

РАЗРАБОТКА МЕТОДА И АЛГОРИТМА РАСПОЗНАВАНИЯ ДВУХМЕРНЫХ КОНТРАСТНЫХ ИЗОБРАЖЕНИЙ ОБЪЕКТОВ ПО ИНВАРИАНТНЫМ ИНФОРМАТИВНЫМ ПРИЗНАКАМ

Близкая О.В., группа КСД-00а

Руководитель проф. Скобцов Ю.А.

Постоянный интерес к достижениям в области распознавания видеоизображений обусловлен стремлением исключить участие человека из процесса принятия решения о типе или свойствах объекта, находящегося в поле зрения телевизионной системы, и тем самым повысить надёжность функционирования систем диагностики, контроля, управления, охраны. Актуальность этой задачи делает её популярной среди исследователей, работающих в различных областях математики, средств вычислительной техники и программирования. В результате сведения об имеющихся достижениях в области распознавания изображений оказываются весьма разрозненными, а подходы к решению проблемы отличаются многообразием не только используемых методов распознавания (наиболее распространёнными можно считать экстремально-корреляционные, статистические, структурно-лингвистические методы, а также метод геометрических инвариантов), но и критериев оценок достигаемого качества распознавания.

В настоящей статье предпринята попытка проведения комплексного анализа свойств системы распознавания видеоизображений объектов, основанная на принципах функционирования нейронных систем (НС). Исходная информация для распознавания содержится в функции распределения интегральной интенсивности излучения $I(x, y)$ по поверхности фотодетектора, с которой связана двумерная система координат XOY . После детектирования видеообраз объекта будет представлять собой дискретную функцию $g(x_i, y_i)$ с шагом дискретизации $\Delta x = \Delta y$. Для обеспечения инвариантности распознавания к изменению $I(x, y)$ в пределах объекта

проводится квантование по уровням 0 и 1 функции $g(x_i, y_i)$ путём сравнения $g(x_i, y_i)$ с наперёд заданным порогом. Эта операция уменьшает избыточность информации и как следствие увеличивает устойчивость процедуры распознавания, которую теперь можно рассматривать как распознавание плоских геометрических фигур.

Осуществление эффективного распознавания произвольно расположенных в плоскости двумерных фигур требует решения следующих задач:

- ❖ формирования множества информативных признаков для распознавания, которые должны быть инвариантными как к смещению фигур в плоскости, так и к их поворотам;
- ❖ определения параметров НС, параметров обучающего алгоритма;
- ❖ проведения корректного обучения НС.

Инвариантность информативных признаков к смещению объектов в пределах поля зрения телевизионной системы можно обеспечить в результате выполнения функционального преобразования, обладающего свойствами непрерывности и взаимной однозначности. Такими свойствами обладает преобразование Фурье. Для случая непрерывной двумерной функции $g(x, y)$ оно имеет вид:

$$F\{g(x, y)\} = G(f_x, f_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \exp(-j2\pi(f_x x + f_y y)) dx dy, \quad (1)$$

где $G(f_x, f_y)$ – Фурье-образ исходной функции $g(x, y)$.

Из свойств Фурье-преобразования известно, что если:

$$F\{g(x, y)\} = G(f_x, f_y), \text{ т } F\{g(x-a, y-b)\} = G(f_x, f_y) \exp\{j2\pi(af_x + bf_y)\}, \quad (2)$$

т.е. смещение функции $g(x, y)$ относительно начала координат приводит к возникновению линейного фазового сдвига в области частот (f_x, f_y) . Амплитудный спектр $|G(f_x, f_y)|$ не изменяется - он инвариантен к сдвигу.

Амплитудный спектр не инвариантен к повороту функции $g(x, y)$ относительно начала координат: Фурье-образ функции вида:

$$g_\alpha(x, y) = g(x \cos \alpha + y \sin \alpha, -x \sin \alpha + y \cos \alpha) \quad (3)$$

получим поворотом Фурье-образа $G(f_x, f_y)$ функции $g(x, y)$ на угол α в плоскости частот.

Для обеспечения инвариантности признаков распознавания к повороту $g(x, y)$ предложено проводить дополнительное функциональное преобразование амплитудного спектра $|G(f_x, f_y)|$, которое заключается в его интегрировании по окружностям с радиусами $0 < \rho < \infty$. Используя полярные координаты в плоскости частот, получаем функцию вида

$$S(\rho) = \int_0^{2\pi} |G(\rho \cos \varphi, \rho \sin \varphi)| d\varphi, \quad (4)$$

которая инвариантна как к смещению, так и к повороту $g(x, y)$ в системе координат XOY .

Перейдя к дискретному представлению $g(x, y)$ с помощью преобразования Фурье, получим дискретную функцию признаков $S(p_i)$. Для этого по отсчетам быстрого преобразования Фурье $|G(f_{xi}, f_{yi})|$, функции $g(x_i, y_i)$ найдем дискретный амплитудный спектр, а затем вычислим суммы его отсчетов на окружностях с радиусами $p_i = 0, 1, 2, \dots, L$ с центром в точке $(0, 0)$ пространства частот (f_{xi}, f_{yi}) , причем $L = x_{\max} = y_{\max}$.

Полученные таким образом для всех типов распознаваемых объектов трижды инвариантные L -мерные векторы признаков предъявляются на вход системы распознавания для ее обучения.

