

Исходя из вычисленных коэффициентов корреляции и анализа форм исходного и восстановленного сообщения, можно сделать вывод о достаточно высоком качестве восстанавливаемого сообщения

Литература

1. Гантмахер, Ф.Р. Теория матриц. М., Физматлит, 2004. 560 с.
2. Хорн Р., Джонсон Ч. Матричный анализ. М., Мир, 1989. 655 с.
3. Прохоренко, Е.И., Болдышев А.В., Фирсова А.А., Эсауленко А.В. Метод оптимального субполосного преобразования в задаче сжатия речевых данных. – «Вопросы радиоэлектроники», сер. ЭВТ, 2010, вып. 1, с. 49-55.

Работа выполнена при поддержке гранта РФФИ 12-07-00514

Статья поступила 09.12.2011

Д.т.н., проф. Е.Г. Жилияков, А.В. Курлов
(НИУ «БелГУ»)

E.G. Zhilyakov, A.V. Kurlov

КОМПЬЮТЕРНАЯ КОМПЕНСАЦИЯ ЭХО ИСКАЖЕНИЙ НА ОСНОВЕ РЕШЕНИЯ ИНТЕГРАЛЬНЫХ УРАВНЕНИЙ

COMPUTER ECHO DISTORTION COMPENSATION BASED ON THE SOLUTION OF INTEGRAL EQUATIONS

В статье предложена компьютерная процедура компенсации эхо искажений, основанная на решении интегральных уравнений типа свертки. С целью проверки работоспособности проведены вычислительные эксперименты, показывающие влияние частоты дискретизации на погрешность восстановления речевого сигнала.

Ключевые слова: речь, компенсация эхо, интегральные уравнения, аппаратная функция.

In the article the computer echo distortion compensation procedure, which is based on the solution of integral equations of convolution type. In order to verify that the computational experiments carried out, showing the influence of sampling frequency on the error recovery of the speech signal.

Key words: speech, echo distortion, integral equations, apparatus function.

Введение

Возникновение эффекта акустического эхо-сигнала представляет собой проблему, которая решается не один десяток лет и сохраняет свою актуальность до сегодняшнего дня[1]. Акустическое эхо является общей проблемой для всех устройств громкой связи, мобильных телефонов, переносных телекоммуникационных устройств, систем голосового управления и громкой связи для автомобилей, устройств для аудио конференций. Оно возникает при отражении звука от предметов в помещении, в случае расположения микрофона и динамика в одном корпусе, при попадании сигнала из вашего динамика в микрофон. В этом случае собеседник слышит не только речь другого абонента, но и свою собственную, вернувшуюся к нему как акустическое эхо. Искажения подобного рода в значительной степени ухудшают качество передаваемой речи, снижают разборчивость и повышают утомляемость слушателя.

Математические основы

В статье в качестве метода, позволяющего компенсировать эхо искажения, предлагается следующее. В качестве модели воздействия эхо будем использовать интегральное соотношение

$$u(t) = \int_0^t R(t, \tau)x(\tau)d\tau, \tag{1}$$

Здесь $u(t)$ - сигнал на выходе микрофона, $x(\tau)$ – входные акустические воздействия, обусловленные различными факторами, в качестве которых можно указать отражения от предметов, находящихся в помещении, где генерируется звук. Эти предметы создают эхо, которое воздействует на микрофон. $R(t, \tau)$ - функция, с помощью которой учитывается наличие эхо-сигналов. Эту функцию принято называть аппаратной функцией

В предположении постоянства свойств отражающих звук предметов, аппаратную функцию можно считать зависящей от разности аргументов, когда имеют место

$$R(t, \tau) = \begin{cases} R(t - \tau), 0 \leq \tau \leq t \\ 0, \tau > t \end{cases} \tag{2}$$

Ноль в качестве нижнего предела интегрирования означает начало генерации звука. Задача компенсации эха сводится к восстановлению входного воздействия $x(\tau)$ по зарегистрированным значе-

ниям отклика $u(t)$. В математическом отношении эта задача означает решение интегрального уравнения вида свертки.

$$u(t) = \int_0^t R(t-\tau)x(\tau)d\tau, \quad 0 \leq t \leq T \quad (3)$$

где $u(t)$ - предполагается известной (зарегистрированные значения). Аппаратная функция определяется конфигурацией расположения микрофона, говорящего, других предметов в помещении.

