

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СТРУКТУР СКРЫТЫХ МАРКОВСКИХ МОДЕЛЕЙ, ИСПОЛЬЗУЕМЫХ В ЗАДАЧЕ УСТАНОВЛЕНИЯ ЛИЧНОСТИ ЧЕЛОВЕКА ПО ИЗОБРАЖЕНИЮ ЛИЦА

Двойной И.Р.¹, Сальников И.И.²

¹ООО «Новая система», Пенза, Россия (440039, Пенза, ул. Гагарина, 16), e-mail: ilyadvoinoi@gmail.com

²ГОУ ВПО «Пензенский государственный технологический университет», Пенза, Россия (440039, Пенза, проезд Байдукова/ул. Гагарина, 1а/11), e-mail: iis@pgta.ru

С целью выявления структуры скрытой марковской модели, дающей наилучший результат в задачах идентификации человека по изображению лица, проведен сравнительный анализ. В статье приводятся общие сведения о скрытых марковских моделях, дается описание характерных областей изображения лица человека, приводятся алгоритмы их описания с помощью лево-правой и псевдодвухмерной структуры. Для оценки качества распознавания было использовано четыре различных набора тестовых изображений, бесплатно распространяемых в сети Internet. В результате исследования было установлено, что применение псевдодвухмерной модели дает незначительное увеличение качества работы системы распознавания за счет более точного описания локальных деформаций и взаимного расположения отдельных участков изображения. Дополнительным фактором, сдерживающим применение псевдодвухмерных моделей в реальных приложениях, является существенный рост вычислительной сложности, негативно влияющий на производительность системы распознавания.

Ключевые слова: изображение лица, скрытая марковская модель (СММ), идентификация.

COMPARATIVE ANALYSIS OF HIDDEN MARKOV MODELS STRUCTURES USED IN THE PROBLEM OF THE IDENTIFICATION HUMAN BY FACE IMAGE

Dvoynoy I.R.¹, Salnikov I.I.²

¹JSC New System, Penza, Russia, (440039, Penza, Gagarina St., 16), e-mail: ilyadvoinoi@gmail.com

²Penza State Technological University, Penza, Russia, (440039, Penza, Baydukova Drive/Gagarina St., 1a/11), e-mail: iis@pgta.ru

In order to identify the structure of a hidden Markov model, which gives the best result in identification of human faces by the image, a comparative analysis was conducted. This article provides an overview of hidden Markov models, a description of specific areas of human face's image. Algorithms, described in this article, show the usage of the left-right and pseudo-dual-in-line structure. To assess the quality of recognition, four different sets of test images (distributed free of charge in the Internet) have been used. The study found that the usage of pseudo-dual-in-line model gives a slight increase in the recognition quality of the system due to a more precise description of local deformations and mutual arrangement of individual parts of the image. A significant development of computational complexity impinging upon recognition system's productivity is an additional constraint, restraining the usage of pseudo-dual-in-line models in real applications.

Keywords: face image, hidden Markov model, identification.

Скрытая марковская модель (СММ) – модель, имитирующая работу реального процесса, позволяющая определить его неизвестные параметры на основе наблюдаемых. Теория СММ формирует основу индустриального стандарта для приложений осуществляющих распознавание речи [1].

Каждая СММ $\lambda=(A,B,\pi)$ представляет собой набор N состояний $S=\{S_1,S_2,\dots,S_N\}$ между которыми возможны переходы. В момент времени t система может находиться только в одном из состояний. При переходе из одного состояния в другое генерируется наблюдаемый символ, который соответствует физическому сигналу с выхода моделируемой системы. Набор символов для каждого состояния – $V=\{v_1,v_2,\dots,v_M\}$, количество символов – M . Символ

в состоянии $q_t=S_j$, в момент времени t генерируется с вероятностью $b_{jk}=P(v_k/q_t=S_j)$. Набор всех таких вероятностей составляет матрицу $B=\{b_{jk}\}$.

Матрица $A=//a_{ij}//$ определяет вероятность перехода из одного состояния в другое: $a_{ij}=P(q_{i+1}=S_j/q_i=S_i)$, $1 \leq i, j \leq N$. Вероятность нахождения системы в начальный момент времени описывается как $\pi=\{\pi_i\}$, где $\pi_i=P(q_1=S_i)$ [1].

