

- диагностика психофизического потенциала пациента;
- профессиональный отбор;
- семейное консультирование;
- регистрация измененных состояний сознания.

Исследования различных жидкостей, веществ и материалов:

- оценка структурных изменений в жидкофазных объектах;
- исследование пищевых продуктов;
- энергетические свойства растений, камней и минералов.

В настоящее время разработан ряд модификаций программно-аппаратных комплексов для исследования параметров газоразрядного свечения объектов различной природы. На сегодняшний день методом ГРВ только в нескольких Санкт-Петербургских клиниках обследовано более 2000 пациентов. Работы ведутся на базе СПб гос. Института Точной Механики и Оптики, Института Биохимии АН РФ, СПб гос. Медицинского Университета им. И.П. Павлова, СПб Военно-Медицинской Академии, Онкологического Центра Грузии, СПб НИИ Спорта, СПб Академии Спорта, Скандинавского международного Университета.

УДК 621.3.06

Г.Н. Метан

ПРИМЕНЕНИЕ НЕЙРО-НЕЧЕТКОЙ ТЕХНОЛОГИИ ANFIS В ЗАДАЧАХ ФИЛЬТРАЦИИ ЗВУКОВЫХ СИГНАЛОВ*

Введение. Эволюция средств хранения аудиоинформации от аналоговых (пластинка, магнитная лента) к цифровым (компакт-диск и др.), а также эволюция студийных аппаратных, а позже программных средств поставили вопрос о цифровой реставрации звука – восстановлении старых записей и переиздании их на современных носителях. Основная работа реставратора – подавление шипения и шелчков и (по возможности) восстановление исходного звука. В данной работе под фильтрацией подразумевается подавление шумов, т.е. задача реставрации решается только в первой ее части.

Сегодня широко распространены средства реставрации на базе метода спектрального анализа (основной алгоритм – быстрое дискретное преобразование Фурье). В данной работе предложена и протестирована более перспективная технология реставрации (цифровой адаптивный фильтр), которая предполагает обработку конкретного поступающего сигнала в два этапа:

- а) распознавание сигнала (с помощью аппарата нейросетей);
- б) реакция на сигнал по нечетким правилам, заложенным разработчиком.

Дальнейшая эволюция средств реставрации должна привести, на наш взгляд, к появлению в будущем активных цифровых средств (программных агентов). Это будут средства, манипулирующие хранилищами звука и библиотеками шумов с использованием определенной стратегии и опирающиеся на гибридные интеллектуальные технологии [Тарасов, 2001]. Такие устройства уже нельзя будет отнести собственно к фильтрам, скорее их можно назвать автоматизированными системами

*Работа выполнена при финансовой поддержке РФФИ по совместному гранту РФФИ-БРФФИ, проект № 02-01-81037

нелинейного звукового монтажа, способными при необходимости работать в автономном режиме (без участия человека).

1. Обоснование необходимости разработки новых средств шумоподавления. Теоретической основой фильтров, реализованных программным путем в составе звуковых редакторов, служит спектральный анализ. Этот подход развивается уже более 50 лет и является классическим. Здесь алгоритмическую основу формируют быстрые преобразования Фурье (БПФ). Сегодня на основе этого метода строятся программно-аппаратные комплексы для монтажных студий. Примерами служат продукты фирм Steinberg (www.steinberg.de), Sonic Foundry (www.sonicfoundry.com), difitec (www.difitec.de) и множество дополнительно подключаемых звуковых модулей третьих фирм.

В такой, ставшей уже классической постановке задачи, звуковой сигнал рассматривается с точки зрения электротехники, а его обработка проводится с помощью электрических схем (или их компьютерных моделей). Такие схемы в принципе не могут содержать знания об элементарных свойствах звука, в частности, волновой природе звука, тембре, гармониках и пр. Поэтому их эффективность в решении задачи подавления шумов весьма ограничена. А ведь есть и более сложные задачи. К тому же такие средства принципиально не дают возможность осуществлять автоматизацию деятельности человека. Применение спектрального анализа требует аккуратной настройки множества параметров в зависимости от конкретной задачи и специфики анализируемых сигналов, а для правильной интерпретации результатов нужен некоторый практический опыт.

В работе предложен новый нейро-нечеткий подход к реставрации звуков, опирающийся на современные достижения искусственного интеллекта и направленный на повышение эффективности фильтрации.