Определение аппаратной функции можно осуществлять эмпирически, регистрируя на выходе микрофона сигнал, возникающий при акустических воздействиях, образующихся в результате короткого звукового сигнала, который создается либо хлопком в ладони, либо иным способом. Короткий звуковой сигнал создается в различных помещениях при неподвижном микрофоне и результаты усредняются.

Решение задачи компенсации эхо, т.е. восстановление исходного речевого сигнала, заключается в решении интегрального уравнения вида (3). При этом следует иметь ввиду, что реально отклик $u(t)$ может быть зарегистрирован только в конечном наборе точек интервала генерации, т.е. в наличии имеется только набор значений

$$u_i = u(t_i), \quad i = 0, 1, \dots, N, \quad (4)$$

которое в соответствии с (4) удовлетворяет соотношениям

$$u_i = \int_0^{t_i} R(t_i - \tau)x(\tau)d\tau, \quad i = 0, 1, \dots, N. \quad (5)$$

Таким образом, для восстановления сигнала $x(\tau)$ можно воспользоваться только таким конечным набором интегральных соотношений.

Положим

$$\varphi_i(\tau) = R(t_i - \tau) \quad (6)$$

Согласно (2) выполняются условия

$$\varphi_i(\tau) \equiv 0, \quad \tau \geq t_i. \quad (7)$$

Тогда, в соответствии с основными положениями функционального анализа [2], некоторую функцию можно представить в виде суммы

$$x(\tau) = x_1(\tau) + x_2(\tau), \quad (8)$$

а норма сигнала $x(\tau)$ равна

$$\|x\|^2 = \|x_1\|^2 + \|x_2\|^2. \quad (9)$$

В выражении (8) компонента x_1 является элементом

$$x_1(\tau) = \sum_{k=0}^N \beta_k \varphi_k(\tau), \quad (10)$$

а вторая компонента ортогональна по всем функциям вида (6), т.е. имеет место

$$(x_2, \varphi_i) = \int_0^T \varphi_i(\tau) x_2(\tau) d\tau = 0, \quad i = 0, 1, \dots, N. \quad (11)$$

С учетом этого подстановка представления (8) в соотношение (5) дает

$$u_i = \int_0^T \varphi_i(\tau) x_1(\tau) d\tau, \quad (12)$$

т.е. информация о второй компоненте в отклике отсутствует.

Таким образом, по результатам регенерации отклика восстановить можно только ту компоненту воздействия, которая описывается соотношением (10). Такие решения интегральных уравнений принято называть решениями, представимыми через ядро.

Коэффициенты $\beta_i, i=0, \dots, N$, используемые в представлении (10), определяются выражением

$$\bar{\beta} = A^{-1} \bar{u} \quad (13)$$

$$\bar{u} = (u_0, \dots, u_N)^T, \quad A = \{a_{ik}\};$$

$$a_{ik} = \int_0^T \varphi_i(\tau) \varphi_k(\tau) d\tau, \quad i, k = 1, \dots, N \quad (14)$$

Отметим, однако, что среди собственных чисел матрицы A могут быть близкие к нулю. Тогда необходимо использовать специальные приемы регуляризации, позволяющие получить устойчивые к воздействиям помех компоненты вектора β .

Вычислительные эксперименты

Для эксперимента в качестве исходного использовался речевой сигнал, представляющий собой фразу «Заяц» с частотой дискретизации 22050Гц, 16 бит, моно, рис. 1а и соответствующая ему

трансформанта Фурье показана на рис. 1б. Аппаратная функция была получена в помещении, где проявляется эффект эхо при помощи хлопка в ладоши, отклик зафиксирован при помощи микрофона и устройства записи звука, рис. 2а и соответствующая ему трансформанта Фурье показана на рис. 2б. Параметры сигнала: частота дискретизации 22050Гц, 16 бит, моно.

Выражение (3) в дискретном виде может быть представлено как

$$\bar{u}(kM\Delta\tau) = \Delta\tau \sum_{i=0}^{kM} R((kM - i)\Delta\tau) \bar{x}(i\Delta\tau), \quad (15)$$

где: $\Delta\tau$ - интервал дискретизации, M -- номер выборки, $x(i\Delta\tau)$ - вектор значений исходного сигнала без воздействия эхо, $u(kM\Delta\tau)$ - вектор значений смоделированной свертки. Задача эксперимента заключалась в определении значения параметра M , при котором погрешность восстановленного сигнала $x_i(\tau)$ в (8) получается приемлемой.

