

СИСТЕМА ЭКСПРЕСС-ИДЕНТИФИКАЦИИ ГОЛОСА ЛИЧНОСТИ МЕТОДОМ КЛОНИРОВАНИЯ АКУСТИЧЕСКИХ ХАРАКТЕРИСТИК РЕЧИ

А.Г. Давыдов, В.В. Киселёв, Б.М. Лобанов, Л.И. Цирульник

*Объединённый институт проблем информатики НАН РБ
Беларусь 220012, Минск, ул. Сурганова, 6
Тел. (017)284-2773, факс (017)231-8403
Эл.почта: lobanov@newman.bas-net.by*

Создаваемые в правоохранительных органах учетные массивы звуковых баз данных (фонограммы речи потенциальных подозреваемых) могут достигать объёма в десятки тысяч голосов. Оперативный поиск нужных образцов голоса по некоторой контрольной записи речи произвольного содержания при таких объёмах весьма затруднён. В докладе предлагается новый метод оперативного поиска фонограмм речи в учетном массиве и система идентификации голоса личности.

Система оперативной идентификации голоса личности путём автоматического доступа к специализированным звуковым базам данных (СЗБД) построена на базе технологии клонирования голоса [1,2] и состоит из трех подсистем, представленных на рис. 1-3.

На первом этапе процесса идентификации необходимо создать “цифровые” портреты индивидуальности голоса личности (ЦПГЛ), фонограммы которых хранятся в специализированной звуковой базе данных (СЗБД). Эту задачу решает подсистема (рис. 1) путем автоматической разметки фонограмм СЗБД на фонетические элементы (слова и аллофоны). Основой системы являются многоголосый синтезатор речи по тексту и программные средства нелинейного по времени сопоставления естественных и синтезированных речевых сигналов методами динамического программирования (ДП-сопоставление). Результатом работы подсистемы является создание “цифровых” портретов индивидуальности голоса и артикуляции речи личности (ЦПГЛ) для каждого из голосов СЗБД, путём считывания отмеченных реализаций аллофонов в естественной речи, их отбора по различным критериям и записи их параметрических описаний.

На втором этапе производится оперативный синтез индивидуальных образцов речи по заданному тексту на основе наборов ЦПГЛ, хранящихся в СЗБД, и в результате создается база индивидуальных образцов речи (БИОР). Этот процесс отображен на рис.2. При поступлении фонограммы речевого сигнала неизвестной личности, речевое сообщение переводится в соответствующий орфографический текст. Орфографический текст речи исследуемой фонограммы преобразуется в аллофонный текст, размеченный на фонетические слова и синтагмы. Главной решаемой задачей на этом является оперативный синтез из наборов ЦПГЛ, хранящихся в СЗБД, индивидуальных образцов речи, соответствующих полученному аллофонному тексту, и создание на их основе базы индивидуальных образцов речи (БИОР), соответствующих тексту исследуемой фонограммы.

На последнем, третьем, этапе реализуется собственно процесс идентификации голоса личности по имеющейся фонограмме речевого сигнала (см. рис. 3). Многокритериальные оценки и принятие вероятностных идентификационных решений осуществляется на основе фонетических и словесных мер сходства, получаемых с помощью ДП-процедур сопоставления исследуемых сигналов. Получаемые оценки позволяют реализовать отбор голосов по заданному числу претендентов или по заданной вероятности принадлежности к голосу конкретной личности.

Создание ЦПГЛ и БИОР с помощью базовой программы синтеза речи по произвольному тексту принципиально позволяет организовать процесс идентификации голоса личности по фонограмме речи произвольного содержания путем последовательного сопоставления параметрических описаний из базы индивидуальных образцов речи, сгенерированных под текст исследуемой фонограммы, с параметрическим описанием собственно предъявленной фонограммы.

