

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

Государственное образовательное учреждение
высшего профессионального образования

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ

ЦИФРОВАЯ ОБРАБОТКА И РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ

Методические указания
к выполнению лабораторных работ

Санкт-Петербург
2006

Составители: И. Л. Ерош, Н. В. Соловьев

Рецензент кандидат технических наук, доцент *О.И.Красильникова*

Содержатся методические указания по выполнению цикла лабораторных работ, предназначенных помочь студентам направления «Информатика и вычислительная техника» в изучении некоторых разделов следующих дисциплин: «Цифровая обработка и распознавание изображений», «Системы искусственного интеллекта».

Изложены методы коррекции пространственных искажений изображений, основанные на групповом представлении преобразований точек плоскости, а также детерминистские и вероятностные методы распознавания образов в пространстве признаков.

Редактор *А. Г. Ларионова*

Компьютерная верстка *А. Н. Колешко*

Сдано в набор 13.04.06. Подписано к печати 20.04.06. Формат 60×84 1/16. Бумага офсетная. Печать офсетная. Усл. печ. л. 1,3. Уч. -изд. л. 1,4. Тираж 100 экз. Заказ №

Отдел электронных публикаций и библиографии библиотеки
Отдел оперативной полиграфии
ГУАП

190000, Санкт-Петербург, ул. Б. Морская, 67

© ГОУ ВПО «Санкт-Петербургский
государственный университет
аэрокосмического приборостроения»,
2006

ПРЕДИСЛОВИЕ

В методических указаниях представлены четыре учебно-исследовательские лабораторные работы, оформленные в виде отдельных программ.

Работа № 1 знакомит студентов с описанием изменения изображения объекта при его движении в пространстве аффинными и проективными группами преобразований точек плоскости. Студент изучает особенности центрального и параллельного проецирования, а также исследует связь между параметрами вращения способом проецирования и коэффициентами группового преобразования.

Работа № 2 знакомит студентов с методами коррекции пространственных искажений изображений, описываемых групповыми преобразованиями. Студент изучает особенности коррекции проективных искажений на стандартном носителе и исследует связи между размерами носителя, параметрами искажения и ошибкой коррекции.

Работа № 3 знакомит студентов с методами распознавания растровых изображений при наличии яркостных искажений на примере печатных символов. Студент изучает подход к распознаванию образов по расстоянию в пространстве признаков, возможные системы признаков для конкретного набора изображений, а также исследует зависимость качества распознавания от выбранной системы признаков, степени близости образов и уровня яркостных искажений.

Работа № 4 знакомит студентов с методами распознавания образов при их статистическом описании. Студент изучает задание многомерных статистических описаний классов векторами математического ожидания и дисперсионными матрицами признаков, сокращение размерности пространства признаков с помощью критерия Фишера, распознавание по расстоянию Махаланобиса, а также моделирует наилучшее с точки зрения качества распознавания разделение пространства признаков на области, содержащие образы одного класса.

Все работы снабжены набором контрольных вопросов, десять из которых случайным образом выбираются для тестирования.

ЛАБОРАТОРНАЯ РАБОТА № 1

Изменение проекции объекта при его вращении в пространстве

Цель работы: Исследование изменения коэффициентов группы преобразования изображения плоского объекта в процессе его вращении вокруг трех неподвижных осей декартовой системы координат при центральном и параллельном проецировании.

Теоретические пояснения

Изображение объекта получается в результате его проецирования оптической системой на плоскость проецирования. Изменение положения объекта в пространстве, т. е. его вращение вокруг трех неподвижных осей декартовой системы координат и смещение вдоль этих осей, приводит к изменениям изображения объекта. Связь между точками изображений объекта в исходном и смещенном положениях описывается уравнениями группового преобразования. Изменение изображения объекта при центральном проецировании описывается проективной группой преобразования (рис. 1)

$$\begin{cases} x' = \frac{a_1 x + a_2 y + b_1}{a_5 x + a_6 y + 1} \\ y' = \frac{a_3 x + a_4 y + b_2}{a_5 x + a_6 y + 1} \end{cases}, \quad (1)$$

где x, y, x', y' – исходные и преобразованные проективной группой координаты точки изображения, $a_1, \dots, a_6, b_1, b_2$ – параметры или коэффициенты проективной группы.

Изменение изображения объекта в случае параллельного проецирования описывается аффинной группой преобразования (рис. 1)

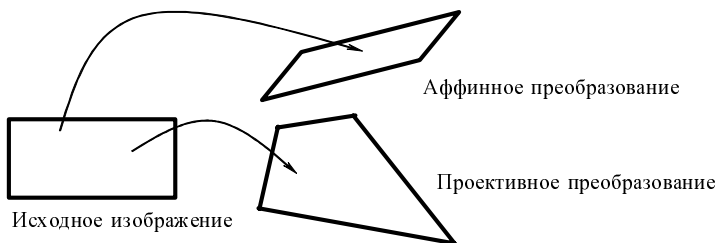


Рис. 1. Групповые преобразования изображения

$$\begin{cases} x' = a_1x + a_2y + b_1 \\ y' = a_3x + a_4y + b_2 \end{cases}, \quad (2)$$

где x, y, x', y' – исходные и преобразованные аффинной группой координаты точки изображения, $a_1, \dots, a_4, b_1, b_2$ – коэффициенты аффинной группы.