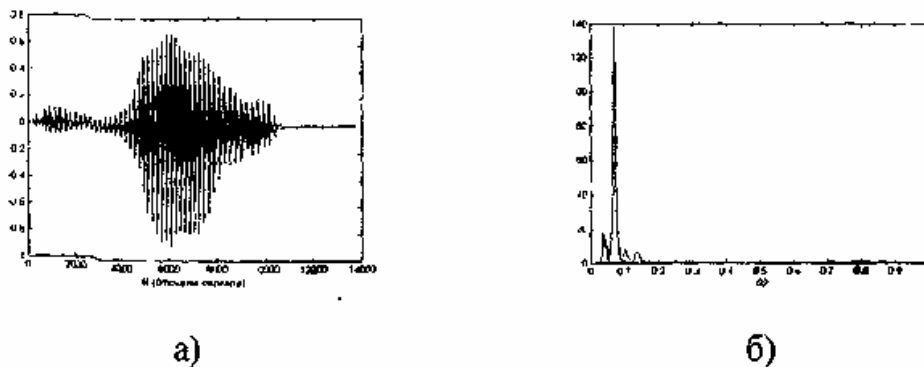


Рис. 1

Исходный сигнал во временной области (а) и соответствующая ему трансформанта Фурье (б)

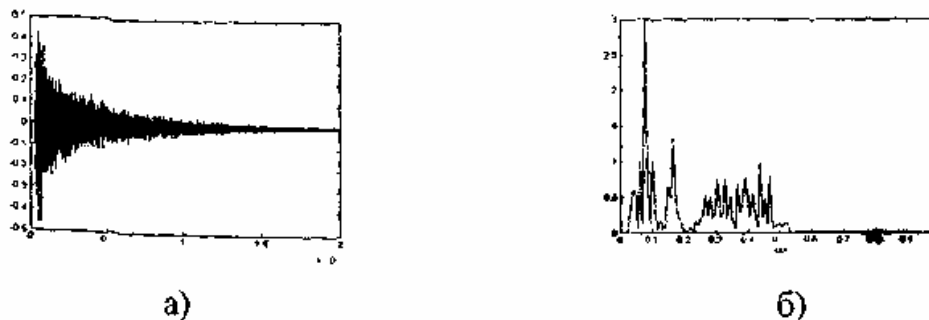


Рис 2

Аппаратная функция во временной области (а) и соответствующая ей трансформанта Фурье (б)

АЛГОРИТМЫ И ПРОГРАММЫ

Для оценки погрешности восстановления речевого сигнал рассчитывалось среднеквадратическое отклонение (СКО) по формуле:

$$\sigma = \sqrt{\sum_{i=1}^N (x_i - ay_i)^2 / \sum_{i=1}^N x_i^2}, \quad a = \sum_{i=1}^N x_i y_i / \sum_{i=1}^N y_i^2. \quad (16)$$

где: x_i - исходный сигнал с отсутствием эффекта эхо, y_i - восстановленный сигнал, a - поправочный коэффициент, N - длительность сигнала. Результаты эксперимента представлены на рисунках ниже.