2. Основы нейро-нечеткого подхода. Ключевым преимуществом нечеткой логики является ее способность описывать желаемое поведение системы простыми «если-то»-правилами [Заде, 1976]. При этом можно использовать все доступные инженерные ноу-хау, для непосредственного улучшения системы. Во многих приложениях это дает более простое решение за меньшее время. Однако, во многих случаях, знания, описывающие желаемое поведение системы, содержатся в наборах данных. Здесь разработчик получает «если-то»-правила из наборов данных вручную. Большие объемы данных требуют больших вычислений. В этом случае хорошо подходит методология нейронных сетей, которые обучаются на наборах данных.

Для объединения сильных сторон обоих подходов создаются гибридные методологии и соответствующие инструментальные системы. Одну из них – ANFIS (Adaptive Network Based Fuzzy Inference System) [Ярушкина, 2001] – мы взяли на вооружение. В новой постановке задачи речь идет о создании адаптивного фильтра, который сначала обучается распознавать звуковой образ шума, а затем, получая на вход смесь полезного сигнала и шума, распознает шум и вычитает его (здесь используются операции нечеткой арифметики [Дюбуа и Прад, 1990]).

Общая методика создания нейро-нечетких моделей включает следующие шаги:

- 1) Получение наборов данных для обучения. В данных содержатся неявные знания, которые предстоит выявить с помощью нейронной сети.
- 2) Создание системы нечеткой логики. Здесь производится подбор функций принадлежности, нечетких правил и метода дефазификации.

3) Выбор метода обучения и настройка его параметров. Метод обучения определяет, как по рассогласованию между реальным и желаемым выходом системы должны изменяться правила и функции принадлежности.

4) Собственно обучение.

5) Оптимизация и проверка.

3. Алгоритм обучения. Обучение нейронной сети производится методом обратного распространения ошибки. Это наиболее известный метод обучения многослойной нейросети без обратных связей. Он является итеративным градиентным алгоритмом. Его суть – распространение сигналов ошибки от выходов нейронной сети к ее входам в направлении, обратном прямому распространению сигналов в обычном режиме работы [Короткий, 2000]. При обучении ставится задача минимизации ошибки нейронной сети, которая определяется методом наименьших квадратов. Для сети с одним выходом (как в нашем случае) ошибка определяется соотношением:

$$E = \frac{1}{2} \sum_{j=1}^p (y_j - d_j)^2, \quad (1)$$

где y_i – значение выхода нейросети для каждого компонента входного вектора (дискретных звуковых отсчетов в нашем случае), а d_j – желаемое значение выхода нейросети.

Полный алгоритм обучения нейронной сети с помощью процедуры обратного распространения строится так:

1) Подать на входы сети один из возможных образов и в режиме обычного функционирования нейронной сети, когда сигналы распространяются от входов к выходам, рассчитать значения последних. Напомним, что выход нейросети рассчитывается по формуле

$$s_j^{(n)} = \sum_{i=0}^M y_i^{(n-1)} \cdot w_{ij}^{(n)}, \quad (2)$$

где M – число нейронов в слое $n-1$ с учетом нейрона с постоянным выходным состоянием $+1$, задающего смещение;

$y_i(n-1) = x_{ij}(n)$ – i -й вход нейрона j -го слоя n ;

$y_j(n) = f(s_j(n))$, где $f()$ – функция активации нейрона;

$y_q(0) = I_q$, где I_q – q -я компонента вектора входного образа;

2) Рассчитать $\delta^{(N)}$ для выходного слоя (слоя N)

$$\delta_l^{(N)} = \frac{\partial E}{\partial y_l} \cdot \frac{dy_l}{ds_l} = (y_l^{(N)} - d_l) \cdot \frac{dy_l}{ds_l}. \quad (3)$$

3) Рассчитать изменения весов $\Delta w(N)$ слоя N .

$$\Delta w_{ij}^{(n)} = -\eta \cdot \delta_j^{(n)} \cdot y_i^{(n-1)}, \quad (4)$$

где η – коэффициент скорости обучения, $0 < \eta < 1$.

4) Рассчитать по формулам (3) и (4) соответственно $\delta^{(n)}$ и $\Delta w^{(n)}$ для всех остальных слоев, $n = N-1, \dots, 1$.

