

УДК 681

Н.А. Сергиевский.

Московский физико-технический институт (государственный университет)

Иерархический подход к распознаванию лиц, основанный на применении искусственных нейронных сетей

В настоящее время самым распространенным способом идентификации и верификации личности человека является распознавание лица полученного с видео или фотокамеры. В данной работе предлагается подход к распознаванию лиц, основанный на поэтапном иерархическом разбиении изображения лица на информативные области, формировании эталонов изображения в этих областях, и ассоциативном сравнении входного образа с эталонными образами.

Распознающая система может состоять только из нейронной сети [1], или поэтапно анализировать изображение, выполняя предобработку изображения, фильтрацию изображения, сокращение размерности изображения, классификацию изображения или вектора признаков полученного на предыдущем этапе.

Главной проблемой в распознавании изображений является выделение ключевых признаков и сохранение информации о топологии изображения. Использование нейронной сети при анализе изображений позволяет сохранить информацию о связях небольших фрагментов изображения на плоскости, и представить изображение в виде более короткого адресного кода, который можно упорядочить используя механизмы кластеризации.

Для иерархического разбиения изображения, формирования эталонов и сравнения могут быть использованы искусственные нейронные сети на основе нейроподобных элементов с временной суммацией сигналов [2]. Такой нейроподобный элемент выполняет свертку фрагмента бинарной последовательности длины n символов - $(x_{t-n+1}, x_{t-n+2}, \dots, x_t)$, $x_i \in \{0,1\}$, с последовательностью весовых коэффициентов (w_1, w_2, \dots, w_n) , $w_i \in \{-1, 1\}$:
$$S = \sum_{i=1}^n x_{t-n+i} w_i.$$

Свертка будет иметь наибольшее значение, если n -членный фрагмент входной последовательности соответствует последовательности (положительных) весовых коэффициентов нейрона, то есть если $w_i = -1$, то $x_i = 0$, а если $w_i = +1$, то $x_i = 1$. Такой фрагмент последовательности называется адресом нейрона. Наибольшее значение свертки равно числу единиц в адресе - sum .

В качестве нелинейной функции используется пороговое преобразование $f(*) = H_{\text{адр}}$ с порогом $h_{\text{адр}}$. Если порог преобразования $h_{\text{адр}}$ равен числу единиц в адресе - sum , то нейрон будет откликаться строго на свой адрес. То есть такой нейроподобный элемент выполняет ассоциативное сравнение.

Нейроподобный элемент, таким образом, моделирует одну из точек n -мерного сигнального пространства R^n . В случае бинарной входной последовательности - это

вершина n -мерного единичного гиперкуба G_n . Множество нейроподобных элементов с разными адресами моделирует весь гиперкуб.

Вся нейронная сеть выполняет преобразование последовательности в многомерное пространство $\hat{A} = F(A)$. Произвольная символьная последовательность A таким множеством нейроподобных элементов отображается в последовательность точек в многомерном пространстве (вершин гиперкуба) – траекторию \hat{A} . Этот тип нейронных сетей обладает рядом интересных, с точки зрения обработки информации, свойств. Входная информация может запоминаться, воспроизводиться и распознаваться (с помощью ассоциативного сравнения). Такие сети позволяют реализовать структурную обработку информации. С их помощью могут быть автоматически восстановлены словари элементов разного уровня, а также автоматически могут быть выявлены связи слов этих словарей во входной информации.

Нейронные сети этого типа можно эмулировать с помощью памяти, адресуемой по содержанию [2]. То есть этот тип обработки информации может быть реализован аппаратно (а значит быстро). Последнее очень важно при обработке больших массивов информации. В первую очередь это относится к обработке изображений. Рассмотрим алгоритмы иерархической обработки изображений лица, которые можно интерпретировать в терминах искусственной нейронной сети описанного типа, и которые в последствие могут быть реализованы аппаратно с использованием памяти, адресуемой по содержанию.

Общий механизм распознавания лица состоит в общем случае из двух этапов: (1) локализации или поиска элемента; и (2) его классификации.

Классификация при распознавании осуществляется на основе предварительной кластеризации множества эталонных элементов полученных на этапе обучения. Обучающая выборка формируется из набора элементов соответствующего уровня.

Рассмотрим применение иерархического механизма распознавания на примере распознавания глаз. Разделим изображение глаза на шесть элементов, равномерно разбив изображение на шесть частей. Эталонное изображение глаза формируется как последовательность окон $n \times n$ пикселей, расположенных на внутреннем контуре изображения. При поиске изображение сканируется окном $n \times n$ слева-направо и сверху-вниз со сдвигом на один пиксель. Сравнение при поиске заключается в вычислении разницы входного и эталонного изображений в пределах окна в дифференциальном представлении:

$$D = \sqrt{D_x^2 + D_y^2}, \text{ где } D_x = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}, D_y = D_x^T.$$

Мера используется стандартная евклидова.

Процесс поиска сводится к локализации эталонных участков на изображении другого глаза. После локализации соответствующих элементов на изображении другого

глаза, множество таких локализаций для всех глаз кластеризуется на группы исходя из минимизации суммарного отклонения локализации методом ближайшего соседа: $\text{If } P = \text{argmin}_p'(D(X, P)) \text{ then Class}(X) = \text{Class}(P)$, и на основе итеративного приближения нейроподобного элемента.

После формирования кластеров входное изображение анализируется с помощью характерных представителей классов, расстояние до которых у элементов анализируемого объекта будет достаточно мало. Операция классификации для каждого элемента дает вектор признаков, набор которых для всех элементов изображения глаза идет на следующий этап классификации более сложного объекта – всего лица.

Сравнение полученных элементов, так же происходит по евклидоподобной мере

$$d(y_1, y_2) = \sqrt{\sum_{i=1}^n \lambda_i (y_{1,i} - y_{2,i})^2}.$$

Заключение.

В работе представлен подход, основанный на классификации элементов глаза, что позволяет с достаточной точностью идентифицировать глаз, даже используя простые сопоставляющие и классифицирующие алгоритмы.

СПИСОК ЛИТЕРАТУРЫ

1. *Галушкин А.И.* Теория нейронных сетей. Серия «Нейрокомпьютеры и их применение», книга 1, - М., ИПРЖР, 2000г.
2. *Харламов А.А., Аллахвердов С.А.*, Искусственная нейронная сеть для структурного анализа информации на основе микросхем памяти, адресуемой по содержанию. // Информационные технологии, N 5, 2007г. С. 27 – 33.