

ПРИМЕНЕНИЕ ДВУМЕРНЫХ СПЕКТРОВ ДЛЯ АВТОМАТИЧЕСКОГО
 РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Г.Я. Воложин, Н.Г. Загорушко

Решение проблемы автоматического распознавания образов существенно зависит от результатов поиска системы параметров, которая достаточно хорошо описывала бы любую реализацию данного образа и не реагировала бы на различного рода случайные воздействия, такие, например, как аддитивная помеха, перенос начала и поворот осей координат пространства наблюдений, изменение масштаба объекта (изображения) и др.

В этом плане полезным оказывается применение спектрально-го описания реализаций распознаваемых образов. Если спектры широко используются для описания одномерных сигналов (меняющихся только во времени), то многомерные спектры не получили такого распространения. Вместе с тем известно [1], что спектральный подход к описанию функций 2-х и большего числа переменных весьма полезен, так как позволяет увеличить помехозащищенность линий связи, оптимизировать методы квантования функций нескольких переменных, в частности двумерных изображений в телевидении, фотографии, полиграфии.

Под многомерным спектром абсолютно интегрируемой функции $f(x_1, \dots, x_m)$ понимается комплексная в общем случае функция $S(\omega_1, \dots, \omega_m)$, связанная с $f(x_1, \dots, x_m)$ следующим соотношением:

$$S(\omega_1, \dots, \omega_m) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f(x_1, \dots, x_m) \exp\left\{-j \sum_{k=1}^m \omega_k x_k\right\} dx_1 \dots dx_m \quad (1)$$

Существует [2] и обратное преобразование (если $S(\omega_1, \dots, \omega_m)$ абсолютно интегрируема):

$$f(x_1, \dots, x_m) = \frac{1}{(2\pi)^m} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} S(\omega_1, \dots, \omega_m) \exp\left\{j \sum_{k=1}^m \omega_k x_k\right\} d\omega_1 \dots d\omega_m \quad (2)$$

Многомерные спектры обладают рядом свойств, которые оказываются весьма полезными при автоматическом распознавании образов [3].

Модуль спектральной функции

$$\varphi(\omega_1, \dots, \omega_m) = |S(\omega_1, \dots, \omega_m)|$$

не зависит от переноса начала координат (параллельного переноса функции $f(x_1, \dots, x_m)$ в пространстве наблюдений).

Спектральные параметры обладают определенной помехозащищенностью, так как в тех случаях, когда спектральные структуры полезной функции и помехи различны, можно с помощью m -мерного фильтра увеличить отношение сигнал/помеха.

Поворот исходной функции $f(x_1, \dots, x_m)$ вокруг произвольной точки приводит к повороту спектральной функции $S(\omega_1, \dots, \omega_m)$ вокруг начала координат с соответствующим изменением фаз составляющих (изменение фаз не оказывает влияния на $\varphi(\omega_1, \dots, \omega_m)$). Таким образом, при необходимости можно устранить влияния поворота координатных осей на систему спектральных параметров, проинтегрировав, например, модуль спектральной функции, записанной в полярных координатах по соответствующему углу в пределах от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$. Некоторые другие свойства многомерных спектров описаны в работе [3].

Многомерные обобщенные спектры (в том числе и (1)) позволяют осуществить единый подход к параметрическому описанию образов любой мерности. Затруднение состоит в том, что расчет многомерных спектров достаточно сложен и требует выполнения большого числа элементарных вычислительных операций. Действительно, при вычислении спектральной функции типа (1) двумерного изображения размером 30×30 без табулирования значений $\sin \alpha_x$ и $\cos \alpha_x$ необходимо выполнить около 200 арифметических действий на каждую точку спектральной функции (предполагается, что вычислительная машина трехадресная).