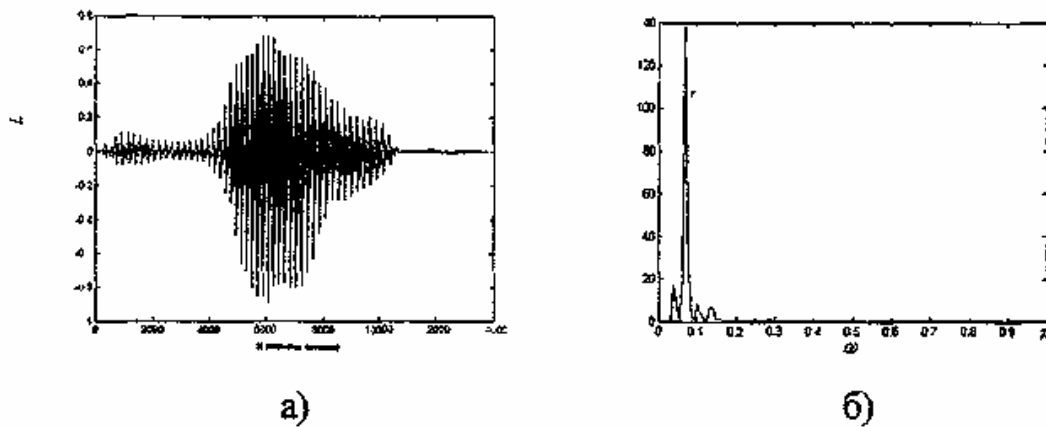


Рис 3

Восстановленный сигнал во временной области при частоте дискретизации 22050Гц (а) и соответствующая ему трансформанта Фурье (б), $\sigma=0.0078$

Из рис. 3 видно, что восстановленный сигнал практически идентичен исходному (рис.1а), о чем свидетельствует достаточно малое значение σ .

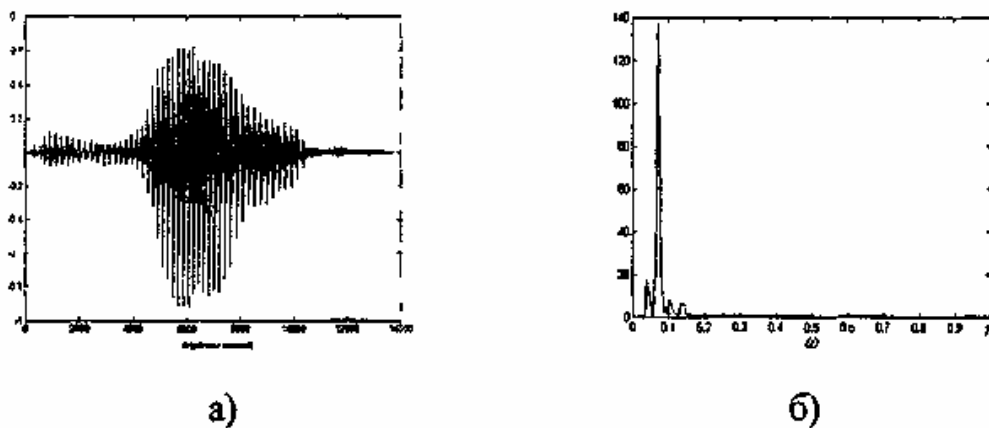


Рис.4

Восстановленный сигнал во временной области при частоте дискретизации 11025Гц (а) и соответствующая ему трансформанта Фурье (б), $\sigma=0.1642$

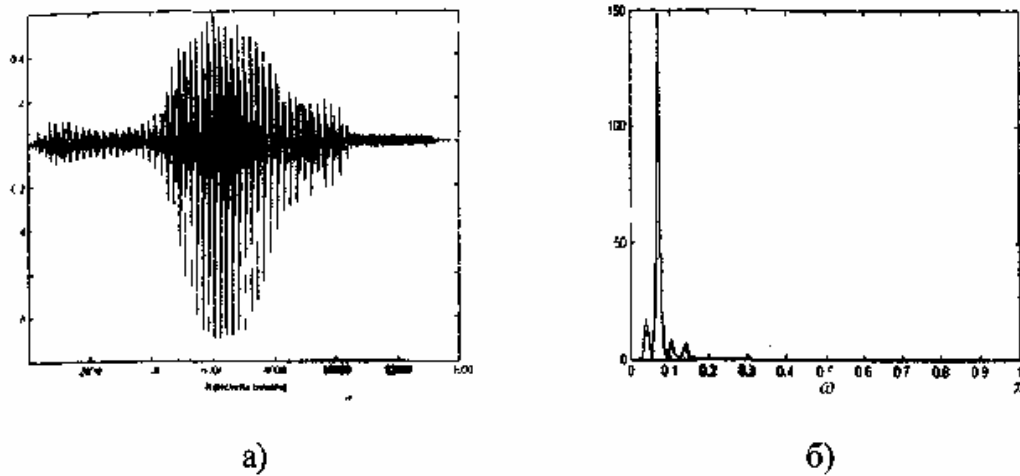


Рис 5

Восстановленный сигнал во временной области при частоте дискретизации 7350Гц (а) и соответствующая ему трансформанта Фурье (б), $\sigma = 0.2523$