Из уравнений (1) и (2) видно, что аффинная группа является подгруппой проективной группы при $a_5 = 1, a_6 = 1$. Конкретные значения коэффициентов a_1, \dots, a_6 определяются величиной, последовательностью и направлением вращения объекта вокруг осей. Коэффициенты b_1, b_2 определяются только поступательным движением объекта, т. е. при вращении объекта без смещения $b_1 = 0, b_2 = 0$.

При центральном проецировании (рис. 2, а) координаты (x, y) точки на изображении, соответствующей точки объекта с координатами (x_0, y_0, z_0) , определяются из уравнений проецирования

$$\frac{x}{x_0} = \frac{f}{z_0}; \quad \frac{y}{y_0} = \frac{f}{z_0}, \quad (3)$$

где f – фокусное расстояние объектива оптической системы. При параллельном проецировании (рис. 2, б) координаты (x, y) точки изображения равны координатам (x_0, y_0) соответствующей точки объекта, т. е. принимается $f/z_0 = 1$.

Из уравнений (3) видно, что отношение f/z_0 определяет степень проективных искажений изображения. Последнее означает, что проекции точек, имеющих одинаковые координаты x_0 или y_0 , но разные координаты z_0 , т. е. находящиеся на разном расстоянии от плоскости проеци-

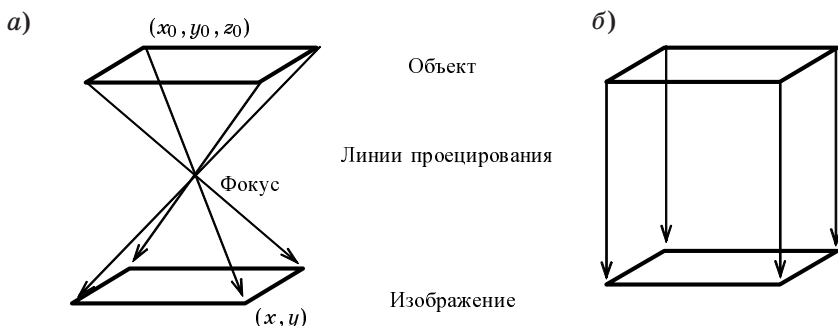


Рис. 2. Проецирование объекта на плоскость: а – центральное; б – параллельное

рования, при движении объекта в пространстве будут изменяться по-разному. Строго говоря, все оптические системы обеспечивают только центральное проецирование, но если изменение координаты z_0 точки объекта при его движении в пространстве незначительно по сравнению с расстоянием от этой точки до плоскости проецирования, то центральное проецирование можно свести к параллельному. Более того, если отношение f/z_0 постоянно для всех точек объекта при его движении в пространстве, то проецирование тоже сводится к параллельному, а отношение f/z_0 определяет коэффициент масштабирования изображения.

Координаты точки объекта в пространстве при его вращении определяются из уравнения

$$\mathbf{X} = \mathbf{A}\mathbf{X}_0, \quad (4)$$

где $\mathbf{X}_0 = (x_0, y_0, z_0)^T$, $\mathbf{X} = (x, y, z)^T$ – векторы координат точки объекта в начальном положении и после поворота объекта соответственно; \mathbf{A} – полная матрица поворота объекта.

Полная матрица поворота определяется как последовательное произведение матриц поворота вокруг осей X , Y и Z , т. е. $\mathbf{A} = \mathbf{A}_x \mathbf{A}_y \mathbf{A}_z$. Причем, от последовательности задания осей зависят значения элементов полной матрицы. Элементами матрицы поворота вокруг оси являются синусы и косинусы угла поворота вокруг данной оси. Например, матрица поворота вокруг оси Z имеет вид

$$\mathbf{A}_z = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

где γ – угол поворота вокруг оси Z в радианах. Аналогично определяют и матрицы поворота вокруг других осей.

Используя уравнения (1), (4) и соотношение (3) при центральном проецировании или уравнения (2) и (4) при параллельном проецировании, можно получить уравнения, связывающие параметры группового преобразования и величины, определяющие изменения положения объекта в пространстве. По известным значениям вектора \mathbf{X}_0 , углам поворота объекта, фокусному расстоянию f и координатам характерных точек на изображении объекта в исходном положении можно определить параметры группового преобразования. Обратное утверждение в общем случае неверно. Действительно, вращение объекта вокруг оси X или Y приводит к изменению его изображения, которое можно описать сжатием изображения вдоль соответствующей оси. Параметр сжатия связан с углом поворота, но направление вращения невозможно определить только из анализа изображений объекта в исходном и изме-

ненном положении. Для определения направления поворота необходима дополнительная информация о расстоянии до одной из точек объекта.

В данной работе объект представляет собой многоугольник, плоскость которого в исходном положении параллельна плоскости проецирования. Вершины многоугольника задаются в декартовой системе координат (начало – в центре экрана дисплея, ось X – горизонтальна, ось Y – вертикальна, ось Z направлена перпендикулярно экрану, представляющему собой плоскость проецирования).

Вращение производится вокруг осей, проходящих через начало координат. Если центр формы объекта не совпадает с началом координат, то в процессе вращения объекта его изображение будет не только изменять свою форму, но и смещаться в поле зрения. Последнее не приводит к появлению отличных от нуля значений b_1, b_2 , так как вычисление коэффициентов группы преобразования производится в системе координат, связанной с центром экрана, а не с центром формы изображения объекта. Если в процессе вращения объекта его изображение вырождается в линию, то вычисление параметров преобразования становится невозможным и дальнейшие вычисления прекращаются. В связи с этим, при центральном проецировании не рекомендуется задавать в качестве шага угла поворота вокруг осей X или Y делители 90° (например, 10, 9, 5, 2 и т. п.).

