

УДАЛЕНИЕ ИЗ РЕЧЕВОГО СИГНАЛА СЛЕДОВ ИНЫХ ФОНООБЪЕКТОВ

Д.т.н. В.Р.Женило

Удаление из речевого сигнала следов фонообъектов, мешающих правильному восприятию речи или снижающих качество ее звучания, - достаточно актуальная задача. К сожалению, для ее решения многие методы фильтрации, в том числе и адаптивной, иногда оказываются недостаточно эффективными. Предлагается метод, основанный на расслоении всех следов фонообъектов, присутствующих в заданном сигнале, на две группы, с последующим восстановлением двух сигналов во временной области таким образом, чтобы каждый из них содержал следы лишь одной группы.

Основу метода составляет синтез сигнала по следам, отраженным в его динамическом спектральном описании. В зависимости от конкретного типа задачи, существует два способа ее решения. Первый – выделить в исследуемом сигнале следы интересующего фонообъекта и только по ним синтезировать новый сигнал. Второй способ – выделить в исследуемом сигнале следы мешающих фонообъектов, синтезировать по ним сигнал-помеху и затем вычесть его из исходного сигнала. Выбор одного из этих способов в значительной мере определяется относительными уровнями мощности следов интересующего речевого сигнала и следов иных фонообъектов. Причем, в ряде случаев, если это очень значительно, то выделение одной группы следов может производиться автоматически. В противном случае может оказаться, что задачу расслоения следов разных фонообъектов лучше отдать эксперту.

Важной технологической тонкостью реализации такого подхода является точность динамического параметрического описания исходного сигнала¹ и точность синтеза речевого сигнала по выделенным параметрам. Если все расчеты производить корректно и в полном объеме, то можно получить синтезированный речевой сигнал, мало отличающийся от исходного. Для иллюстрации на рис. 1 и 2 показаны фрагменты сонограмм и осциллограмм исходного речевого сигнала и синтезированного по всему множеству следов, имеющихся в исходном сигнале.

¹ Подробнее см. «Компьютерная фоноскопия» (1995).

