

## ПРИМЕНЕНИЕ КОСИНУС-ПРЕОБРАЗОВАНИЯ ФУРЬЕ В ТЕХНОЛОГИИ РЕЧЕВОЙ ПОДПИСИ

В настоящее время активно развивается направление распознавания голоса. Для систем контроля доступа предлагается метод голосовой аутентификации. Основная задача состоит в преобразовании голоса как функции амплитуды от времени в функцию частоты, что значительно облегчает аутентификацию. В данной работе изложены результаты исследований применения косинус-преобразования Фурье.

### 1 Понятие косинус-преобразования Фурье.

Пусть функция  $f(x)$  определена на всей числовой оси, удовлетворяет условиям Дирихле (условия, при которых эта функция может быть разложена в ряд Фурье) на любом конечном интервале и абсолютно интегрируема по всей оси (т.е. сходится несобственный интеграл  $\int_{-\infty}^{+\infty} |f(x)| dx$ ). Тогда справедливо следующее представление функции  $f(x)$  интегралом Фурье [1]

$$f(x) = \frac{1}{\pi} \int_0^{+\infty} \left\{ \int_{-\infty}^{+\infty} f(t) \cos[\omega(t-x)] dt \right\} d\omega .$$

Если  $f(x)$  – четная функция, то ее представление интегралом Фурье может быть переписано в виде

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} \left[ \sqrt{\frac{2}{\pi}} \int_0^{+\infty} f(t) \cos \omega t dt \right] \cos \omega x d\omega .$$

Полагая

$$F(\omega) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} f(t) \cos \omega t dt , \quad (1)$$

получим

$$f(x) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} F(\omega) \cos \omega x d\omega .$$

Функция  $F(\omega)$  называется косинус-преобразованием Фурье функции  $f(t)$ . Приведенная пара формул устанавливает закон взаимности: если  $F(\omega)$  – косинус-преобразование Фурье четной функции  $f(x)$ , то  $f(x)$  есть косинус-преобразование Фурье функции  $F(\omega)$ .

Косинус-преобразование Фурье можно использовать как инструмент спектрального анализа непериодических сигналов, т.е. это преобразование ставит в соответствие функции  $f(t)$  времени функцию  $F(\omega)$  частоты.

**ТАКОЕ ПРЕОБРАЗОВАНИЕ ЯВЛЯЕТСЯ ВЗАИМНО ОДНОЗНАЧНЫМ, ПОЭТОМУ ПРЕДСТАВЛЕНИЕ СИГНАЛА В ЧАСТОТНОЙ ОБЛАСТИ СОДЕРЖИТ РОВНО СТОЛЬКО ЖЕ ИНФОРМАЦИИ, СКОЛЬКО И ИСХОДНЫЙ СИГНАЛ, ЗАДАННЫЙ ВО ВРЕМЕННОЙ ОБЛАСТИ.**

## 2 Поиск частотной характеристики.

Пусть функция  $f(t)$  задана на отрезке  $[0; T]$  в виде совокупности из  $N_t$  точек (звуковых отсчетов или сэмплов), полученных после прочтения WAV-файла (положим  $\Delta t$  – отрезок между двумя соседними точками;  $T = (N_t - 1)\Delta t$ ). В каждом таком отрезке построим на левом значении функции прямоугольник. В интервале  $(T; +\infty)$   $f(t) = 0$ . Четным образом продолжим эту функцию на отрицательную полуось. Тогда имеет место преобразование (1):

$$F(\omega) = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} f(t) \cos \omega t dt = \sqrt{\frac{2}{\pi}} \left( \int_0^T f(t) \cos \omega t dt + \int_T^{+\infty} f(t) \cos \omega t dt \right) = \sqrt{\frac{2}{\pi}} \int_0^T f(t) \cos \omega t dt = \Delta t \sqrt{\frac{2}{\pi}} \sum_{i=1}^{N_t-1} f_i \cos(\omega(i-1)\Delta t). \quad (2)$$

Отведем в памяти одномерный массив  $zf$  вещественных чисел размером  $N_\omega = 7000$ . Выбирать это значение необходимо исходя из того, что говорящие люди имеют голоса с основными частотами  $\nu < 1000$  Гц;  $\omega = 2\pi\nu$  – циклическая частота.

Элемент массива  $zf$  с индексом  $j$  как следствие из выражения (2) с учетом нормировки будет рассчитываться следующим образом:

$$zf_j = \frac{\sum_{i=1}^{N_t-1} f_i \cos(j(i-1)\Delta t)}{\max_j \left| \sum_{i=1}^{N_t-1} f_i \cos(j(i-1)\Delta t) \right|}, \quad j = \overline{0, N_\omega - 1}.$$

## 3 Алгоритм распознавания голоса:

- при обучении системы записываем голос одного лица, который проговаривает букву, преобразуя каждый раз функцию  $f(t)$  по Фурье;
- первый раз частотная характеристика голоса заносится в массив  $zf_0$ , потом – в массив  $zf$ ;
- каждый раз при обучении мы смотрим различие данных в двух массивах, как  $\sum_j |zf_j - zf_0_j| = \varepsilon$ . Из всех  $\varepsilon$  выбираем максимальное значение  $\varepsilon_{\max}$  и принимаем его за интервал (коридор) для данного голоса;
- при входе в систему программа смотрит различие голосов  $\varepsilon_{\max}$  и если  $\varepsilon \leq \varepsilon_{\max}$ , то пропускает входящее лицо, иначе – нет.

1 Полянин А.Д. Краткий справочник для инженеров и студентов. М.: Международная программа образования, 1996. 432 с.

**Кафедра "Системы автоматизированного проектирования"**