Порядок выполнения работы

1. Изобразить на дисплее некоторую фигуру в виде ломаной линии.
2. Записать координаты вершин рисунка и зарисовать исходное изображение.
3. Выбрать последовательность осей вращения, число шагов и величину шага для выбранных поворотов, способ проецирования, фокусное расстояние и расстояние от фокуса до плоскости объекта для центрального проецирования. Полный поворот вокруг одной оси должен быть около 360° , а шаг на должен превышать 20° .
4. Выполнить вращение объекта при заданных параметрах. Проследить за тем, чтобы точки изображения не выходили за пределы плоскости проецирования.
5. Зарисовать графики изменения параметров группы, записать максимальные и минимальные значения параметров группы.
6. Повторить пп. 3 – 5 для других последовательностей осей вращения и способов проецирования. Общее число экспериментов – не менее четырех.
7. Получить положительную оценку за ответы на тестовые вопросы.

Пример задания параметров вращения и проецирования

Эксперимент 1: Последовательность осей вращения – XYZ , шаг для каждой оси – 14° , число шагов – 26, проецирование – параллельное.

Эксперимент 2: Последовательность осей вращения – YXZ , шаг для каждой оси – 14° , число шагов – 26, проецирование – параллельное.

Эксперимент 3: Последовательность осей вращения – XYZ , шаг для каждой оси – 14° , число шагов – 26, проецирование – центральное, расстояние до объекта – 200, фокусное расстояние – 100.

Эксперимент 4: Последовательность осей вращения – XYZ , шаг для каждой оси – 14° , число шагов – 26, проецирование – центральное, расстояние до объекта – 1000, фокусное расстояние – 100.

Выводы: сравнить результаты экспериментов 1 и 2, в которых изменен порядок осей вращения; сравнить результаты экспериментов 1 и 3, в которых выбраны разные системы проецирования; сравнить результаты экспериментов 2 и 3, в которых выбраны разные расстояния до объекта.

Содержание отчета

1. Рисунок и координаты исходного изображения объекта.
2. Выбранные параметры вращения и проецирования, графики и экстремальные значения параметров группы для каждого эксперимента.
3. Выводы о соотношениях коэффициентов группового преобразования и параметрах вращения объекта, а также об особенностях проецирования плоских объектов.

Контрольные вопросы

1. Как получается изображение объекта?
2. Как связаны координаты точек изображений объекта в исходном и повернутом положениях?
3. Чем отличается параллельное проецирование от центрального?
4. При каких условиях центральное проецирование переходит в параллельное?
5. Как определяется полная матрица поворота объекта?
6. Как выбор последовательности осей вращения влияет на конечное положение объекта?

ЛАБОРАТОРНАЯ РАБОТА № 2

Коррекция проективных искажений на стандартном носителе

Цель работы: Исследование методов коррекции пространственных искажений изображений, представленных в растровом виде, получение общего представления об аффинных и проективных группах преобразований точек плоскости.

Теоретические пояснения

При построении компьютерных систем, использующих различные визуальные устройства для получения растровых изображений с последующим распознаванием объектов, представленных на этих изображениях, часто приходится компенсировать пространственные искажения, возникающие при вводе изображения. Примерами таких систем являются:

- визуальные терминалы для ввода графической и текстовой информации при свободной ориентации носителя;
- интеллектуальные роботы с техническим зрением промышленного и исследовательского назначения;
- системы обработки аэрокосмических снимков для поиска полезных ископаемых, оценки природных ресурсов, экологического мониторинга.

Для построения визуальных устройств в таких системах используются разнообразные датчики бесконтактного типа (сканеры, цифровые фото- и видеокамеры, широкополосные спектрометры и т. п.), позволяющие формировать в памяти компьютера растровое изображение в виде последовательности значений функции яркости на плоскости.

В результате изменения в пространстве положения объекта или точки съемки изображение объекта тоже изменяется. Это изменение принято называть пространственным искажением изображения объекта относительно изображения этого же объекта в некотором эталонном положении. Известно, что такие искажения могут быть эффективно описаны с использованием аппарата теории непрерывных групп, в частности проективными (1) и аффинными (2) группами Ли на плоскости, которые рассматриваются в лабораторной работе № 1.

В матричной форме аффинное преобразование можно представить как

$$\mathbf{X}' = \mathbf{A} \mathbf{X} + \mathbf{B}, \quad (5)$$

где \mathbf{X} , \mathbf{X}' – соответственно векторы-столбцы координат точки исходного и искаженного изображения; \mathbf{A} , \mathbf{B} – матричное представление элемента g аффинной группы G , т. е.

$$\mathbf{A} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Аффинная группа G может быть разложена на подгруппы G_1, G_2, \dots, G_n , т. е. элемент группы $g \in G$ может быть представлен как произведение элементов ее подгрупп, т. е. $g = g_1 g_2 \dots g_n$. Наиболее часто используются следующие подгруппы аффинной группы:

– центроаффинная ($\mathbf{A}_{ca} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$, $b_1 = 0, b_2 = 0$);

– вращение ($\mathbf{A}_\alpha = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$, где α – угол поворота вокруг оси,

перпендикулярной плоскости изображения);

– симметричное отражение ($\mathbf{A}_o = \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$);

– масштабирование ($\mathbf{A}_m = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}$, где k_1, k_2 – масштабные коэффициенты);