Существуют (см. [4]) физические аналоговые двумерные анализаторы спектра в виде оптических приборов. При определенных ограничениях, накладываемых на соотношение между размером объекта и длиной световой монохроматической волны, дифракционная картина объекта соответствует его двумерному спектру. Однако законы распространения световых лучей лишают дифракционную картину некоторых полезных свойств, присущих двумерным спектрам. Например, дифракционная картина зависит от точности юстировки оптической системы, сдвиг изображения приводит к сдвигу дифракционной картины, при повороте изображения координаты точки, вокруг которой поворачивается дифракционная картина, зависят от координат точки, вокруг которой поворачивается изображение.

Трудности, связанные с вычислением многомерных спектров, не позволяют использовать последние для практического создания устройств, распознающих, например, двумерные изображения. Однако исследовательские работы, направленные на поиск информативных многомерных спектральных параметров и упрощение методики их выделения, представляются, на наш взгляд, целесообразными.

В настоящей статье приводятся результаты некоторых экспериментов по автоматическому распознаванию простых геометрических фигур и рукописных цифр по их двумерным спектрам

$$\varphi(\omega_1, \omega_2) = |S(\omega_1, \omega_2)|.$$

Три геометрические фигуры (круг, треугольник и прямоугольник) были заданы на поле размером 36×36 . Изображения фигур были закодированы и перенесены на перфокарты. Затем поле с изображениями засеивалось случайной равномерно распределенной помехой интенсивностью 5%, 10%, 15% и т.д. до 40%. Проценты означают относительное число точек исходного поля, пораженных помехой. Поражение точки помехой состоит в том, что если в ней был "0", то он заменяется на "1", и наоборот. Одновременно изображения подвергались параллельному сдвигу на исходном поле. На рис. 1 представлены изображения треугольника, круга и прямоугольника, засеянных 5%-й и 35%-й помехой.

Комплексная спектральная функция достаточно достоверно может быть рассчитана внутри прямоугольника, ограниченного прямыми, параллельными осям координат плоскости (ω_1, ω_2) и проходящими через точки $+\omega_{1max}, +\omega_{2max}, -\omega_{1max}, -\omega_{2max}$, где ω_{1max} и ω_{2max} - частоты, равные $\frac{1}{2}$ частоты квантования