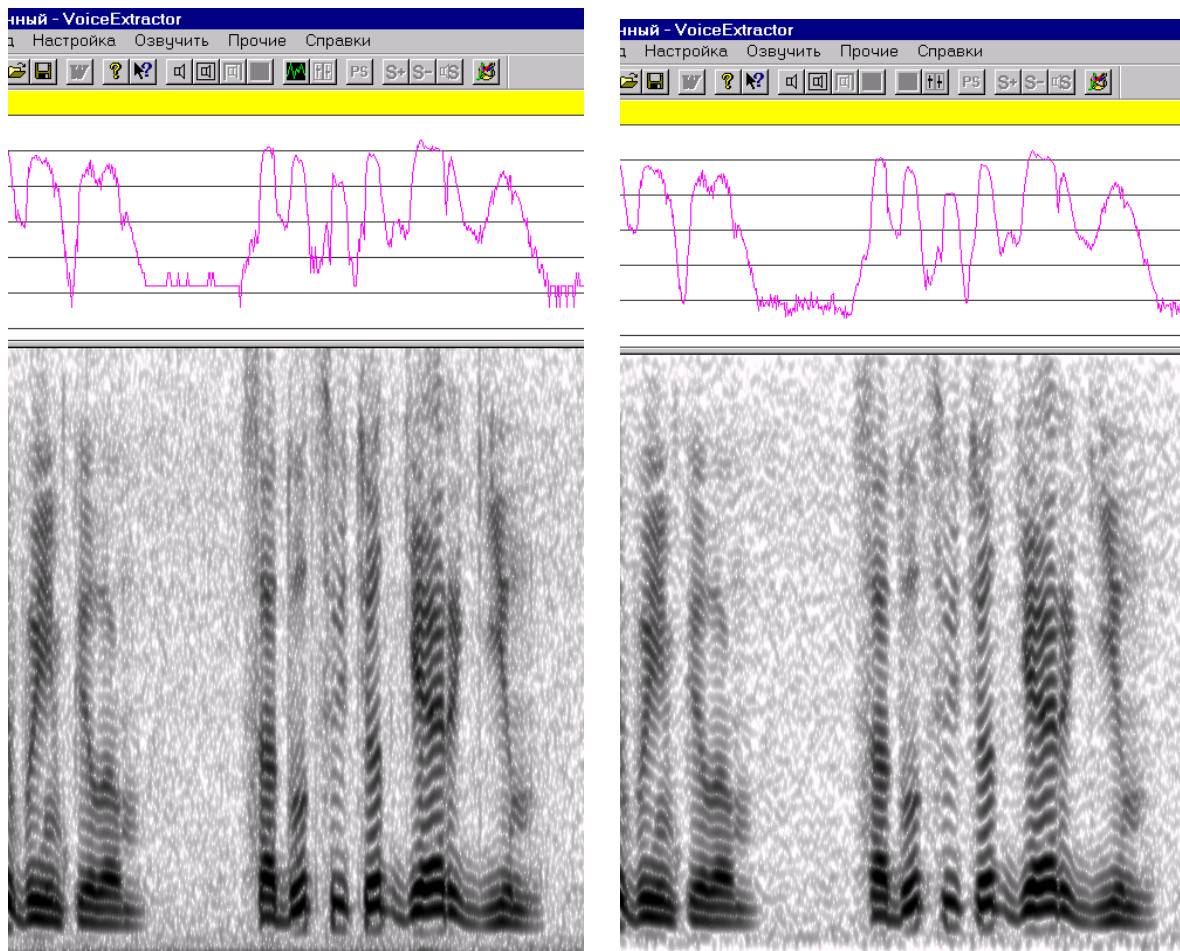


Рис. 1. Сонограмма исходного речевого сигнала (слева) и синтезированного по всему множеству следов (справа).



Рис. 2. Осциллограмма фрагмента исходного речевого сигнала (вверху) и синтезированного по всему множеству следов (снизу).

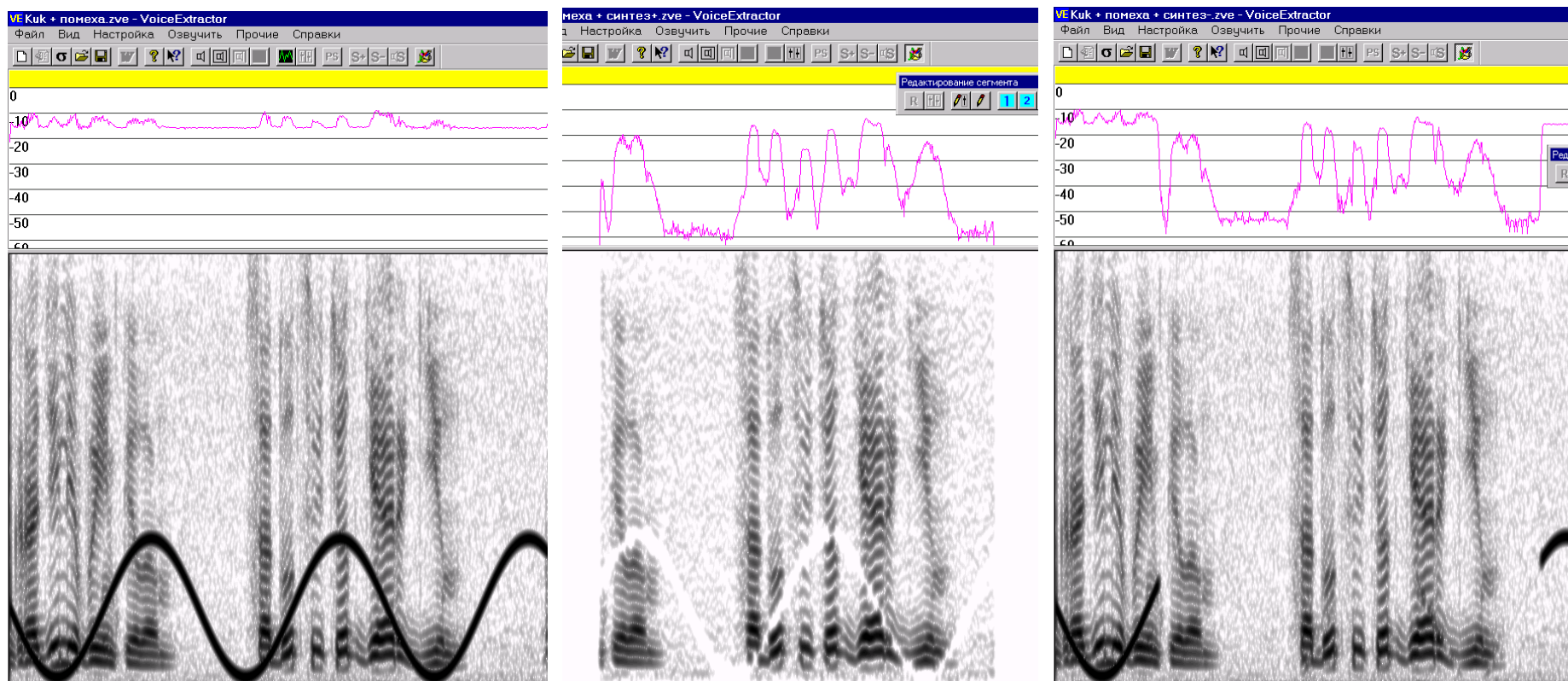


Рис. 3. Два способа удаления помехи из исходного сигнала (слева):
 а) синтез нового сигнала по следам, явно не относящимся к помехе (посередине);
 б) вычитание сигнала-помехи (справа), синтезированной по ее следам, из исходного сигнала.

