УДК 681.3.06

Т. В. Грищук, к. т. н.

ПОЛУЧЕНИЕ ХАРАКТЕРИСТИЧЕСКОЙ ОБСЕРВАЦИИ СКРЫТОЙ МАРКОВСКОЙ МОДЕЛИ

В статье рассмотрен процесс получения на основе натренированной скрытой марковской модели характеристической обсервации голосовой команды. Данная информация может использоваться для повышения эффективности процесса распознавания речи. Введено понятие оценочной функции, нахождение максимума которой дает характеристическую обсервацию. Даны рекомендации по выбору начального приближения.

Ключевые слова: *характеристическая обсервация, скрытая марковская модель, распознавание речи, оценочная функция.*

В связи с быстрым развитием компьютерной техники все большую популярность завоевывают голосовые интерфейсы, которые позволяют сделать процесс взаимодействия с компьютером более естественным для пользователя и значительно увеличить круг пользователей программного обеспечения с возможностью голосового общения. Необходимо отметить, что подобное программное обеспечение из разряда прикладного постепенно переходит в разряд системного программного обеспечения, основными требованиями к которому являются высокое быстродействие и небольшой объем занимаемой памяти [1].

Основной задачей, которая решается в системах голосового общения, является задача определения вероятности принадлежности входного сигнала одной из команд голосового интерфейса. На рис. 1 показана упрощенная схема автоматического классификатора.

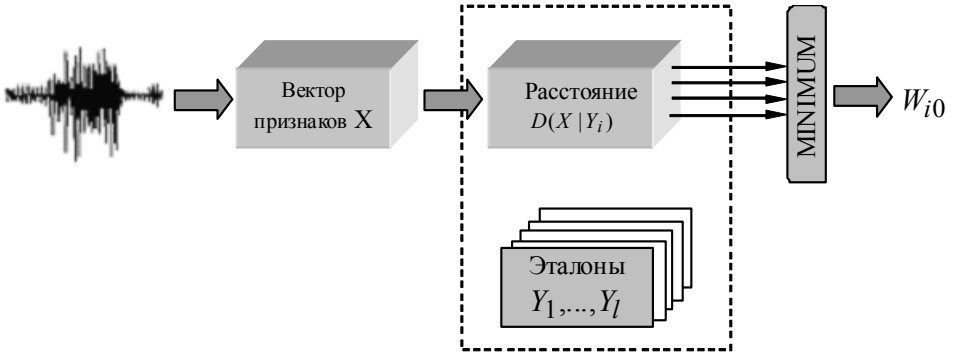


Рис. 1. Структура автоматического классификатора

Классическим и общепринятым подходом к классификации голосовых команд является подход, основывающийся на использовании математического аппарата скрытых марковских моделей (СММ). Этот подход предусматривает хранение в памяти не самих эталонов речевых сигналов, а наборов параметров скрытых марковских моделей, которые по своей сути являются «генераторами» векторов наблюдения сигнала (обсерваций). Неформально, обсервация – это последовательность векторов, координаты которых соответствуют временным и частотным признакам речевого сигнала в каждом окне анализа (приблизительно 256 отсчетов). Задача классификации сводится к задаче определения вероятности генерации входной обсервации моделями, которые соответствуют фразам входного командного языка [2, 3].

Первым этапом построения системы распознавания является этап обучения скрытых марковских моделей, то есть расчет параметров законов распределения в каждой вершине модели. В данной статье рассматривается обратный обучению процесс – получение на основе наученной модели эталонной (то есть характеристической) обсервации голосовой команды. Получение характеристической обсервации модели позволило бы решить не только задачу упрощения вычислительной сложности процесса классификации голосовых команд, но и предоставить дополнительную базу для решения задач адаптации и построения моделей с шумом.

Постановка задачи

В статье рассматривается задача обратная задаче построения СММ по алгоритму Баума – Велха [3].

> Дана марковская модель. Необходимо найти обсервацию, которая дает максимальную оценку в данной модели по Бауму – Велху.

Более неформально, пусть M – скрытая марковская модель. Необходимо найти обсервацию x', такую что:

$$x' = \arg\max_x P\left(x \middle| M\right). \tag{1}$$

Метод нахождения характеристической обсервации

Для начала введем необходимые обозначения и дадим определения основных понятий, которые будут использованы в этой статье.

Векторы обсерваций будем обозначать так:

$$\bar{x} = x_1, x_2, ..., x_n,$$

где n – количество координат (признаков) векторов обсерваций.