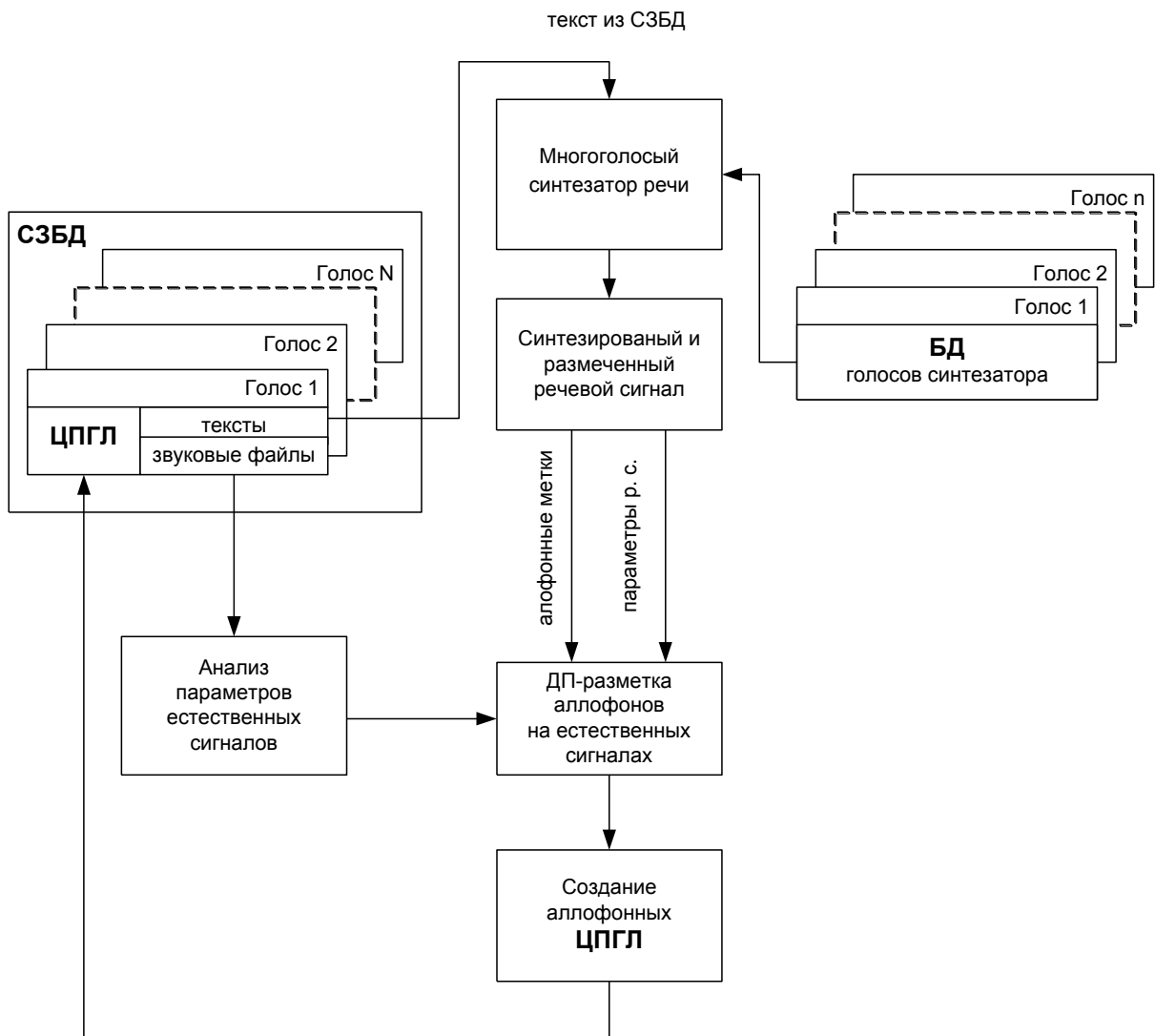


Рис.1. Структурная схема автоматизированной системы создания наборов персональных “цифровых” портретов голоса личности – ЦПГЛ