– подобие ($\mathbf{A}_p = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$, где k – коэффициент подобия);

– косой сдвиг по оси X ($\mathbf{A}_{sx} = \begin{pmatrix} 1 & k_x \\ 0 & 1 \end{pmatrix}$, где k_x – коэффициент сдвига);

– косой сдвиг по оси Y ($\mathbf{A}_{sy} = \begin{pmatrix} 1 & 0 \\ 1 & k_y \end{pmatrix}$, где k_y – коэффициент сдвига);

– смещение ($\mathbf{A}_s = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $b_1 = t_x, b_2 = t_y$, где t_x, t_y – смещение по осям

X и Y);

– а также их комбинации, например, смещение и масштабирование

$$(\mathbf{A}_m = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}, \quad b_1 = t_x, b_2 = t_y).$$

Компенсацией пространственного искажения изображения называется приведение искаженного изображения к исходному виду, т. е. вы-

полнение преобразования, обратного (5): $X = g^{-1}(X')$, где g^{-1} – элемент группы преобразования, обратный к g . Как правило, параметры элемента группы преобразования g и g^{-1} неизвестны. Однако могут быть известны исходное положение объекта и некоторые особенности изображения. Параметры группы преобразования справедливы для всех точек изображения, так как изображение преобразуется как единое целое. Следовательно, зная соответствие между несколькими точками исходного и искаженного изображения, можно вычислить параметры преобразования и по ним привести изображение к исходному виду. Из уравнений (1) и (2) видно, что для нахождения параметров аффинного преобразования необходимо найти соответствие между тремя точками на исходном и искаженном изображениях, а для нахождения параметров проективной группы – между четырьмя точками, так как по каждой точке формируются два уравнения, связывающие параметры группового преобразования и координаты соответствующих точек на исходном и искаженном изображениях. Как следствие, для нахождения параметров группового преобразования необходимо решить систему из шести линейных уравнений в случае аффинного преобразования или восьми – в случае проективного преобразования.

Заметим, что уравнения (1) нелинейны относительно координат точек, но линейны относительно групповых параметров.

В лабораторной работе рассматривается случай, когда заранее известно, что исходное изображение (график, схема, текст и т. п.) находится на плоском прямоугольном носителе высотой H (вертикальный размер) и шириной L (горизонтальный размер). Предполагается, что в результате работы визуальной системы получено изображение этого листа, искаженное поворотом и смещением листа в пространстве перед визуальной системой. Под исходным положением, к которому необходимо привести искаженное изображение, принимается такое расположение листа, при котором центр листа совпадает с оптической осью визуальной системы, а стороны листа параллельны строкам и столбцам системы дисплея.

В качестве характерных точек, по координатам которых можно определить параметры группового преобразования, целесообразно использовать вершины листа. Известные методы яркостной обработки изображений позволяют выделить на искаженном изображении вершины листа, определить их координаты и найти соответствие между вершинами исходного и искаженного листов. Последнее существенно, так как ошибка в соответствии вершин приводит к неправильному определению параметров группы и, как следствие, к неправильной компенсации искажения.

Порядок выполнения работы

1. Выбрать размеры листа и изобразить на нем некоторую фигуру в виде ломаной линии.
2. Выбрать несколько подгрупп аффинной группы, для каждой из них выбрать параметры, вычислить полную аффинную группу как произведение выбранных подгрупп.
3. Произвести искажение исходного изображения полученной в п. 2 группой.
4. Произвести восстановление изображения, указав правильное соответствие между вершинами исходного и искаженного изображений листа.
5. Повторить несколько раз п. 4, изменяя соответствие вершин и размеры листа.
6. Произвести искажение исходного изображения проективной группой, в которой параметры $a_1, \dots, a_4, b_1, b_2$ совпадают с параметрами аффинной группы, вычисленной в п. 2, а параметры a_5, a_6 подобраны, так, чтобы максимизировать искажения.
7. Повторить п. 4, обратив внимание на точность восстановления.
8. Подобрать такие параметры искажения проективной или аффинной группой, при которых ошибка восстановления имеет существенное значение, например, при значительном сжатии изображения.
9. Получить положительную оценку за ответы на тестовые вопросы.

Содержание отчета

1. Размеры листа, рисунок и координаты исходного изображения объекта.
2. Для каждой пары искажение – восстановление привести параметры искажения и восстановления, выбранные размеры листа и соответствие вершин при восстановлении, рисунки искаженного изображения и восстановленного изображения, если оно заметно отличается от исходного.
3. Выводы о влиянии соответствия вершин и задания размеров листа на результаты восстановления, о связи между степенью искажения и точностью восстановления и др.

Контрольные вопросы

1. Какие визуальные системы применяются для получения изображения?
2. Чем вызваны пространственные искажения изображения?

3. Какая особенность пространственных искажений позволяет использовать описанный метод компенсации?
4. На какие подгруппы может быть разложена аффинная группа?
5. Почему важно правильное соответствие между вершинами исходного и искаженного изображений?
6. Для какого количества точек нужно знать координаты при вычислении параметров аффинного и проективного преобразования?

ЛАБОРАТОРНАЯ РАБОТА № 3

Методы распознавания зашумленных символов

Цель работы: Исследование различающихся по числу и виду используемых признаков методов распознавания изображений символов, представленных в растровой форме, при различной степени их зашумления.