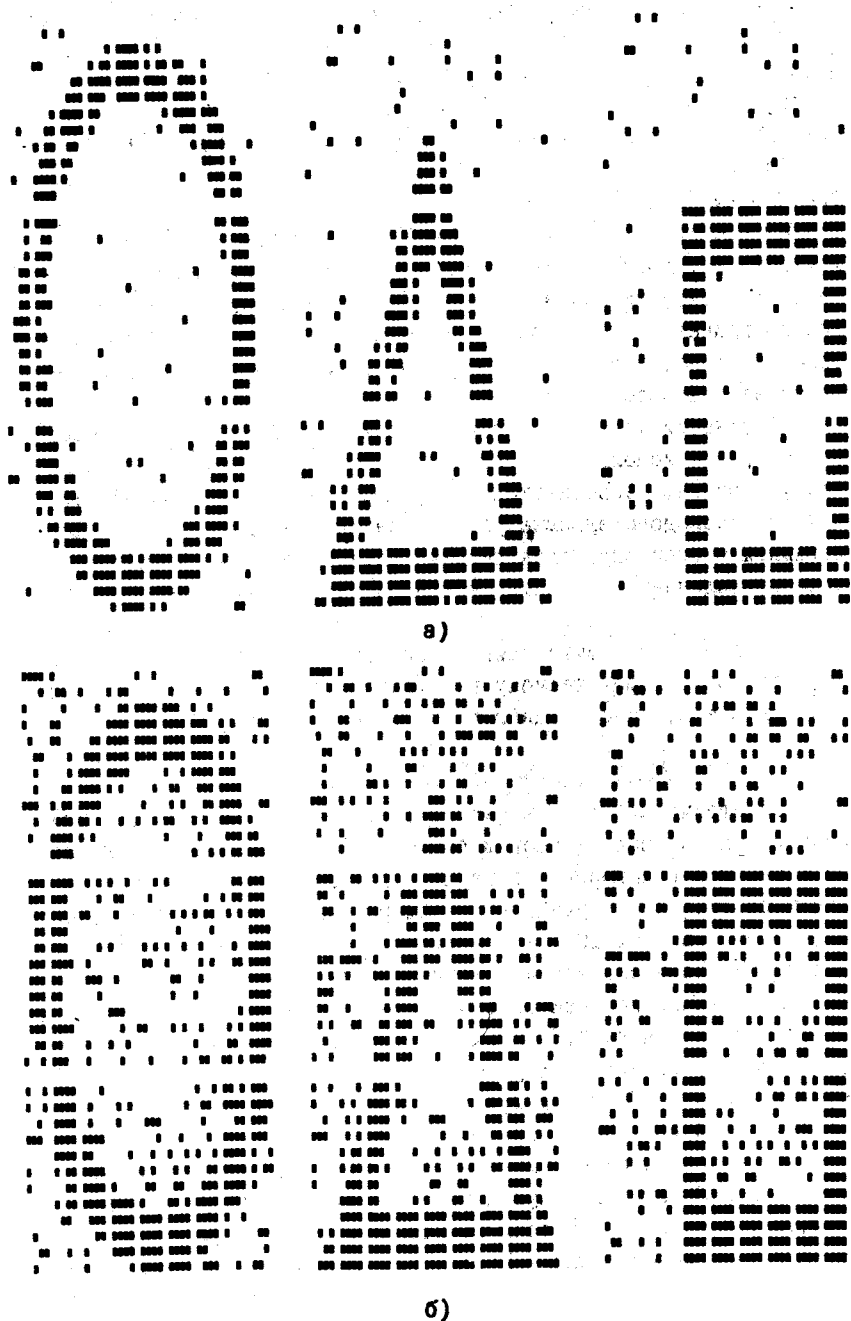


Рис. 1. Изображения фигур с помехами различного уровня: а) уровень помех 5%; б) уровень помех 35%.

исходной функции по соответствующей оси (см. [1]). Для того, чтобы найти модуль спектральной функции, достаточно рассчитать $S(\omega_1, \omega_2)$ лишь на половине указанного прямоугольника.

Действительно, если через начало координат плоскости частот провести прямую линию (см. рис. 2а), то для любой точки A расчетного прямоугольника, лежащей по одну сторону прямой, имеется центросимметричная точка A^* , причем $S(A)$ и $S(A^*)$ комплексно сопряжены. Поэтому для расчета $\varphi(\omega_1, \omega_2)$ достаточно знать значения $S(\omega_1, \omega_2)$ лишь по одну сторону прямой.

В данной работе спектр рассчитывался в прямоугольнике со сторонами, параллельными координатным осям и проходящими через точки O , ω_{1max} , ω_{2max} и $-\omega_{1max}$ (рис. 2б).

Вычисление производилось по приближенной формуле:

$$\begin{aligned} \varphi(\omega_1, \omega_2) \approx & \left[\sum_{i=1}^{36} \cos \omega_1 x_1^{(i)} \cdot \sum_{j=1}^{36} f(x_1^{(i)}, x_2^{(j)}) \cos \omega_2 x_2^{(j)} - \right. \\ & \left. - \sum_{i=1}^{36} \sin \omega_1 x_1^{(i)} \cdot \sum_{j=1}^{36} f(x_1^{(i)}, x_2^{(j)}) \sin \omega_2 x_2^{(j)} \right]^2 + \\ & + \left[\sum_{i=1}^{36} \cos \omega_1 x_1^{(i)} \cdot \sum_{j=1}^{36} f(x_1^{(i)}, x_2^{(j)}) \sin \omega_2 x_2^{(j)} + \right. \\ & \left. + \sum_{i=1}^{36} \sin \omega_1 x_1^{(i)} \cdot \sum_{j=1}^{36} f(x_1^{(i)}, x_2^{(j)}) \cos \omega_2 x_2^{(j)} \right]^2 \cdot \Delta x_1 \Delta x_2, \quad (3) \end{aligned}$$

где $f(x_1, x_2)$ - анализируемое изображение;
 $\Delta x_1, \Delta x_2$ - шаг квантования по осям x_1 и x_2 ; соответственно.