На рис. 3 показан пример результата расслоения следов речевого сигнала и мощной частотно модулированной помехи двумя методами: а) методом синтеза речевого сигнала по всему множеству исходных следов, из которого удалены следы мощной помехи и б) методом синтеза сигнала-помехи и вычитания ее из исходного сигнала. Из рисунка видно, что в данном случае второй метод более эффективен, поскольку он практически не затрагивает следы полезного речевого сигнала (они не проходят через весь механизм синтеза), оставляя его практически аутентичным. И хорошо видно, что если первый метод не восстанавливает маскированные слабые следы речи, то второй – делает их видимыми и, стало быть, слышимыми. Последнее делает предложенную технологию достаточно перспективной.

Хорошее совпадение во временной области синтезированного сигнала (по его следам) с оригиналом позволяет получить технологию вычитания мощных помех из сигнала, в котором речевой сигнал на своих самых мощных участках лежит ниже уровня помех более чем на 20 дБ. Для примера на рис. 4 показаны сонограммы и уровни мощности исходного речевого сигнала, заглушенного мощными частотно-модулированными помехами, из-за которых следы речи вообще не слышны. Там же показан обработанный сигнал, в котором для иллюстрации на малом участке удалены все следы мощных помех. При этом уровень отрисовки сонограммы на рисунке усилен так, чтобы было видно: на месте бывших помех не появились белые полосы, и

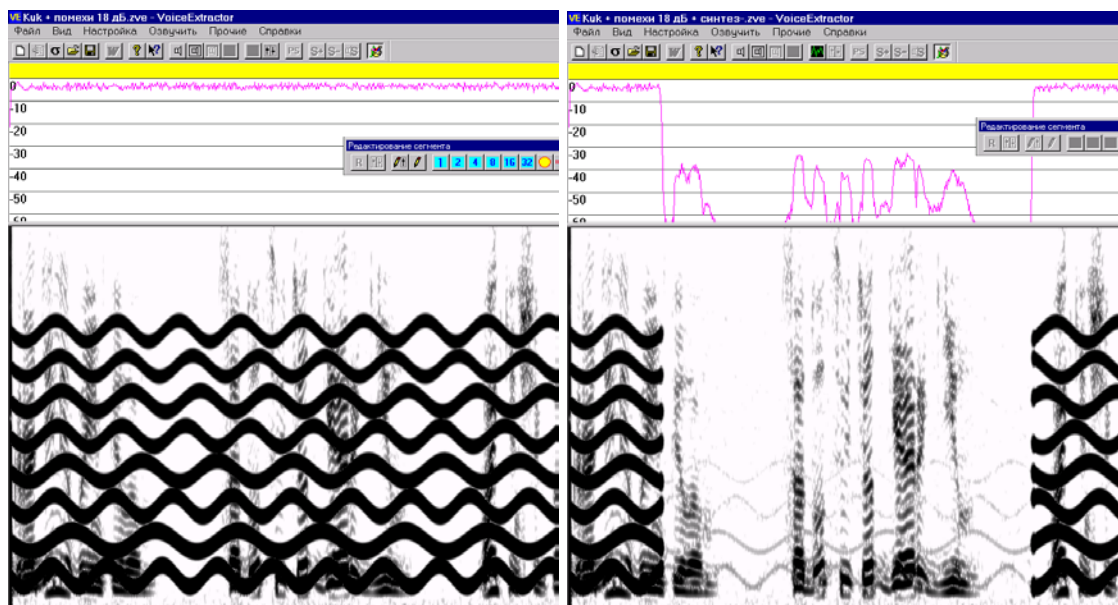


Рис. 4. Неслышимая, полностью заглушенная помехами, речь (слева), и результат удаления помех с полным восстановлением разборчивости.

сонограмма самого речевого сигнала мало отличается от оригинала на рис. 1. Звучание восстановленного фрагмента после его усиления имеет весьма удовлетворительное качество.

Следует отметить, что такие результаты могут быть достигнуты лишь в тех случаях, когда следы речевого сигнала реально присутствуют в исследуемом сигнале, а не потеряны в результате клиппирования, сильных нелинейных искажений или когда полезный сигнал слабее помехи более того, что в принципе может зафиксировать аппаратура.

Предложенные технологии могут применяться при решении широкого класса задач, начиная от реставрации архивных фонограмм с сохранением их аутентичности, и заканчивая протоколированием речи, заглушенной мощными помехами.