

Инвариантное к ориентации и масштабу распознавание визуальных образов с использованием нечеткой нейросети

Станкевич Л. А., Хоа Н. Д.

stankevich_lev@inbox.ru, ndkhoa82@mail.ru

Санкт-Петербург, СПбГПУ

Проблема распознавания образов широко исследовалась в течение двух прошедших десятилетий. Основная цель исследований — разработка обучаемых систем распознавания образов, инвариантных к масштабу, положению и ориентации объектов.

В данной работе предлагается классифицирующая нейронная сеть SFAM (Simplified Fuzzy ARTMAP), которая является упрощенным вариантом сети FAM (Fuzzy ARTMAP) [1]. SFAM определяет принадлежность входных векторов к соответствующим классам, которым она может быть обучена. В качестве специального фильтра, используемого для предварительной обработки распознаваемых объектов, предлагается использовать What and Where (W&W) фильтр [2], с помощью которого можно определить положение, размер и ориентацию объекта в изображении.

Классифицирующая сеть SFAM

Сеть SFAM разработана специально для классификации путем удаления большой избыточности сети FAM. Структура SFAM показана на Рис. 1. Здесь нейрон P_j представляет подкласс, и его весовой вектор является прототипом подкласса. Узел C_k и веса w_{ij} являются меткой класса и весовым вектором нейрона в слое выходных подклассов P_j . Параметр $\rho \in [0; 1]$ — фактор вигильности.

В сети SFAM все непомеченные нейроны с единичными весами $w_{ij} = 1, \forall i$, названы некоммутированными. После того, как нейрон назначен в класс, он помечается и называется коммутированным. В режиме работы сети, если некоммутированный нейрон выигрывает конкуренцию со всеми коммутированными нейронами, то входной вектор относится к новому классу, см. Алгоритм 1.

W&W фильтр

В фильтре используются пространственные ориентированные рецептивные поля с различными размерами и ориентациями, которые свертываются с входным изображением. Рецептивное поле имеет ядро с ориентацией φ градусов, размером s пикселей и протяженностью a в точке (x, y) , функция ядра от r изображена на Рис. 2.

$$K(x, y, \varphi, s) = (1 - r^6) \exp(-r^4 / (1 + r^2)), \quad r^2 = (x'/a \cdot s)^2 + (y'/s)^2 \quad (1)$$

$$x' = x \cos \varphi + y \sin \varphi; \quad y' = y \cos \varphi - x \sin \varphi \quad (2)$$

Алгоритм 1. Обучение сети SFAM.

-
- 1: Определить дополнительный код из входного вектора a :
 $|a| := \sum_{i=1}^M a_i$;
 для всех $i = 1, \dots, M$: $a_i := a_i/|a|$; $a_i^c := 1 - a_i$;
 $A := (a, a^c)$;
 Вход A автоматически нормализуется, поскольку $\sum_{i=1}^{2M} A_i = M$.
 - 2: Вычислить активности нейронов и найти победительный нейрон J :
 для всех $j = 1, \dots, N$: $T_j := |A_j \wedge w_j|/(\alpha + |w_j|)$, где $\alpha > 0$;
 $J = \arg \max_j T_j$;
 - 3: **если** J — коммитированный нейрон **то**
 - 4: **если** класс J — обучаемый класс **то**
 - 5: $w_J := \beta(A \wedge w_J) + (1 - \beta)w_J$;
 - 6: **иначе**
 - 7: $T_J := 0$; новый победитель $J = \arg \max_j T_j$; повторить шаг 3;
 - 8: **иначе**
 - 9: $w_J := A$;
-

Входное изображение свертывается с каждым рецептивным полем. Активность узла A , расположенного в (x, y) , рецептивное поле которого имеет ориентацию φ и размер s , определяется дискретной сверткой с входом $I(x, y)$. Свертка между входным изображением и фильтрами различных ориентаций и размеров образует четырехмерную матрицу нейронных узлов, которая зависит от x , y , φ и s . Активность каждого из этих узлов дает степень соответствия между фигурой объекта в изображении и рецептивным полем. Позиция, ориентация и масштаб объекта может быть оценена нахождением максимальной активности.