Были рассчитаны спектры 27 изображений (3 образа, каждый из которых представлен 9-ю реализациями: от чистой до замурманенной 40%-й помехой). На основе полученных результатов были построены эталоны образов. Распознавание осуществлялось по минимуму евклидова расстояния от распознаваемой реализации до эталонов. Распознаваемые реализации отличались друг от друга положением на исходном поле и мощностью помехи. В результате правильно были опознаны все реализации, засаженные помехой мощностью от 5% до 35% (см. рис. 1). Этот эксперимент иллюстриру-

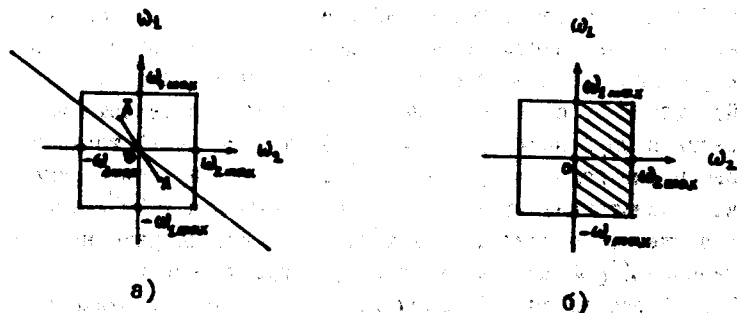


Рис. 2. Изображение области определения двумерной спектральной функции: а) общий случай; б) область, использованная в описываемом эксперименте.

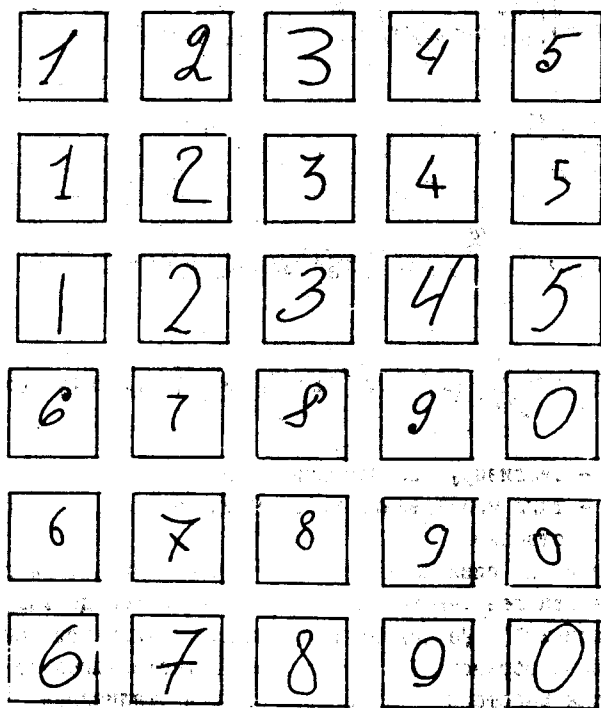


Рис. 3. Изображения рукописных цифр, использованные в эксперименте.

ет помехозащищенность многомерных спектральных параметров.

Эксперимент, который заключался в опознавании рукописных цифр, осуществлялся следующим образом. Каждая из 10 арабских цифр была написана 12-ю разными людьми, причем написание цифр практически не накладывалось никаких ограничений. Задан был лишь прямоугольник размером 25x34, внутри которого должна уместиться цифра. *) Примеры различных написаний цифр приведены на рис. 3. Предполагалось, что контурные изображения цифр свободны от помех.

