрольные наборы всех дикторов. В таблице 2 приведены средние ошибки первого рода по каждому интонационному типу, полученные в результате проведенного тестирования по стандартной методике, использующей десять признаков, описанных в первой части статьи, и по предложенной методике. Классификация по ПАЕ проводилась на разных словарях эталонов методом DTW, по стандартным признакам — методом k ближайших соседей (в качестве функции близости использовалось скалярное произведение векторов) [16]. Поскольку пары ИК-1, ИК-2 и ИК-3, ИК-7 не отличаются друг от друга направлением тона, уровнями тона в центре и пост-центре, то каждая из пар была объединена в один интонационный тип.

Таблица 2. Ошибки (в %) первого рода при классификации интонационного типа по ПАЕ и стандартным признакам

Типы ИК	Направление тона	Ошибки классификации, %	
		ПАЕ	Ст. признаки
ИК-1, 2	Нисходящее: (=), (=+)	3	6
ИК-3, 7	Восходящее: (/), (/q)	5	9
ИК-4	Нисходяще-восходящее (V)	4	8
ИК-5	Восходяще-нисходящее (/≠)	6	10
ИК-6	Восходящее (/≠)	8	12



Полученные результаты показали более высокую разделительную способность ПАЕ при классификации интонационного типа синтагмы по сравнению со стандартными признаками (точность повысилась более чем на 3%). Как видно из таблицы, точность определения ИК, содержащих нисходящий мелодический тон, выше, чем ИК с восходящим и ровным тоном. Худшие результаты относятся к классификации синтагм с ровным тоном. Полученные результаты и преобладание синтагм с нисходящим тоном объясняется общей мелодической деклинацией в синтагме, которая накладывается на мелодический тип синтагмы. Так, в синтагмах с ровным мелодическим типом всё равно присутствует понижение частоты основного тона от начала синтагмы к концу.

Выводы

Анализируя полученные результаты, можно сделать вывод о перспективности использования мелодического, энергетического и ритмического контуров для определения интонационного типа синтагмы в рамках просодического анализа. Повышению точности классификации способствовала обработка мелодического контура с целью удаления микропросодических явлений. Численные исследования показали повышение точности определения интонационных типов более чем на 3%, при использовании ПАЕ по сравнению со стандартными признаками темперированной музыкальной шкалы для задания границ полос пропускания фильтров, на основе которых вычисляются спектральные признаки, характеризующие речевой поток в целом.

Предложенный подход к определению интонационного типа применим не только в задачах фоноскопической экспертизы цифровых фонограмм. На его основе можно создать модули, которые будут являться неотъемлемой частью систем синтеза речи, определения эмоционального состояния говорящего и биометрической идентификации, что свидетельствует о большой практической значимости данной работы.

Литература

1. Артёмов В.А. Метод структурного анализа речевой интонации. М., 1962.
2. Брызгунова Е.А. Интонация / Е.А. Брызгунова // Русская грамматика, М., 1980. Т.1. С. 96–123.
3. Брызгунова Е.А. Звуки и интонация русской речи. М., 1977.
4. Светозарова Н.Д. Интонационная система русского языка. Л., 1982.
5. O'Connor J. Intonation of colloquial English / J. O'Connor, G. Arnold. London: Longman, 1973. — 576 p.
6. Hart J. A Perceptual Study of Intonation: An experimental-phonetic approach to speech melody / J. Hart, R. Collier, A. Cohen. Cambridge: Cambridge University Press, 1990. 380 p.
7. Смирнова Н.С. Идентификация дикторов на основе сравнения параметров реализации мелодических контуров высказываний. Доклад на конф. Диалог. 2007.
8. Lobanov B. The phonemophon text-to-speech system // International Congress of Phonetic Sciences: proc. of the 11-th seccion ICPhS'87, Tallin, USSR, 6–10 August 1987. Tallin, 1987. V.1. P. 120–124.
9. Лобанов Б.М. Микроволновой синтез речи по тексту // Анализ и синтез речи: сб. науч. трудов / научн. ред. Б.М. Лобанов. Мн.: Институт технической кибернетики АН БССР, 1991. С. 57–73.



Ермоленко Т.В.

Автоматическая классификация типа мелодических контуров синтагм при проведении просодического анализа

10. Lobanov, B., Tsiurulnik, L., Sizonov, O. AUP's Modeling of Speaker Specific Intonation Contour Peculiarities // Speech and Computer: proceedings of the 12-th International conference SPECOM'2007, Moscow, Russia, 15–18 October, 2007 / Moscow State Linguistic University. Moscow, 2007. V.1. P. 312–317.
11. Лобанов Б.М., Елисеева О.Е. Речевой интерфейс интеллектуальных систем. Мн.: БГУИР, 2006. 152 с.
12. Lobanov, B., Tsiurulnik, L. Development of multi-voice and multilanguage TTS synthesizer (languages: Belarussian, Polish, Russian) // Speech and Computer: proceedings of the 11-th International conference SPECOM'2006, St. Petersburg, Russia, 25–29 June, 2006 / Institute of Informatics and Automation of RAS, Speech Informatics Group. St.-Petersburg: Anatolia, 2006. P. 274–283.
12. Бабкин В.В. LPC вокодер 1000–1200 бит/с // Труды 3-й межд. конф. Цифровая Обработка Сигналов и её Применение (DSPA-2000) М., 2000.
13. Rabiner L. Fundamentals of Speech Recognition / L. Rabiner, B.-H. Juang. Prentice Hall PTR, 1993. 507 p.
14. C. S. Myers and L. R. Rabiner. A comparative study of several dynamic time-warping algorithms for connected word recognition//The Bell System Technical Journal, 60(7):1389–1409, September, 1981.
15. Журавлев Ю.И., Рязанов В.В., Сенько О.В. Распознавание. Математические методы. Программная система. Практические применения. М.: Фазис, 2006, 176 с.

Сведения об авторах

Ермоленко Татьяна Владимировна —

кандидат технических наук, научный сотрудник отдела распознавания речевых образов Института проблем искусственного интеллекта МОН и НАН Украины, г. Донецк. Распознаванием и обработкой речевых сигналов занимается с 2002 года.