Теоретические пояснения

Обучить машину тому, что умеет человек, создать машину, способную, например, читать – давняя мечта ученых. В течение последних 50 лет эта мечта стала воплощаться в реальность. Распознавание символов – одно из самых плодотворных направлений исследований в области искусственного интеллекта. Современные программные продукты, предназначенные для оптического распознавания печатных символов, успешно заменяют оператора, вводящего в компьютер текст с готового печатного документа, особенно при наличии документов хорошего и среднего качества печати. Специализированные системы, ориентированные на промышленное использование, могут вводить более 100 000 документов в сутки, что позволяет обрабатывать огромные объемы информации, ранее компьютерной обработке не подлежавшие.

Системы оптического распознавания символов (OCR – Optical Character Recognition) предназначены для перевода растрового изображения в текст. Для получения растрового изображения страницы текста обычно используется монохромный сканер с разрешением от 300 dpi, который позволяет получить вполне удовлетворительное с точки зрения распознавания растровое изображение.

Программа распознавания должна преобразовать это изображение в текстовый формат, т. е. присвоить каждому символу некоторый код. На первом этапе программа выполняет предварительную обработку растрового изображения, которая включает в себя:

- подавление имеющихся яркостных помех;
- коррекцию возможных пространственных искажений;
- бинаризацию, т. е. присвоение пикселям изображения значения яркости из множества целых чисел (0, 1);
- сегментацию, т. е. выделение фрагмента изображения, включающего только один символ.

В результате каждому распознаваемому символу ставится в соответствие фрагмент бинарного изображения, на котором пиксели, имеющие значение 1, считаются относящимися к символу, а пиксели, имеющие значение 0 – к фону.

На втором этапе производится последовательное распознавание отдельных символов путем сравнения их с имеющимися эталонными образцами, полученными ранее в процессе обучения системы. В результате распознаваемому символу присваивается код того эталона, с которым он совпадает наилучшим образом.

Все алгоритмы распознавания печатных символов можно разделить на два класса: шрифтовые и безшрифтовые. Шрифтовые алгоритмы используют априорную информацию о шрифте (тип, размер), которым напечатаны буквы, т. е. программе OCR на этапе обучения необходимо предъявить полноценную выборку текста, напечатанного данным шрифтом. Программа измеряет и анализирует различные признаки каждого символа шрифта и заносит их в свою базу эталонных образов. По окончании этого процесса программа готова к распознаванию символов данного конкретного шрифта.

Безшрифтовые алгоритмы не имеют априорных знаний о шрифте символов, поступающих к ним на вход, т. е. измеряют и анализируют различные признаки, присущие печатным символам как таковым независимо от типа и размера шрифта. В предельном случае для безшрифтового алгоритма процесс обучения может отсутствовать и характеристики символов задает человек. Более часто база эталонных образов создается в процессе обучения программы на выборке реальных символов.

Сравнительные недостатки и достоинства обоих классов алгоритмов распознавания текстовых символов приведены в табл. 1.

Результаты экспериментов показывают, что разные изображения одного и того же печатного символа практически всегда отличаются друг от друга, причем степень отличия зависит как от качества печатной продукции, так и от технических параметров сканера (число градаций яркости, величина разрешения). В лабораторной работе эти отличия имитируются равномерным зашумлением, уровень которого можно регулировать.

Вероятность правильного распознавания символа зависит как от выбранного метода распознавания, так и от степени различия эталонных символов. В лабораторной работе можно попиксельно формировать растровые изображения эталонных символов с разной степенью отличия.

В лабораторной работе не рассматривается первый этап распознавания, т. е. предполагается что бинаризация и сегментация уже проведены и положение фрагментов изображения текста, в которых находятся распознаваемые символы, известно. Таким образом, остается лишь задача последовательного распознавания изображений отдельных символов, которая решается в работе тремя методами (по площади, по век-

Таблица 1

Свойства	Шрифтовые алгоритмы	Бесшрифтовые алгоритмы
Универсальность	Малая (обусловлена необходимостью предварительного обучения всему, что предъ- является для распознавания)	Большая (обусловлена неза- висимостью обучающей выборки от какой-либо сис- темы априорной классификации символов)
Точность распозна- вания	Высокая (обусловлена деталь- ной классификацией символов в процессе обучения и тем, что материал распознавания нахо- дится строго в рамках классов, созданных при обучении)	Низкая (обусловлена высокой степенью обобщения и огрубленными измерениями признаков символов)
Техноло- гичность	Низкая (обусловлена значи- тельными накладными расхода- ми, связанными с поддержкой классификации символов)	Высокая (обусловлена отсутствием какой-либо априорной системы классификации символов)
Поддержка процесса распозна- вания пользова- телем	Необходима (при обучении для задания системы классификации, при распознавании для указания конкретных классов символов)	Не требуется

тору признаков и методом маски) путем последовательного сравнения признаков текущего символа с признаками эталонных символов.

В методе распознавания по площади используется единственный признак – площадь символа, т. е. число пикселей в прямоугольном фрагменте известного размера (в работе – квадрат со стороной 65 пикселей), априори относящихся к символу (в работе – светлые пиксели). В результате зашумления распознаваемого символа значение его признака (площади) увеличивается на число ярких пикселей в области фона и уменьшается на число темных пикселей в области эталонного изображения данного символа, что приводит к значительной флуктуации значения признака на разных изображениях одного и того же символа. Расстояние между распознаваемым символом и текущим эталоном определяется как

$$d_s^{(i)} = S - S_i, \quad (6)$$

где $d_s^{(i)}$ – расстояние между распознаваемым символом и i -м эталоном по методу площади; S , S_i – площадь распознаваемого символа и i -го эталона соответственно.