При расчете $\Phi(\omega_1, \omega_2)$ осуществлялось нормирование по размеру изображения. Известно (см. [3]), что если функция

$$f(x_1, x_2) \text{ имеет спектр } S(\omega_1, \omega_2),$$

то функция $f(\eta_1, \eta_2)$, связанная преобразованием подобия с $f(x_1, x_2)$ и равная

$$f(ax_1, ax_2),$$

где a - коэффициент изменения масштаба ($a > 1$ соответствует уменьшению, $a < 1$ - увеличению),

имеет спектр

$$\frac{1}{a^2} S\left(\frac{\omega_1}{a}, \frac{\omega_2}{a}\right).$$

При строгом преобразовании подобия увеличение контурного изображения должно сопровождаться утолщением линий. Практически этого нет. Поэтому использовать свойства преобразования подобия можно лишь с определенной степенью приближения.

Размер рукописной цифры определялся по площади поля (S_n), занятой изображением (а не по $\sqrt{S_n}$, как следовало бы делать при преобразовании подобия). Затем вычислялся относительный размер изображения и вносились соответствующие поправки в численную формулу расчета спектра (3).

Описанный метод определения размера изображения можно применять лишь при отсутствии помех.

В условиях зашумленности изображения необходимо предварительно его отфильтровать.

*) Исходный материал собран и подготовлен для ввода в ЭВМ Силаевой Г.Б.

Поле частот было представлено $27 \times 10 = 270$ числами

$$\omega_{1 \max} = 9 \Delta \omega_1, \omega_{2 \max} = 13 \Delta \omega_2, \Delta \omega_i = \frac{2\pi}{T_{i \max}}$$

Были проведены следующие эксперименты по обучению и опознаванию.

I. Случайным образом были выбраны 6 реализаций каждого образа (всего 60 реализаций). На основе этой обучающей последовательности были построены эталоны 10 образов. Затем для проверки успешности обучения было проведено опознавание обучающей последовательности. Опознавание производилось (как и во всех других опытах) по минимуму евклидова расстояния до эталонов образов.

Из 60 реализаций обучающей последовательности неправильно классифицированы 2, что соответствует надежности опознавания 96,7%.

После этого для распознавания была предъявлена контрольная последовательность изображений, не участвовавших в обучении. Из 60 реализаций (6 реализаций каждого образа) были неправильно опознаны 24, что соответствует надежности распознавания 60%. Такое снижение надежности по сравнению с опознаванием обучающей последовательности объясняется, по всей вероятности, недостаточной представительностью последней. Можно ожидать, что увеличение числа обучающих реализаций приведет к увеличению надежности опознавания контрольной последовательности. Это иллюстрируется сравнением I-го эксперимента со следующим:

II. Для построения эталонов были использованы по 9 случайно выбранных реализаций каждого образа. При опознавании обучающей последовательности из 90 реализаций неправильно были классифицированы 11, что соответствует надежности распознавания 88,9%.

В контрольной последовательности из 30 реализаций (3 реализации каждого образа) неправильно опознаны 10, что соответствует надежности 66,7%.

В заключение следует отметить, что полученные результаты являются предварительными, поскольку основываются на недостаточно большом статистическом материале. По этой же причине не изучен вопрос о построении оптимального решающего правила для опознавания рукописных цифр по их двумерным спектрам.

Вместе с тем достигнутая в процессе экспериментов надежность опознавания (при условии, что на способ написания

цифр практически не накладывалось ограничений) позволяет надеяться, что многомерные спектральные параметры окажутся весьма полезными для автоматического распознавания образов.

Л И Т Е Р А Т У Р А

1. Игнатьев Н.К. Теория дискретизации и ее применение к задачам связи. М., 1962 г., МЭИС (докторская диссертация).
2. Привалов И.И. Ряды Фурье. ОНТИ. ГТТИ. М-Л., 1934 г.
3. Волошин Г.Я. Некоторые свойства многомерных спектров (данный сборник). Стр. 13 - 20.
4. Рытов С.М. О методе фазового контраста в микроскопии. Успехи физических наук, 1950 г., 41 т, № 4.

Поступила в редакцию
25.VI.1965