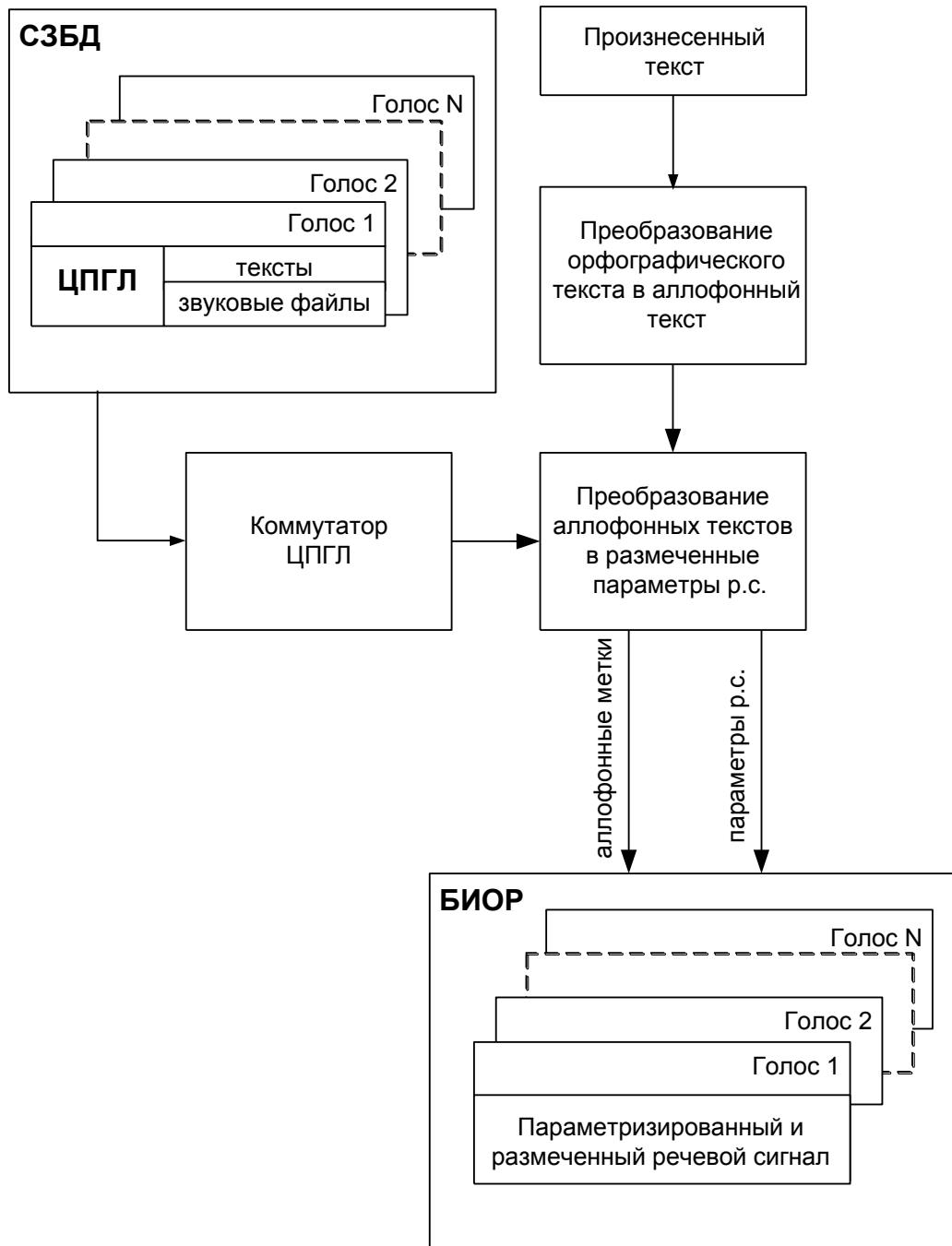


Рис.2. Структурная схема системы оперативного синтеза индивидуальных образцов речи

- СЗБД – специализированная звуковая база данных
- ЦПГЛ – цифровые портреты голосов личности
- БИОР – база индивидуальных образцов речи

