Одесский Национальный Политехнический Университет

Кафедра информационных систем

Лабораторная работа № 1

по дисциплине: «Интеллектуальный анализ данных»

на тему «Исследование геометрических мер близости объектов и классов в системах распознавания»

Выполнил:

Ст. группы АИ-166

Дидух Э. Г.

Проверил:

Герганов М. Л.

Бойко В. Д.

Одесса, 2017

**ЗАДАНИЕ НА ЛАБОРАТОРНУЮ РАБОТУ:**

1. Розробити алгоритм прийняття рішення в детермінованій системі розпізнавання на основі використання відомих геометричних мір близькості.

2. Виконати програмну реалізацію розробленого алгоритму, налагодити програму.

3. Виконати контрольні розпізнавання невідомих об'єктів по векторах їхніх ознак (задати самостійно).

4. Порівняти ефективність прийняття рішень про віднесення невідомих об'єктів до заданих класів для різних методів обчислення відстані між об’єктом та класом. Визначити найкращу комбінацію з запропонованих.

**Додаткові завдання:**

1. Запропонувати інші методи обчислення міри близькості двох об’єктів та відстані між об’єктом та класом (розробити або відшукати в літературі).

2. Порівняти запропоновані методи з тими, що задані згідно варіанту.

**Етапи поступового виконання роботи:**

1. Визначити значення ознак п’яти еталонів для кожного з чотирьох класів (початкові дані — згідно табл. 1.1 та 1.4).

2. Відобразити на графіку визначені еталони всіх класів, екземпляри кожного класу у вигляді точки певної форми та кольору.

3. Реалізувати послідовне інтерактивне введення векторів ознак невідомих об’єктів, які треба розпізнати та віднести одного з чотирьох класів (значення ознак можна визначати за допомогою натискань клавіші мишки, запитувати у користувача або задавати фіксовані значення в програмному коді).

4. Реалізувати функції обчислення міри близькості між невідомим об’єктом та заданим об’єктом певного класу (методи обчислення міри близькості між двома об’єктами — згідно табл. 1.2 та 1.4).

5. Реалізувати функції обчислення відстані від кожного невідомого об’єкту до заданого класу (методи обчислення відстані між об’єктом та класом — згідно табл. 1.3 та 1.4).

6. Реалізувати можливість вибору комбінацій методів обчислення міри близькості та відстані між об’єктами.

7. Обчислити відстані від невідомого об’єкту до кожного з чотирьох класів.

8. Визначити клас, відстань від невідомого об’єкту до якого є мінімальною (у відповідністю з обраною комбінацією методів обчислення міри близькості між об’єктами та відстані між об’єктом та класом).

9. Відобразити розпізнаний об’єкт на графіку вигляді точки, яка за формою та кольором збігається з класом, приналежність до якого було визначено.