В общем случае каждое состояние S_j модели λ может быть достигнуто за один шаг из любого другого состояния S_i . Такая модель называется эргодической. Каждая переходная вероятность a_{ij} эргодической СММ строго положительна. На практике моделируемый сигнал может не отвечать данному требованию, в результате за годы использования СММ в системах распознавания, было разработано множество топологий и конфигураций. Выбор структуры СММ во многом определяется областью применения и предшествующим опытом разработчика. Применительно к задаче установления личности наиболее очевидным решением является выбор структуры СММ, отражающей физиологические особенности строения лица человека.

Глядя на изображение лица, можно выделить ряд верикальных областей: волосы, лоб, брови, глаза, нос, рот и подбородок, при этом области следуют строго друг за другом, пропуск или изменение порядка следования областей невозможен, если человек не обладает ярко выраженными физическими отклонениями. В роли начальной области всегда выступают «волосы», в роли конечной – «подбородок». На основании сказанного можно заключить, что наиболее подходящей для описания лица является лево-правая модель СММ (модель Бакиса) [2].

Последовательность состояний модели Бакиса обладает тем свойством, что с увеличением времени, индекс состояния не может уменьшаться, иными словами, состояния переходят из одного в другое только слева направо. Поскольку в лево-правой модели не разрешены переходы в состояния, индекс которых меньше индекса текущего состояния, основное свойство всех лево-правых СММ выражается через значения переходных вероятностей следующим образом:

$$a_{ij} = 0, \quad j < i$$

Кроме того, поскольку последовательность состояний должна начинаться в состоянии 1, а заканчиваться в состоянии N , начальное распределение вероятностей состояний обладает свойством:

$$\pi_i = \begin{cases} 0, & i \neq 1 \\ 1, & i = 1 \end{cases}$$

Для того чтобы избежать значительных скачков в индексах состояний при использовании лево-правых моделей на переходные вероятности, как правило, налагаются дополнительные ограничения:

$$a_{ij} = 0, \quad j > i + \Delta$$

Переходные вероятности для последнего состояния лево-правой модели определяются следующим образом:

$$\begin{aligned} a_{NN} &= 1 \\ a_{Ni} &= 0, \quad i < N \end{aligned}$$

Таким образом, лево-правая модель хорошо согласуется с особенностями изображения человеческого лица, получаемого с систем фото- и видеофиксации. Поскольку обработка изображений в системах распознавания производится слева-направо, сверху-вниз, гарантируется переход только между соседними состояниями.

На практике зачастую производят объединение области бровей с областью глаз, а области рта – с областью подбородка, это дает возможность получить начальные значения СММ путем разбиения изображения на пять равных частей [3, 4]. Недостатком лево-правой модели является то, что она не учитывает пространственных особенностей изображения, представляя его в виде линейной цепочки событий.

Псевдодвухмерные СММ имеют большой потенциал в области обработки двумерных массивов данных. Псевдодвухмерная СММ является обобщением СММ, где каждое состояние само по себе является одномерной СММ. Таким образом, псевдодвухмерная СММ состоит из набора супер состояний, каждое из которых содержит набор вложенных состояний. Такая СММ характеризуется следующим набором параметров:

- Количество суперсостояний N_0 . Множество супер состояний обозначается как $S_0 = \{S_{0,i} \mid 1 \leq i \leq N_0\}$.
- Начальное распределение вероятностей супер состояния $\Pi_0 = \{\pi_{0,i}\}$, где $\pi_{0,i}$ вероятность нахождения в супер состоянии i в начальный момент времени.
- Матрица распределения вероятности перехода между супер состояниями, $A_0 = \{a_{0,ij}\}$, где $a_{0,ij}$ вероятность перехода из супер состояния i в супер состояние j .

Матрица распределения вероятности появления состояния, $B^{(k)} = \{b_i^{(k)}(O_{t_0,t_1})\}$, где O_{t_0,t_1} – вектор наблюдений для строки t_0 и столбца t_1 . Функция распределения вероятности, может быть представлена как:

$$b_i^k(O_{t_0,t_1}) = \sum_{m=1}^{M_i^k} c_{im}^k N(O_{t_0,t_1}, \mu_{im}^k, U_{im}^k),$$

где $N(O_{t_0,t_1}, \mu_{im}^k, U_{im}^k)$ функция нормального распределения.

Параметрами вложенной СММ \mathcal{L} являются:

- Количество вложенных состояний к k -ом супер состоянию, $N_1^{(k)}$ и множество вложенных состояний, $S_1^{(k)} = \{S_{1,i}^{(k)}\}$.
- Начальное распределение вероятностей состояний, $\Pi_1^{(k)} = \{\pi_{1,i}^{(k)}\}$, где $\pi_{1,i}^{(k)}$ вероятность нахождения в состоянии i супер состояния k в начальный момент времени.
- Матрица вероятности переходов между состояниями, $A_1^{(k)} = \{a_{1,jk}^{(k)}\}$, описывающая вероятность перехода из состояния k в состояние j .