При распознавании по вектору признаков в качестве отдельного признака используется число светлых пикселей в строке и столбце изображения. Следовательно, число элементов в векторе признаков в лабораторной работе равно 130 (65 строк и 65 столбцов). Расстояние между распознаваемым символом и текущим эталоном определяется как

$$d_p^{(i)} = \sum_{k=1}^{135} (r_k - r_k^{(i)})^2, \quad (7)$$

где $d_p^{(i)}$ – расстояние между распознаваемым символом и i -м эталоном по вектору признаков; $r_k, r_k^{(i)}$ – значение k -го признака распознаваемого символа и i -го эталона соответственно.

При распознавании методом маски последовательно производится сравнение изображений распознаваемого символа и каждого эталона по всем пикселям. В данном случае каждый пиксель можно считать признаком, имеющим значение 0 или 1. Понятно, что число признаков составляет $65 \cdot 65 = 4225$, а расстояние между распознаваемым символом и текущим эталоном определяется как

$$d_m^{(i)} = \sum_{k=1}^{65} \sum_{l=1}^{65} (z_{kl} - z_{kl}^{(i)}), \quad (8)$$

где $d_m^{(i)}$ – расстояние между распознаваемым символом и i -м эталоном по методу маски; $z_{kl}, z_{kl}^{(i)}$ – значение яркости пикселя с координатами (k, l) распознаваемого символа и i -го эталона соответственно.

При определении расстояния по формулам (6–8) в лабораторной работе не устанавливается порог достоверности, т. е. распознаваемый символ будет отнесен к ближайшему эталону даже при очень значительном расстоянии до него. На практике экспериментально подбирается значение максимально допустимого расстояния, при превышении которого вычисление расстояния до текущего эталона прекращается, а расстояние между этим эталоном и распознаваемым символом считается неопределенным. Если расстояния до всех эталонов неопределенные, то символ относится к классу нераспознанных символов.

Порядок выполнения работы

1. Сформировать несколько (3–4) новых эталонных символов с различной степенью и типом отличия (площадь, форма).
2. Сформировать текст, включающий одинаковые символы.
3. Выбрать уровень зашумления и метод распознавания, зафиксировать результат распознавания и расстояние до ближайшего эталона для каждого распознаваемого символа.

4. Повторить п. 3 для других методов распознавания при выбранном уровне зашумления.
5. Повторить п. 3 при других уровнях зашумления.
6. Повторить пп. 2 и 3 для других символов в сформированном тексте.
7. Получить положительную оценку за ответы на тестовые вопросы.

Содержание отчета

1. Изображения сформированных эталонных символов, их характеристики.
2. Результаты выполнения пп. 2–6 в табличной форме.
3. Подобрать порог достоверности распознавания сформированных символов для каждого метода.
4. Выводы о связи между методами распознавания, уровнем зашумления и степенью отличия символов.

Контрольные вопросы

1. Приведите типичные характеристики изображения, используемого для распознавания печатного текста.
2. Перечислите основные этапы предварительной обработки изображения перед распознаванием.
3. На какие два класса можно разделить алгоритмы распознавания печатных символов? В чем их основные достоинства и недостатки?
4. Какие методы распознавания используются в лабораторной работе?
5. Как определяется расстояние между распознаваемым символом и эталоном в этих методах?
6. Зависят ли приведенные в работе признаки от каких-либо пространственных искажений?

ЛАБОРАТОРНАЯ РАБОТА № 4

Кластерный анализ при распознавании образов

Цель работы: Ознакомиться с наглядной интерпретацией многомерных образов в виде кластеров на плоскости по произвольно выбранным парам признаков. Освоить статистическое описание образов в многомерном пространстве признаков и процедуры оценки информативности признаков для их минимизации. Исследовать особенности распознавания образов по расстоянию Махаланобиса и разделением пространства признаков плоскостями.

Теоретические пояснения

Любой физический объект обладает набором некоторых свойств, которые, собственно, и позволяют отличать один объект от другого. Совокупность свойств, описывающих конкретный объект, называется образом данного объекта. Под классом объектов понимается некоторая совокупность образов, называемых элементами класса, обладающая рядом близких свойств. Измеряемые или вычисляемые свойства объектов, позволяющие отличить классы друг от друга, называются признаками. В пределе каждый класс может состоять только из одного элемента, как, например, при опознавании человека. С другой стороны, вся совокупность образов может быть разделена всего на два класса, например «свой», «чужой».

Пусть число признаков равно n , тогда каждый образ можно представить в виде некоторого набора значений признаков или вектора признаков вида $\mathbf{x} = (x_1, \dots, x_n)$, где x_i – значение i -го признака данного образа. В многомерном пространстве, осями координат которого являются отдельные признаки, каждый образ определяется точкой, причем расстояние от этой точки до начала координат и расстояние между точками определяется выбранной метрикой.

Каждый класс может характеризоваться некоторой областью в n -мерном пространстве признаков. Эта область определяется степенью отличия образов, относящихся к данному классу, и может задаваться границами, например плоскостями. Область, в которую с наибольшей вероятностью попадают векторы образов данного класса, обычно называют кластером, а процедуру разбиения пространства признаков на области, соответствующие разным классам, – кластеризацией совокупности образов.

Если образы, относящиеся к одному классу, имеют различные значения признаков, то для описания класса можно использовать статис-

тические характеристики: вектор средних значений или вектор математических ожиданий по всем признакам и вектор среднеквадратичных отклонений или дисперсий признаков. Для многомерных векторов принято использовать матрицу ковариации, на диагонали которой находятся дисперсии соответствующих признаков.