5) Скорректировать все веса в нейронной сети

$$w_{ij}^{(n)}(t) = w_{ij}^{(n)}(t-1) + \Delta w_{ij}^{(n)}(t), \quad (5)$$

где t – номер текущей итерации, n – номер слоя.

6) Если ошибка сети больше заданной или число итераций обучения не достигло заданного предела, то перейти к шагу 1. В противном случае – конец.

4. Описание эксперимента. В нашей работе адаптивный фильтр был построен на основе нейронечеткой архитектуры ANFIS с использованием модуля Fuzzy Logic Toolbox в среде MATLAB 6.1. Для построения графиков использованы команды из Signal Processing Toolbox.

Порядок обработки звука таким фильтром следующий:

- 1) Подготовить в виде файла исходный материал.
- 2) Выбрать тип удаляемого шума: шипение или щелчки.
- 3) Обучить фильтр на небольшом (0.1..0.5 сек) участке.
- 4) Подать на вход фильтра весь исходный материал.
- 5) Результат записать в новый файл (при необходимости контроля прослушать исходный и полученный результаты).

На первом этапе исследования для выбранного типа полезного сигнала и шума выявлялось, как влияет изменение структуры и параметров простого адаптивного фильтра, построенного по гибридной методологии ANFIS, на качество фильтрации. Ключевые особенности: шкала времени – дискретная, шкала значений – непрерывная. На втором этапе изменялся тип полезного сигнала при неизменных источнике шума и структуре ANFIS. Новый тип полезного сигнала – волновой пакет. Ключевые особенности: сигнал ограничен во времени и по частоте. Ниже рассмотрим подробнее формулировку задачи.

Первый этап. Пусть имеется тестовый синусоидальный сигнал x_0 на интервале времени от 0 до 6 секунд с частотой дискретизации F_s 1000 Гц. По теореме Котельникова, максимально возможная частота – $F_s/2$, т.е. 500 Гц. Таким образом, пусть за 6 секунд частота полезного сигнала меняется от 0,5 кГц до 20 Гц. Такой сигнал называют частотно-модулируемым (ЧМ). Выбор сигнала связан с необходимостью определить эффективность фильтрации в широком диапазоне частот.

Пусть также имеется источник шума: псевдослучайные числа с нормальным распределением. Такой шум называется широкополосным или белым. Параметры генератора случайных чисел: среднее значение - 0, среднее квадратическое отклонение - 1.

$$n1 = \text{randn}(\text{time}). \quad (6)$$

Шум накладывается на источник по сложному нелинейному закону:

$$n2 = \frac{4 \cdot \sin(n1_{d0}) \cdot n1_{d1}}{1 + n1_{d1}^2}, \quad (7)$$

где: $n1_{d0}$ – сигнал в момент времени i , $n1_{d1}$ – сигнал в момент времени $i-1$ (исходный сигнал с задержкой).

Измеренный (входной) сигнал имеет вид:

$$m = x_0 + n2. \quad (8)$$

Этот сигнал поступает на вход адаптивного фильтра на основе ANFIS. Будем полагать, что полезный сигнал $n2$ "загрязнен" сигналом x_0 . Иными словами, на входе ANFIS имеем сигнал $n1$, а на выходе - сигнал m . Обучение ставит целью научить ANFIS воспроизводить функцию $n2(n1)$. Тогда предполагаемый вид сигнала x_f можно получить из соотношения (8):

$$x_f = m - n2(n1)_f, \quad (9)$$

где $n2(n1)_f$ – предполагаемый вид сигнала ошибки – результат обучения нейро-нечеткой модели (предполагается, что порядок нелинейности сигнала шума известен и составляет 2).

Создаем модель ANFIS с двумя входами для $n1_{d0}$ и $n1_{d1}$ соответственно (рис.1). На каждый вход назначаются по две колоколообразные функции принадлежности.

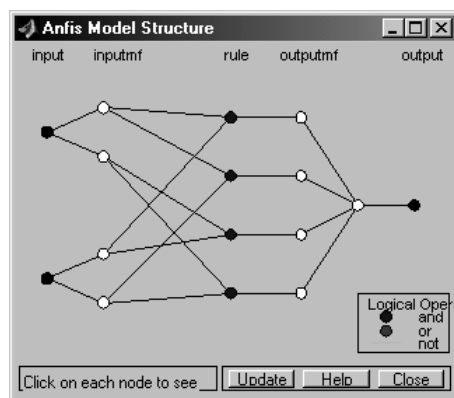


Рис.1. Структура модели ANFIS

Обучение производится по алгоритму обратного распространения ошибки. Параметры этого алгоритма: количество эпох (итераций алгоритма) – 10; целевая функция – минимизация ошибки. Начальный размер шага обучения установим равным 0.2.