На рисунке 1 представлена описанная структура СММ, каждое состояние в которой само по себе является лево-правой моделью СММ [5].

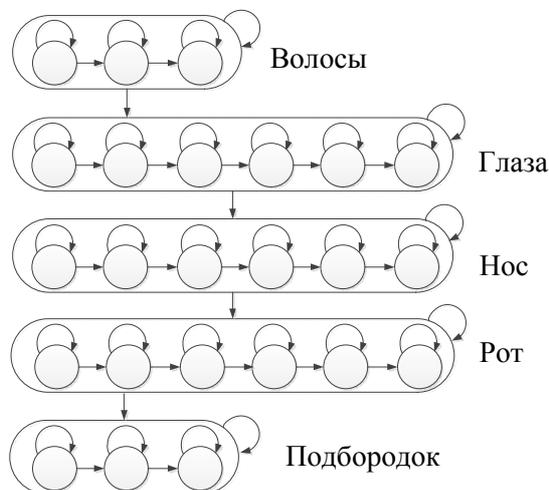


Рисунок 1 – Структура псевдодвухмерной СММ с пятью супер состояниями

Вероятность наблюдения одного из состояний вложенной СММ описывается с помощью распределения Гаусса:

$$b_i^{(k)}(O_{t_0,t_1}) = N(O_{t_0,t_1}, \mu_i^{(k)}, U_i^{(k)})$$

Псевдодвухмерная модель СММ имеет более сложную структуру по сравнению с лево-правой моделью, однако она в большей степени отражает характерные особенности изображения лица, учитывая не только его вертикальную, но и горизонтальную структуру.

В таблице 1 представлены результаты распознавания, полученные с использованием лево-правой и псевдодвухмерной структур, а также производительность каждого варианта системы распознавания.

Таблица 1 – Обобщенные результаты проведенного эксперимента

Тестовая база	Процент правильно распознанных изображений		Количество изображений, обработанных за 1 секунду	
	Лево-правая	Псевдо-двухмерная	Лево-правая модель	Псевдо-двухмерная

	модель	модель		модель
The ORL Database of Faces	96,5%	99,5%	3,92	3,46
Cohn-Kanade AU-Coded Facial Expression Database	91,75%	96,9	3,17	2,69
MIT-CBCL Face Recognition Database	93,8%	95,65%	3,78	3,33
Face Recognition Data, University of Essex, UK	89,76%	93,75%	3,63	3,15

Полученные данные позволяют сделать вывод о том, что использование псевдодвухмерной системы увеличивает количество правильно распознанных объектов в среднем на 4%. Это обусловлено тем, что такая модель позволяет учитывать локальные деформации и взаимное расположение участков изображения. Отрицательной стороной является значительный рост осуществляемых вычислений, что приводит к снижению производительности на 12,9%. Для систем, работающих в режиме реального времени, такой показатель может быть критичен.

Таким образом, применение псевдодвухмерной модели оправданно в том случае, если необходимо обеспечить повышенную точность распознавания, притом что время распознавания не является критичным. К таким задачам можно отнести аутентификацию пользователей персональных компьютеров и смартфонов, контроль доступа персонала к закрытым помещениям и т.д.

Список литературы

1. Рабинер Л.Р. Скрытые марковские модели и их применение в избранных приложениях при распознавании речи// ТИИЭР. – 1989. - №2. – С. 86-120.
2. Miar-Naimi H., Davari P. A new fast and efficient HMM-based face recognition system using a 7-state HMM along with SVD coefficients// Iranian Journal of Electrical & Engineering. – 2008. Vol. 4. – P. 46-57.
3. Nefian A.V., Hayes M. H. Face detection and recognition using hidden markov models// Proc. of image processing international conference. – 1998. Vol. 1. – P. 359-364.
4. Nefian A.V., Hayes M. H. Hidden markov models for face recognition// ICASSP98. – 1998. – P. 2721-2724.
5. Samaria F., Young S. HMM based architecture for face identification// Image and Computer Vision. – 1994. Vol. 12. – P. 537-583.

Рецензенты:

Бабич М.Ю., д.т.н, главный специалист ОАО «НПП «Рубин», г. Пенза.

Светлов А.В., д.т.н, профессор, заведующий кафедрой «Радиотехника и радиоэлектронные системы» Пензенского государственного технологического университета, г. Пенза.