Результат восстановления речевого сигнала, показанный на рис. 4-и 5, все еще является приемлемым

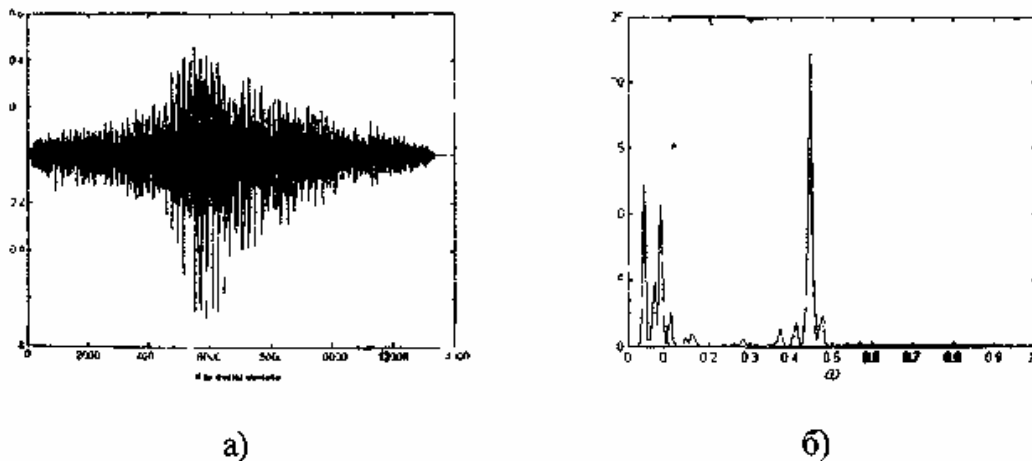


Рис 6

Восстановленный сигнал во временной области при частоте дискретизации 5512Гц (а) и соответствующая ему трансформанта Фурье (б), $\sigma = 0.6908$

Показанный на рис 6 результат при частоте дискретизации 5512 Гц ($M=4$) свидетельствует о появлении искажений в восстановленном сигнале. Это соответствует случаю, когда в выражении (9) $\|x_2\|^2 > \|x_1\|^2$. Таким образом, приемлемый результат восстановления речевого сигнала достигается при значениях $M=1,2,3$. Уже при $M=4$ погрешность достигает неприемлемого значения $\sigma = 0.6908$.

Литература

1. Кузнецов Е.П., Витязев В.В. Цифровая обработка сигналов в задачах эхо-компенсации. – "Цифровая обработка сигналов", 2006, № 3, с. 8-19.
2. Ректорис К. Вариационные методы в математической физике и технике. М., Мир, 1985. 590 с.

Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры для инновационной России» на 2009-2013 гг, гос контракт № 14 740 11 0390

Статья поступила 09 12 2011

Д.т.н., проф. Е.Г. Жиляков, П.Г. Лихолоб (НИУ «БелГУ»)

E.G. Zhilyakov, P.G. Likholob

**СПОСОБ СКРЫТОГО ИНДЕКСИРОВАНИЯ РЕЧЕВЫХ
СООБЩЕНИЙ В МУЛЬТИМЕДИЙНЫХ БАЗАХ ДАННЫХ**

**A WAY TO HIDE INDEXING VOICE MESSAGES
MULTIMEDIA DATABASES**

В статье предложена модель создания метки индекса элемента мультимедийной базы данных. Описан способ, внедряющий метку цифрового водяного знака в речевые данные.

In this paper we propose a model to label the element index multimedynoy database. Describes a method that implements label digital watermark in the speech data. The method meets the proposed criteria characterizing features of the introduction of CEH in speech data.

Ключевые слова: стеганография, цифровой водяной знак, цифровая метка, субполосная матрица, собственный вектор, часть энергии, внедрение, устная речь, речевые данные.

Key words: steganography, digital watermarking, digital label, subpolosnaya matrix, the eigenvector of the energy introduction, speaking, speech data.

Одной из распространенных форм обмена информацией, представляющих собой данные и знания, является устная речь. Для хранения и накопления информации в виде мультимедийных баз, аудио-сигналы устной речи преобразовывают в кодовые комбина-