Теперь можно начинать фильтрацию сигнала. Результат выполнения фильтрации обученным фильтром приведен на рис.2 и 3. Самый лучший способ определить эффективность алгоритма шумоподавления – это прослушать результаты. За отсутствием такой возможности приводятся результаты, полученные двумя способами, общепринятыми в сфере анализа электроакустических сигналов: временное представление и периодограмма. Можно предоставить файлы формата wav в ответ на запрос по электронной почте, а также исходники программ MATLAB.

При увеличении начального шага обучения до 0.5 улучшений не произошло. Увеличение периодов обучения с 10 до 30 привело к некоторому улучшению качества распознавания (но и к увеличению времени). Оставив число периодов обучения 10, увеличим количество функций принадлежности на каждый вход с 2 до 4. Полученный результат лучше предыдущего и гораздо лучше первоначального. Однако затраченное время возросло с 12 до 220 секунд.

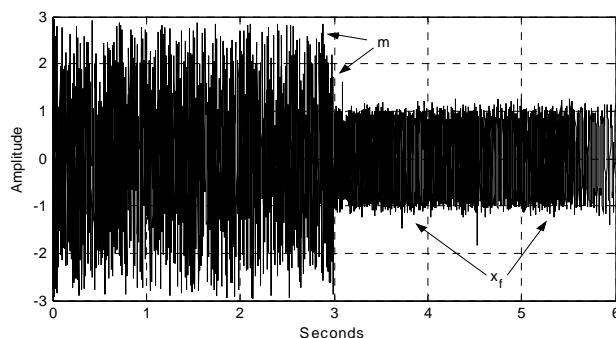


Рис.2. Сигналы: входной (m) [0..3 с], выходной (x_f) [3..6 с]

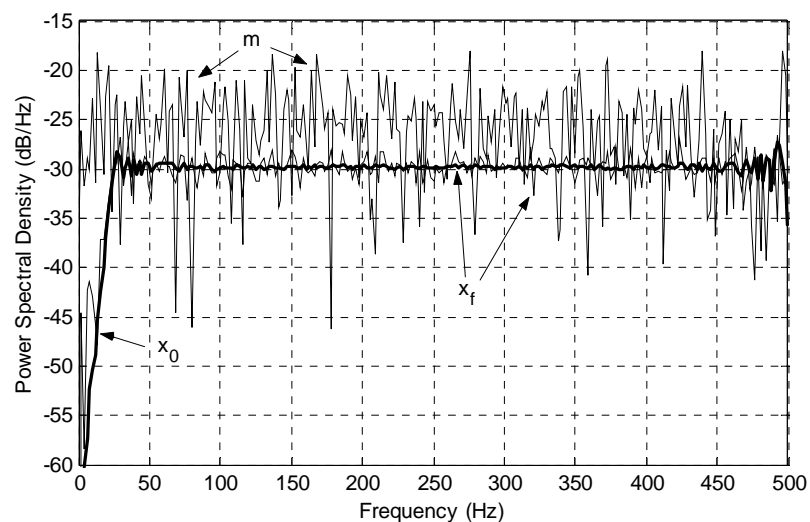


Рис.3. Спектральная плотность мощности сигналов в зависимости от частоты (периодограмма)

Второй этап. Короткий синусоидальный импульс, модулированный функцией Гаусса (частота дискретизации – 44100 Гц, амплитуда – 1, основная частота – 50 Гц, полоса пропускания – 50%, усечение происходит ниже 40 дБ), загрязнен тем же типом шума. После обучения фильтра и фильтрации сигнала, амплитуда шума уменьшилась на ≈ 10 дБ, что видно на рис.4 и 5.