Статистические характеристики признаков класса вычисляются следующим образом. Пусть набор n -мерных векторов, полученных в результате предъявления распознающей системе m объектов, относящихся к одному классу, задан в виде таблицы:

$$\begin{array}{cccccc} x_{11} & x_{12} & \dots & \dots & x_{1m} & \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & \dots & x_{nm} & \end{array}$$

Тогда

$$\mathbf{m} = (\mu_1, \mu_2, \dots, \mu_n); \mu_i = \frac{1}{m} \sum_{k=1}^m x_{ik}, \quad (9)$$

$$\mathbf{Cov} = \begin{pmatrix} D_{11} & D_{12} & \dots & D_{1n} \\ D_{21} & D_{22} & \dots & D_{2n} \\ \dots & \dots & \dots & \dots \\ D_{n1} & D_{n2} & \dots & D_{nn} \end{pmatrix} : \begin{cases} D_{ii} = \frac{1}{m} \sum_{k=1}^m (x_{ik} - \mu_i)^2; \\ D_{ij} = \frac{1}{m} \sum_{k=1}^m (x_{ik} - \mu_i)(x_{jk} - \mu_j), \end{cases} \quad (10)$$

где $i, j=1, \dots, n$ – индексы номеров компонент вектора признаков; m – число образов, составляющих данный класс; x_{ik} – значение i -го признака k -го образа; μ_i – математическое ожидание i -й компоненты вектора признаков; D_{ii} – дисперсия i -го признака; D_{ij} – коэффициент ковариации i -го и j -го признаков; \mathbf{m} – вектор математических ожиданий; \mathbf{Cov} – ковариационная матрица. Из выражения (10) видно, что ковариационная матрица симметрична относительно главной диагонали и, следовательно, необходимо вычислять только половину ее элементов. Ковариация характеризует степень линейной зависимости случайных величин. Если ковариация равна нулю, то величины называются некоррелированными.

Необходимо отметить, один важный, но малоизвестный факт, касающийся ковариационной матрицы. Если число образов, относящихся к некоторому классу, меньше или равно числу признаков, то ковариационная матрица, вычисляемая по этому множеству образов, будет вырожденной при любых значениях признаков каждого образа. Справедливость данного утверждения легко показать, если

представить ковариационную матрицу как результат матричного умножения

$$\mathbf{Cov} = \mathbf{X}_c \mathbf{X}_c^T,$$

где $\mathbf{X}_c = \begin{pmatrix} \mathbf{x}_1 - \mathbf{m} \\ \dots \\ \mathbf{x}_m - \mathbf{m} \end{pmatrix}$.

Как известно из линейной алгебры, определитель полученной таким образом матрицы равен нулю, если $m \leq n$, следовательно, обратная ковариационная матрица существует только при выполнении условия $m > n$.

Таким образом, процесс распознавания включает в себя этап обучения, т. е. определение некоторых характеристик по обучающей репрезентативной выборке образов, отнесенных к известным классам, и собственно распознавание на основе информации, полученной при обучении.

Под распознаванием образа понимается отнесение его к тому классу, расстояние до которого от данного образа в пространстве признаков минимально. Для определения расстояния между точками в метрическом пространстве необходимо ввести понятие метрики, т. е. определить процедуру измерения расстояния d_{lp} между точками l и p в этом пространстве так, чтобы выполнялись следующие аксиомы:

- симметричность расстояния ($d_{lp} = d_{pl}$);
- правило треугольника ($d_{lh} + d_{hp} > d_{lp}$);
- положительность расстояния ($d_{lp} \geq 0$, причем $d_{lp} = 0$, только если $l = p$).

Самой употребительной метрикой является евклидова метрика, в которой расстояние определяется следующим образом:

$$d_{lp} = \sqrt{\sum_{k=1}^n (x_{kl} - x_{kp})^2},$$

где x_{kl} , x_{kp} – k -е координаты точек l и p соответственно.

Для одномерных векторов ($k = 1$), т. е. в случае распознавания по одному признаку $d_{lp} = x_l - x_p$. Расстояние между классами по одному признаку, учитывающее разброс значений признака для образов, относящихся к одному классу, можно вычислить по критерию Фишера

$$F_{lp} = \frac{(\mu_l - \mu_p)^2}{D_l + D_p},$$

где μ и D вычисляются по формулам (9), (10). Из приведенного на рис. 3 сравнения плотностей распределения вероятности значений признака для двух классов при одинаковых значениях математического ожидания и разных дисперсиях видно, что значение критерия Фишера и вероятность правильного распознавания для пары классов 1–2 больше, чем для пары 1–3. Таким образом, с увеличением критерия Фишера растет вероятность правильного распознавания по данному признаку, т. е. критерий Фишера позволяет оценить информативность признака.