Определение 1 (Forward-вероятности):

$$\alpha_1(1) = 1$$

$$\alpha_j(1) = a_{1j}, \text{ для j = 2,...,N – 1}$$

$$\alpha_j(t) = \sum_{i=2}^{N-1} \alpha_i(t-1)\, G_i(x_{t-1}) \cdot a_{ij}, \text{ для j = 2,...,N – 1, t = 2,...,} T_\chi$$

$$\alpha_N(T_\chi) = \sum_{j=2}^{N-1} \alpha_j(T_\chi) G_j(x_{T_\chi}) \cdot a_{jN},$$

где a_{ij} – транзитная вероятность перехода из вершины i в вершину j;

N – количество вершин модели;

$G_j(\bar{x})$ – функция распределения, ассоциированная с эмиссионной вершиной j;

T_χ – математическое ожидание длин обсерваций, генерируемых моделью M:

$$T_\chi = \sum_{T=1}^{\infty} P(\chi(T) = N) \cdot T \text{ или } T_\chi = \sum_{T=1}^{\infty} \alpha_N(T) \cdot T.$$

Определение 2 (Backward-вероятности):

$$\beta_N(T_\chi) = 1$$

$$\beta_j(T_\chi) = a_{jN}, \text{ для j = 2, ..., N –1}$$

$$\beta_i(t) = \sum_{i=2}^{N-1} a_{ij} G_j(x_{t+1}) \beta_j(t+1), \text{ для i = 2, ..., N – 1, t = 1, ..., } T_\chi \text{-1,}$$

$$\beta_1(1) = \sum_{j=2}^{N-1} a_{1j} G_j(x_1) \cdot \beta_j(1).$$

Теперь мы имеем соотношения, которые определяют основную идею решения задачи:

$$P\left(x \middle| M\right) = \sum_{j=2}^{N-1} \alpha_j(t) G_j(x_t) \cdot \beta_j(t) \tag{2}$$

Соотношение (2) справедливо для любого $t = 1, \ldots, T_\chi$.

Таким образом, мы можем поэтапно улучшать оценку Баума – Велха путем нахождения максимума:

$$\arg\max_{x_t} P(x_t \mid M).$$

Для $t = 1$ получим: $P(x \mid M) = \sum_{j=2}^{N-1} \alpha_j(1) G_j(x_1) \cdot \beta_j(1)$, а для $t = T_\chi$ получим: $P(x \mid M) = \sum_{j=2}^{N-1} \alpha_j(T_\chi) G_j(T_\chi) \cdot a_{jN}$.

Для оценки качества обсервации $x = x_1, x_2, \ldots, x_{T\chi}$ примем:

$$e(j,t) = \alpha_j(t) \cdot \beta_j(t),$$

где j = 2, ... , N – 1, t = 1, ..., T_χ

Оценка Баума – Велха определяет наиболее вероятную вершину, в которой находится вектор обсервации в момент времени t [3]. На основе данной оценки ми введем оценочную функцию в виде суммы законов распределения всех эмиссионных вершин модели:

$$E(x_t,t) = e(2,t) G_2(x_t) + \ldots + e(j,t) G_j(x_t) + \ldots + e(N-1,t) G_{N-1}(x_t)$$

Таким образом, для любого t, $t = 1, 2, \ldots, T_\chi$ мы имеем оценочную функцию $E(x_t,t)$ для оптимизации оценки по x_t.

На верхнем уровне детализации алгоритм получения характеристической обсервации имеет вид:

```
Alg   Get Observation
      Get Init Observation          ‘получить характеристическое время Tχ,
                                    инициальную обсервацию x1,…, xTχ
                                    и ее оценку Pnew по Бауму–Велху
      Repeat
          P ⟵ Pnew
          Get Backward Probabilities      ‘ получить все βj (t)
          Get First Forward Probabilities ‘ получить αj (1) = a1j

          Loop  ‘ по t от 1 до Tχ
              Get Estimation Function     ‘получить E (xt,t)
              Get New Observation Vector  ‘получить x’t
              Get New Forward Probabilities ‘получить αj (t+1), если t < Tχ
                                            или αN (Tχ), если t = Tχ
          end
          Pnew ⟵ αN (Tχ)
      until | P – Pnew | < ε
```

Остановимся более подробно на процедурах, составляющих данный алгоритм.

Get Init Observation

Тут мы получаем характеристическое время, начальную обсервация и ее оценку по Бауму – Велху.

Get Backward Probabilities

Рассчитываем для данной модели величины $\beta_j(t)$.

Get First Forward Probabilities

Находим для всех эмиссионных вершин величину $\alpha_j(1) = a_{1j}$.