Основным элементом для построения НС служит искусственный нейрон. Он традиционно представляется линейным сумматором с n входами (каждому из входов ставится в соответствие некоторый весовой коэффициент w_i) и одним выходом, соединенным с нелинейным элементом, реализующим активационную функцию нейрона $F(x)$. В качестве активационной функции используется, как правило, функция гиперболического тангенса или сигмоидальная функция вида:

$$F(x) = 1/(1 + \exp(-x)) \quad (5)$$

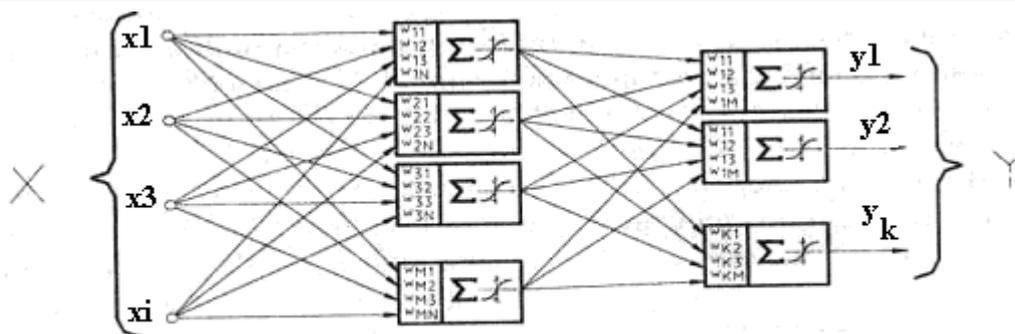


Рисунок 1 - Двухслойная нейронная сеть

Двухслойная нейронная сеть прямого распространения, имеющая N входов и K выходов (рис. 1), содержит два слоя нейронов - скрытый, состоящий из M нейронов, и выходной, состоящий из K нейронов. Результатом работы такой нейронной сети является нелинейное преобразование N -мерного входного вектора-строки X в K -мерный выходной вектор-строку Y .

$$Y = F((F(X * W)) * U), \quad (6)$$

где W - матрица весовых коэффициентов нейронов внутреннего слоя НС размерностью $[N * M]$, U - матрица весовых коэффициентов нейронов выходного слоя размерностью $[M * K]$; $F(x_j)$ - активационная функция.

Уравнение (6) представляет собой матричную форму записи системы из K нелинейных уравнений. Путем подбора соответствующих элементов матриц весовых коэффициентов W и U можно добиться примерного равенства каждого выходного вектора Y_i нейронной сети одному из наперед заданных векторов Y_i , каждый из которых, в свою очередь, ставится в соответствие одному или нескольким входным векторам X_j . В результате реализуется разделение множества входных векторов $\{X\}$ на несколько подмножеств в N -мерном пространстве признаков некоторой гиперповерхностью, неявным образом задаваемой совокупностью коэффициентов W и U . Размерность входного вектора пространства признаков X существенным образом влияет как на качество распознавания объектов, так и на его устойчивость. Обучение НС по интегральным признакам вида (1) изображений объектов позволяет обеспечить инвариантность распознавания к сдвигу и повороту наблюдаемой фигуры, а использование подмножества наиболее информативных признаков, устойчивых к шуму, - повысить эффективность распознавания. При

определении параметров НС важно правильно выбрать размерность M выходного вектора. Как правило, эта величина выбирается в соответствии с общими представлениями о виде получаемого решения - часто M принимается равным числу возможных классов распознаваемых объектов, предопределяя тем самым близость выходных векторов к ортам M -мерного ортогонального базиса.

Для распознавания фигур используется двухслойная НС, обученная при помощи процедуры обратного распространения. Итерации обучения НС повторяются до тех пор, пока для всех эталонных образов не будут достигнуты требуемые значения сигналов на выходах НС с ошибкой, меньшей некоторой наперед заданной ошибки обучения λ . Поскольку вероятность правильного распознавания неявным образом зависит как от значения λ , так и от числа нейронов в каждом слое НС и способов кодирования выходной информации НС, необходимо определить не только параметры самой НС, но и найти оптимальное значение λ , обеспечивающее максимальное значение правильного распознавания.

Эффективность работы системы распознавания с выбранными параметрами оценивается в результате цикла статистических испытаний, по результатам которых принимается решение о необходимости корректирования параметров системы.

Перечень ссылок

1. Балухто А.Н., Булаев В.И. Нейрокомпьютеры в системах обработки изображений. – М.: «Радиотехника» – Книга 7. – 2003 – 192 с.
2. Буянов АА, Власов А.И., Макеев С.С. Исследование нейросетевых алгоритмов, применяемых для распознавания образов// Вторая Международная конференция СНГ. – М, окт. 2000. – ч.2,3. – с. 22-25.
3. Уоссерман Ф. «Нейрокомпьютерная техника: теория и практика». –М.- 1992. – 132 с.
4. Белявцев В.Г., Воскобойников Ю.Е. «Алгоритмы фильтрации изображений с адаптацией размеров апертуры» // Автометрия. – 1998. – № 3. – С. 18 – 25.
5. «Адаптивные методы обработки изображений». – Сборник научных трудов. – под. ред. В.И. Сифоров. – М: «Наука». – 1988. – 242 с.