Когда размер объекта увеличивается, ядра фильтров, определенные в (1)–(2), дают неправильные оценки позиции, ориентации и размера объекта. Чтобы устранить эту неправильность, используется нормализация весов фильтра. При этом активность выходного узла нормированного фильтра определяется как

$$A(\varphi, s) = \sum_x \sum_y K_N(x, y, \varphi, s) I_C(x, y),$$

где $K_N(x, y, \varphi, s)$ — нормированная по площади рецептивного поля функция ядра фильтра, а $I_C(x, y)$ — центрированное изображение объекта.

Путем конкуренции выбирается узел, который имеет максимальную активность, и определяются параметры φ_I и s_I . В итоге, после преобразования изображения W&W фильтром, получаем изображение объекта с нормированным размером и горизонтальной ориентацией.

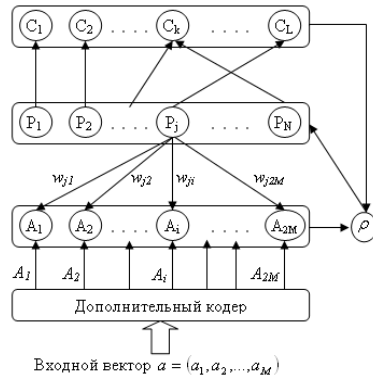
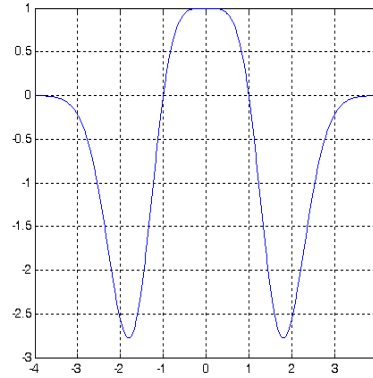


Рис. 1. Структура SFAM

Рис. 2. Функция ядра фильтра в зависимости от t

Эксперимент

В эксперименте использовались бинарные изображения самолета размером 128×128 пикселей. После W&W фильтра получалось изображение с размером 32×32 пикселей, и подавалось в векторном виде на вход сети SFAM. Распознавалось 10000 изображений 10 типов боевых самолетов: по 1000 изображений каждого типа с ориентацией в диапазоне $[0^\circ; 2^\circ]$ и масштабом в диапазоне $[0, 5; 2]$. Обучающая выборка содержала 180 изображений: по 18 изображений для каждого типа самолета с масштабами от 0,5 до 2 по отношению к их оригинальным изображениям. Обучение заняло 0,6 секунды. Результаты эксперимента распознавания объектов с нормализованным и ненормализованным фильтрами и с различным количеством грубых фильтров в банках для простых и зашумленных изображений, показана в таблице.

Число фильтров	6	9	12
С нормализацией	79, 10%	91, 20%	96, 60%
Без нормализацией	25, 25%	27, 40%	31, 00%
С нормализацией (с шумом)	33, 70%	42, 30%	71, 60%
Без нормализацией (с шумом)	14, 60%	18, 55%	24, 95%

Эксперименты показали, что сеть SFAM и нормализованный фильтр дают лучшие результаты распознавания, поскольку точнее определяется размер и ориентации объекта.

Литература

- [1] *Carpenter G. A., Grossber S.* Fuzzy Artmap : A neural network architecture for incremental supervised learning of analog multidimensional Maps // IEEE Transactions on Neural Network. — 1992. — V. 3. — Pp. 698–712.
- [2] *Carpenter G. A., Grossber S. and Leshert G. W.* What-and-Where filter. A partial mapping neural network for object recognition and image understanding // Computervision and image understanding. — 1998. — V. 69, No. 1. — Pp. 1–22.