Get Estimation Function

В данной процедуре нужно получить оценочную функцию $E(x_t, t)$ для расчета вектора x_t. Эта оценочная функция задается при помощи коэффициентов:

$$\alpha_j(t) \cdot \beta_j(t) = e(j,t),$$

$$E(x_t, t) = \sum_{j=2}^{N-1} e(j,t)\, G_j(x_t).$$

Get New Observation Vector

Цель этой процедуры – нахождение x_t',

$$x_t' = \arg\max_{x_t} E(x_t, t).$$

Нахождение максимума такого типа функций мы рассмотрим ниже.

Get New Forward Probabilities

Эта процедура находит новые forward-вероятности с учетом того, что найден новый вектор обсервации x_t':

$$\alpha_j(t+1) = \sum_{i=2}^{N-1} \alpha_i(t)\, G_i(x_t') \cdot a_{ij}$$

для $j = 2, \ldots, N-1$, $t = 1, \ldots, T_\chi - 1$;

$$\alpha_N(T_\chi) = \sum_{i=2}^{N-1} \alpha_i(T_\chi)\, G_i(T_\chi) \cdot a_{iN}.$$

Отметим, что forward-вероятности на данном шаге цикла, который соответствует моменту времени t, рассчитываются только для момента времени $t + 1$ (для $t < T_\chi$).

Если шаг цикла соответствует моменту времени $t = T_\chi$, то рассчитывается только $\alpha_N(T_\chi)$.

Нахождение максимума функции при различных положениях составляющих

Для получения векторов обсерваций необходимо находить максимум функции вида:

$$F(x) = \gamma_2\, G_2(x) + \ldots + \gamma_j\, G_j(x) + \ldots + \gamma_{N-1}\, G_{N-1}(x),$$

где $G_j(x) = G_j(x_1,...,x_n)$ – функция распределения, ассоциированная с вершиной j;

$\gamma_j = \alpha_j(t) \cdot \beta_j(t)$.

В общем случае, нам необходимо найти максимум функции, которая является сумой нескольких нормальных законов распределения и имеет несколько локальных экстремумов. На рис. 2 показан пример такой функции.

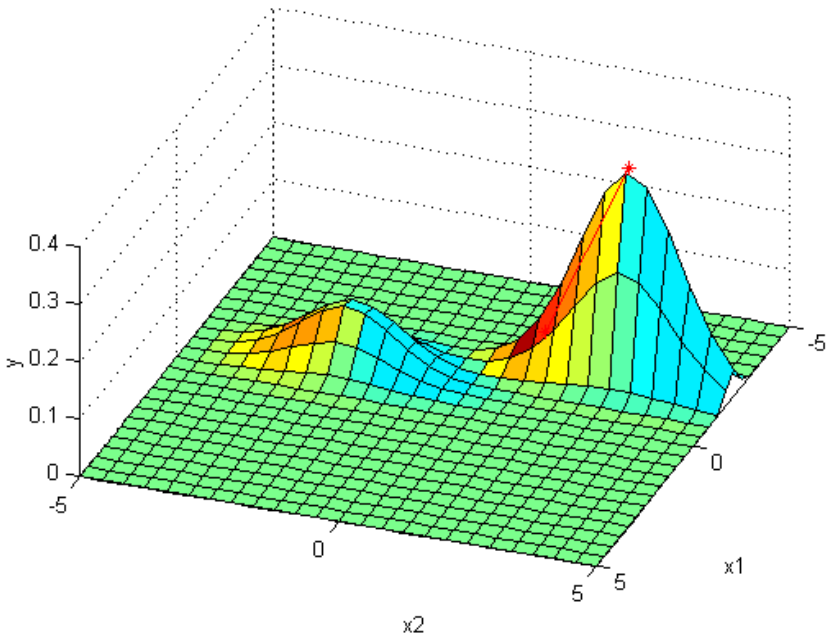


Рис. 2. Пример функции распределения модели для 2-х вершин и 2-х координат

Особенностью данной задачи является то, что глобальный максимум достигается не всегда. Поэтому в качестве начального приближения выберем обсервацию x^o:

$$x^o = \frac{1}{W}\left(W_2 \cdot c_2 + ... + W_{N-1} \cdot c_{N-1}\right),$$

где $W_j = F(c_j) = \sum_{i=2}^{N-1} \gamma_i G_i(c_{j1},...,c_{jn})$,

$W = \sum_{j=2}^{N-1} W_j$.

