

## СИСТЕМЫ МАСКИРОВАНИЯ ГОЛОСА ГОВОРЯЩЕГО

А.С. РЫЛОВ, Ю.А. ВЕЖИК, А.С. ВИСКУП, Д.В. МОТУЗ

Исследования в области разработки систем маскирования (изменения) голоса (СМГ) говорящего представляют сегодня целое направление в речевой информатике. Все СМГ можно разделить на две разновидности.

К первой разновидности относятся СМГ для изменения качества голоса и могут использоваться для защиты прав свидетеля на суде или для удовлетворения требования человека быть неузнанным при публичном раскрытии какой-либо информации. Проведенный анализ показал, что эти системы реализуются, как правило, на основе вокодерных систем, работающих как в частотной, так и во временной областях. Также установлено:

1) системы, работающие в частотной области, позволяют гибко регулировать акустические параметры как тоновые, так и форматные, но имеют недостаток, связанный с возникновением фазовых искажений при больших значениях коэффициента модификации параметров. Это ощущается в виде реверберационных эффектов воспринимаемых на слух. Кроме того, реализация этих систем более сложная, чем систем, работающих во временной области;

2) системы, работающие во временной области, позволяют легко нелинейно изменять частоту основного тона без изменения формантной структуры. Модификация формантной структуры затруднена, однако применение систем анализа-синтеза речи

на основе метода линейного предсказания с выделением линейных спектральных пар и с последующим их преобразованием с помощью билинейного Z-преобразования позволяет преодолеть этот недостаток.

Ко второй разновидности СМГ относятся системы, которые позволяют трансформировать голос истинного диктора в голос определенного диктора (диктора-"мишени"). При этом осуществляется адаптация параметров речи истинного диктора к параметрам речи диктора-"мишени" для определенного речевого фрагмента (предложения, фразы), которые были произнесены обоими дикторами. К наиболее сложным морфологическим системам относятся системы с распознаванием речи произвольного диктора. В этом случае речевой морфинг осуществляется без предварительной фазы адаптации к параметрам определенного диктора-"мишени", но принадлежащими определенной речевой единице. Предварительной фазой в таких системах является распознавание речи от произвольного диктора, а затем выделение параметров у истинного диктора для распознанного речевого фрагмента от диктора-"мишени", и только после этого осуществляется адаптация параметров обоих дикторов и синтез речи похожий на голос диктора-"мишени". В системах этого типа при синтезе речи целесообразно применять трехмерную анимацию лица диктора-"мишени" с целью повышения правдоподобности речевого морфинга.

Системы, относящиеся ко второй разновидности СМГ, как правило, используются для создания систем синтеза речи по тексту, которые могут говорить голосом любого диктора. Кроме того, они могут использоваться в качестве специальных систем.