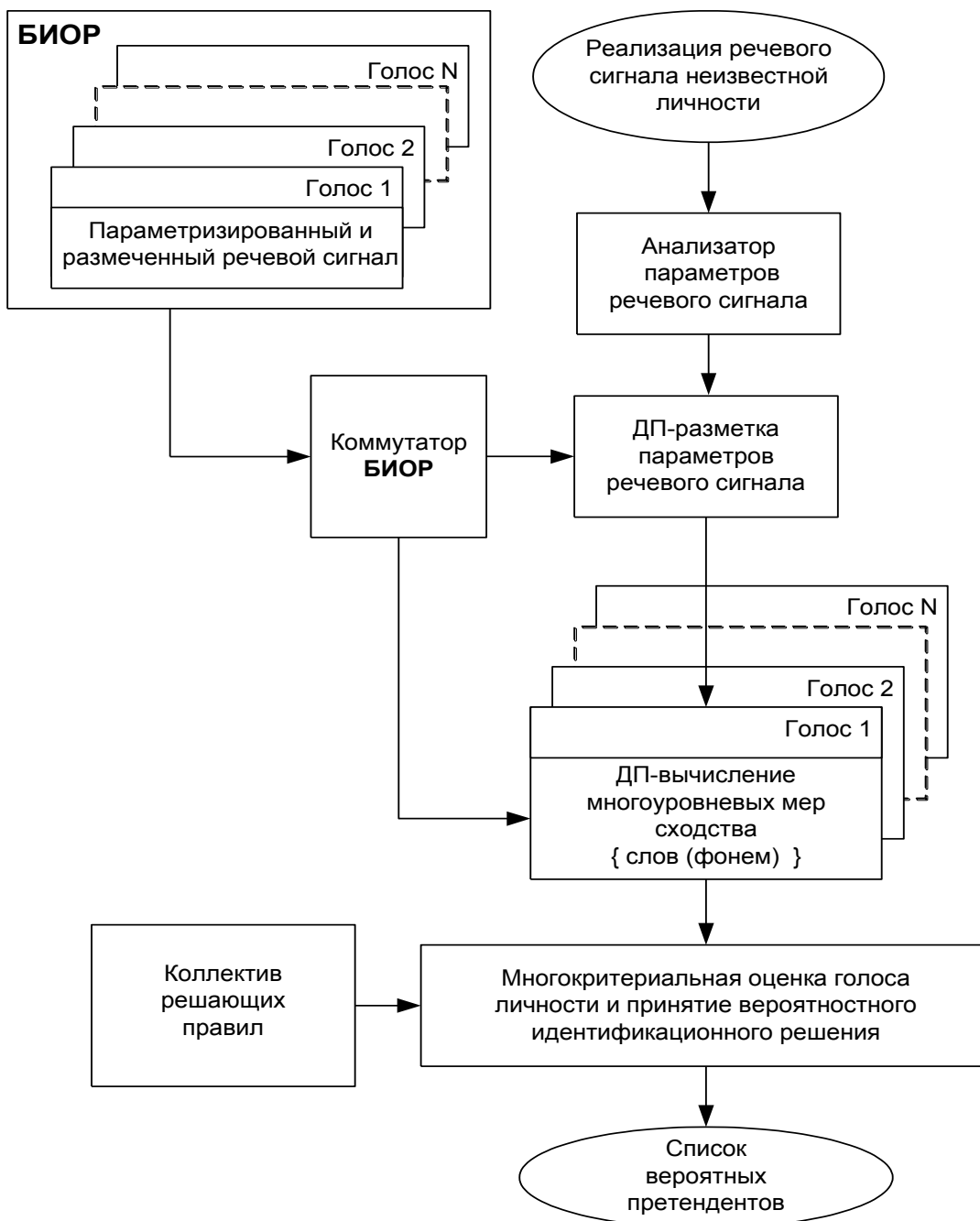


Рис.3. Структурная схема системы автоматической идентификации личности по предъявленному отрезку речи

Литература

1. Б.М.Лобанов, Е.Б. Карневская, Т.В. Левковская. Синтезатор речи по тексту как компьютерное средство “клонирования” персонального голоса. Тр. Международной конференции Диалог-2001, Москва, 2001, сс. 265-272
2. Б.М. Лобанов, В.В. Киселёв. Автоматизация клонирования персонального голоса и дикции для систем синтеза речи по тексту. Тр. Международной конференции Диалог-2003, Москва, 2003, сс 417-424.