Рассмотрим систему уравнений:

$$\begin{cases} \dfrac{\partial E(x_1,...,x_n,t)}{\partial x_1} = 0 \\ - - - - - - - - - - - \\ \dfrac{\partial E(x_1,...,x_n,t)}{\partial x_j} = 0 \\ - - - - - - - - - - - \\ \dfrac{\partial E(x_1,...,x_n,t)}{\partial x_n} = 0 \end{cases}.$$

Выполнив необходимые преобразования, получим систему уравнений для нахождения вектора обсервации, который максимизирует оценочную функцию:

$$x_i = \frac{1}{\sum_{j=2}^{N-1} e_{ji}(x)} \cdot \sum_{j=2}^{N-1} e_{ji}(x) \cdot c_{ji} ,$$

где $i = 1, 2, ..., n$.

$$e_{ji}(x) = \frac{\gamma_j}{\sigma_{ji}^2} \cdot G_j(x),$$

σ_{ji} – среднеквадратическое отклонение для i-й координаты функции распределения $G_j(x)$, ассоциированной с вершиной j,

c_{ji} – математическое ожидание для i-й координаты функции распределения $G_j(x)$.

Решить данную систему уравнений можно любым итерационным методом численной математики.

Экспериментальные исследования

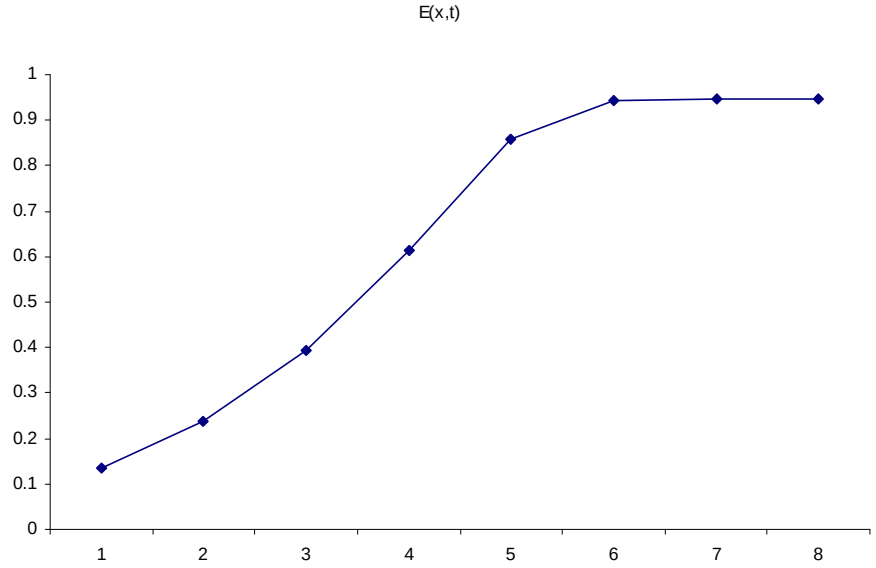
Автором проводились экспериментальные исследования разработанного алгоритма. В качестве иллюстрации приведем пример. В таблице представлены параметры скрытой марковской модели с 3-мя эмиссионными вершинами, для которой необходимо найти характеристическую обсервацию из 2-х координат длиной $T_\chi = 9$.

Таблица

Параметры СММ

Математическое ожидание	СКО	Транзитные вероятности
0 0	0 0	0 1 0 0 0
3.56 1.2	0.6275 0.2115	0 0.6 0.4 0 0
3.5 1.23	0.6169 0.2168	0 0 0.5 0.5 0
2.499 1.189	0.4405 0.2096	0 0 0 0.8 0.2
0 0	0 0	

На рис. 3 показана динамика роста оценочной функции и значения результирующей характеристической обсервации.



Результат

3.5600 1.2000
3.5137 1.2227
2.7548 1.2039
2.5408 1.1919
2.5055 1.1895
2.5000 1.1891
2.4991 1.1890
2.4990 1.1890
2.4990 1.1890

Рис. 3. Значения целевой функции на каждой итерации

Выводы

Основной научный результат, изложенный в данной статье, состоит в разработке метода нахождения характеристической обсервации скрытой марковской модели. Проведенные экспериментальные исследования подтвердили адекватность разработанного метода. Дальнейшие исследования в этой области могут касаться исследования возможности использования характеристических обсерваций для решения задач повышения быстродействия процесса распознавания, адаптации и создания моделей с шумом.

СПИСОК ЛИТЕРАТУРЫ

1. Борисов М. ПО для распознавания речи // Мир ПК. – 1999. – №3. – С. 124 – 129.

2. Normandin Y. Hidden Markov Models, Maximum Mutual Information Training, and the Speech Recognition Problem. – Montreal, Canada: Punlished by the Department of Electrical Engineering, McCall University, 1991. – 266 p.

3. Young S., Kershaw D., Odell J., Ollason D., Valthev V., Woodland P. The HTK Book (for HTK Version 3.0). – 2000. – 278 p.

Грищук Татьяна Викторовна – старший преподаватель кафедры компьютерных систем управления.

Винницкий национальный технический университет