Расстояние между распознаваемым образом и классом l при одном признаке, учитывающее статистические характеристики признака, можно найти по расстоянию Фишера

$$d_F = \frac{(x - \mu_l)^2}{D_l}.$$

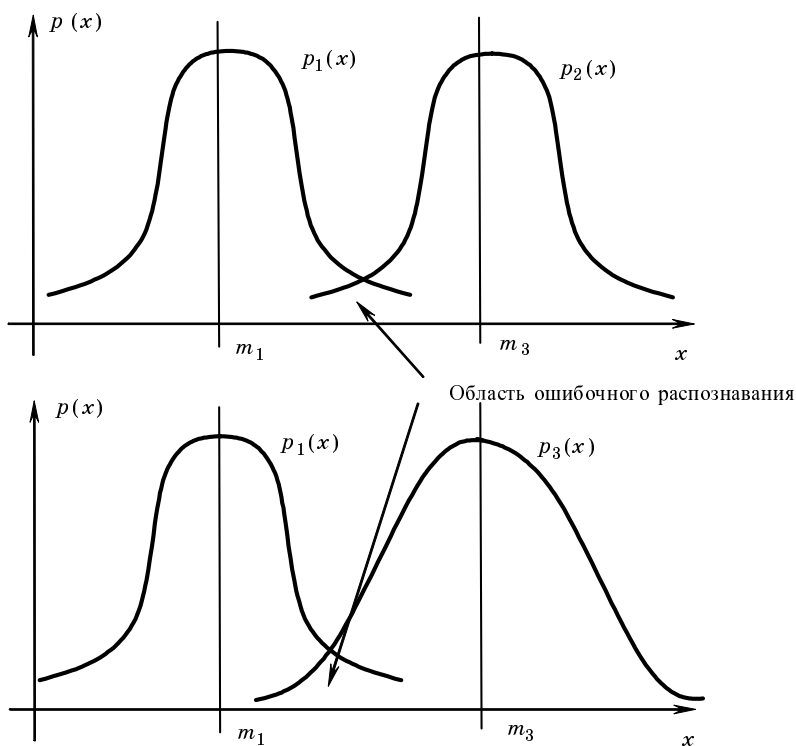


Рис. 3. Плотность распределения вероятности значения признака

Обобщением расстояния Фишера на многомерный случай является расстояние Махаланобиса $d_M = (\mathbf{x} - \mathbf{m}_l) \text{Cov}_l^{-1} (\mathbf{x} - \mathbf{m}_l)^T$, где Cov_l^{-1} – обратная ковариационная матрица для класса l .

Методы распознавания можно разделить на две группы. К первой группе относятся методы, основанные на последовательном вычислении расстояний между распознаваемым образом и центрами кластеров в пространстве признаков с той или иной метрикой, например расстояния Махаланобиса. Образ считается принадлежащим классу, расстояние до которого минимально.

Другая группа методов предполагает предварительное разбиение признакового пространства плоскостями или поверхностями более высокого порядка на области, соответствующие известным классам, и определение принадлежности распознаваемого образа одной из областей, например, для двух признаков производится разбиение плоскости набором прямых.

И те и другие методы имеют ряд достоинств и недостатков, и выбор конкретного метода распознавания в каждом случае зависит от таких особенностей, как число признаков, требуемая надежность и скорость распознавания, технические характеристики системы получения значений признаков и т. п.

Порядок выполнения работы

1. Ввести номер варианта и сформировать значения трех признаков для десяти образов нового класса с учетом представленных статистических характеристик уже имеющихся классов.
2. Проанализировать изменение значения математического ожидания и дисперсии при изменении признаков отдельных образов.
3. Записать вероятность правильного распознавания при кластеризации образов по расстоянию Махаланобиса для всех классов, включая и вновь созданный класс.
4. По критерию Фишера выбрать наиболее информативную пару признаков.
5. Для выбранной пары признаков провести разбиение признакового пространства (плоскости) на области кластеризации с максимальной достижимой вероятностью правильного распознавания.
6. Получить положительную оценку за ответы на тестовые вопросы.

Содержание отчета

1. Значения признаков для всех образов вновь созданного класса.
2. Значения математического ожидания, дисперсии и вероятности распознавания по расстоянию Махаланобиса для всех классов.

3. Таблица информативности признаков по критерию Фишера.
4. Эскиз разбиения признакового пространства для выбранной пары признаков и соответствующие вероятности распознавания образов.
5. Выводы по сравнению двух методов кластеризации.

Контрольные вопросы

1. Чем характеризуется образ в пространстве признаков?
2. Как можно представить класс в пространстве признаков?
3. Как вычисляются статистические характеристики класса?
4. Чем являются диагональные элементы ковариационной матрицы?
5. Что такое расстояние в метрическом пространстве?
6. Как вычисляется расстояние Махаланобиса?

Рекомендуемая литература

1. Ту Дж., Гонсалес Р. Принципы распознавания образов: Пер.с англ. М.: Мир, 1978. 411 с.
2. Ерош И. Л., Игнатьев М. Б., Москалев Э. С. Адаптивные робототехнические системы: Методы анализа и системы обработки изображений: Учеб. пособие / ЛИАП. Л., 1985. 144 с.
3. Путятин Е. П., Аверин С. И. Обработка изображений в робототехнике. М.: Машиностроение, 1990. 320 с.
4. Гонсалес Р., Вудс Р. Цифровая обработка изображений: Пер. с англ. М.: Техносфера, 2005. 1072 с.
5. Ерош И. Л., Сергеев М. Б., Соловьев Н. В. Обработка и распознавание изображений в системах превентивной безопасности: Учеб. пособие / ГУАП. СПб., 2006. 150 с.

Содержание

Предисловие	3
Лабораторная работа № 1. Изменение проекции объекта при его вращении в пространстве	4
Лабораторная работа № 2. Коррекция проективных искажений на стандартном носителе	9
Лабораторная работа № 3. Методы распознавания зашумленных символов	14
Лабораторная работа № 4. Кластерный анализ при распознавании образов	19
Рекомендуемая литература	25