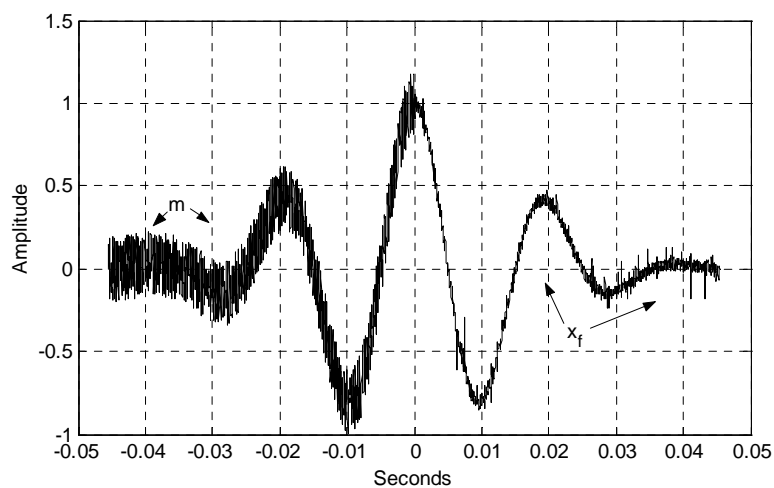


Рис.4. Сигналы: входной (m) $[-0.05..0]$ с, выходной (x_f) $[0..+0.05]$ с

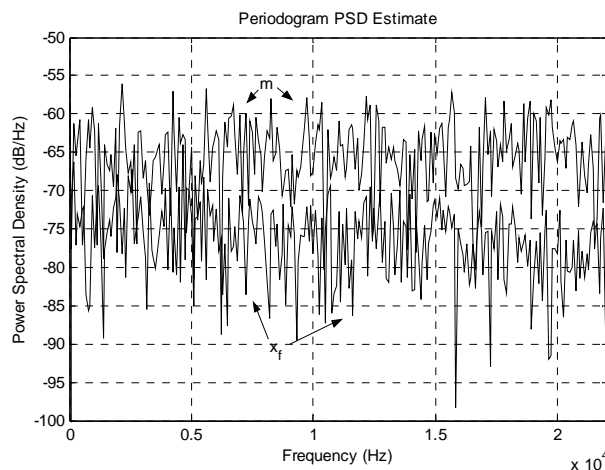


Рис.5. Спектральная плотность мощности сигналов в зависимости от частоты (периодограмма)

Заключение. Фильтр действительно является адаптивным, поскольку обучается на каждом подаваемом сигнале. Для обоих случаев тестового сигнала видно, что фильтр значительно приближает сигнал к оригиналу, а также существенно уменьшает разброс значений. Эффективность фильтрации зависит от качества обучения, которое напрямую связано с вычислительными затратами. Процедура обучения занимает 96% общего времени, т.е. обучение требует на порядок больше времени, чем режим исполнения (фильтрации). Всего на обработку фрагмента размером 6000 отсчетов тратилось 12 ± 2 секунд процессорного времени на компьютере с процессором Intel Pentium II и частотой 233 МГц. Обучение на длинных фрагментах нерационально, поскольку при увеличении наборов данных для обучения затраты времени вырастают до неприемлемых значений. Лучше обучать на небольшом фрагменте (до 5..10 секунд), а затем проводить нечеткие вычисления на полноразмерном фрагменте (любой длины – вплоть до нескольких часов). Однако, установление приемлемого компромисса между качеством обучения и количеством затраченного времени остается прерогативой исследователя. Низкая производительность обусловлена применением интерпретируемых программ MATLAB. При использовании оптимизированных библиотек (например, Intel Signal Processing Library и Recognition Primitives Library) или решений на основе DSP возможна обработка в реальном времени.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Дюбуа Д., Прад А. Теория возможностей. Приложения к представлению знаний в информатике: Пер. с франц. М.: Радио и связь, 1990.
2. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений: Пер. с англ. М.: Мир, 1976.
3. Короткий С.Г. Нейронные сети: алгоритм обратного распространения // ВУТЕ Россия. 2000. №21. С.26-29.
4. Тарасов В.Б. Синергетические проблемы в искусственном интеллекте // Труды Международной научно-практической конференции «Знание – диалог – решение». Т.2 (KDS-2001, Санкт-Петербург, 19-22 июня 2001 г.). СПб: Лань, 2001. С.594-602.
5. Ярушклина Н.Г. Нечеткие нейронные сети. Ч.1 // Новости ИИ. 2001. №2-3. С.47-51.