**Результат работы программы:**

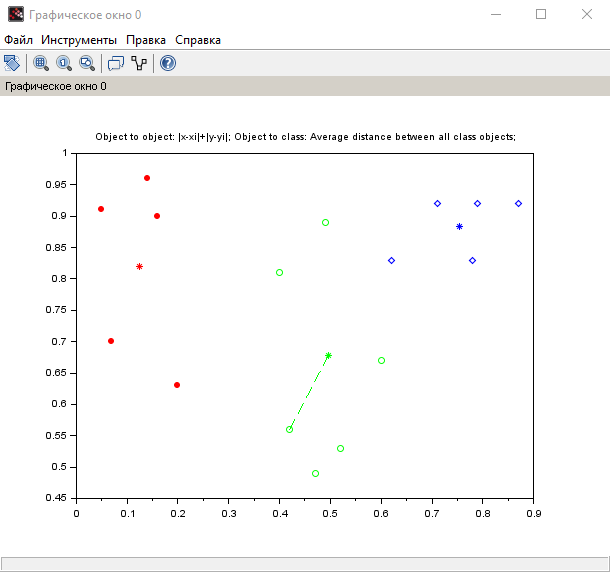


Рис. 1 Классификация методами 3/3

|  |
| --- |
| Код программы (методы 3/3) |
| x1=[0.05,0.14,0.16,0.07,0.20];  y1=[0.91,0.96,0.90,0.70,0.63];  x2=[0.49,0.40,0.60,0.47,0.52];  y2=[0.89,0.81,0.67,0.49,0.53];  x3=[0.62,0.79,0.71,0.78,0.87];  y3=[0.83,0.92,0.92,0.83,0.92];  trigger=1;  plot(x1,y1,'r.',x2,y2,'go',x3,y3,'bd')  while trigger==1    dist\_I=0;  dist\_II=0;  dist\_III=0;  x\_manual=input("Input the x-coordinate of the point: ");  y\_manual=input("Input the y-coordinate of the point: ");  for i=1:length(x1)    dist\_I=dist\_I+abs(x\_manual-x1(i))+abs(y\_manual-y1(i));  dist\_II=dist\_II+abs(x\_manual-x2(i))+abs(y\_manual-y2(i));  dist\_III=dist\_III+abs(x\_manual-x3(i))+abs(y\_manual-y3(i));    end    dist\_I\_avg=dist\_I/length(x1);  dist\_II\_avg=dist\_II/length(x2);  dist\_III\_avg=dist\_III/length(x3);    [class\_det,index]=min([dist\_I\_avg,dist\_II\_avg,dist\_III\_avg]);    if index==1 then  plot(x\_manual,y\_manual,'r.')  elseif index ==2 then  plot(x\_manual,y\_manual,'go')  elseif index==3 then  plot(x\_manual,y\_manual,'bd')  else  plot(x\_manual,y\_manual,'m\*')  end    trigger=input("Continue?(0/1) ");    end |

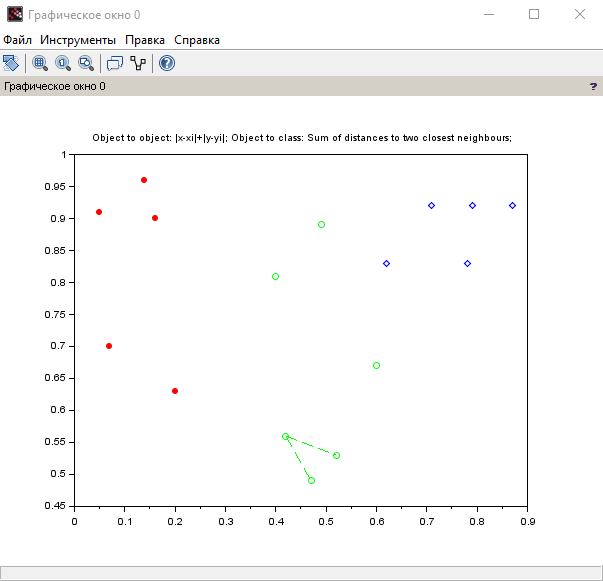


Рис. 2 Классификация методами 3/5

|  |
| --- |
| Код программы (методы 3/5) |
| x1=[0.05,0.14,0.16,0.07,0.20];  y1=[0.91,0.96,0.90,0.70,0.63];  x2=[0.49,0.40,0.60,0.47,0.52];  y2=[0.89,0.81,0.67,0.49,0.53];  x3=[0.62,0.79,0.71,0.78,0.87];  y3=[0.83,0.92,0.92,0.83,0.92];  trigger=1;  while trigger==1    clf("reset");    title('Object to object: |x-xi|+|y-yi|; Object to class: Sum of distances to two closest neighbours;');    plot(x1,y1,'r.',x2,y2,'go',x3,y3,'bd')    dist\_I=0;  dist\_II=0;  dist\_III=0;    dist\_I\_sum=0;  dist\_II\_sum=0;  dist\_III\_sum=0;  x\_manual=input("Input the x-coordinate of the point: ");  y\_manual=input("Input the y-coordinate of the point: ");  for i=1:length(x1)    dist\_I=abs(x\_manual-x1(i))+abs(y\_manual-y1(i));  dist\_I\_vector(i)=dist\_I;  dist\_I\_sum=dist\_I\_sum+dist\_I;    dist\_II=abs(x\_manual-x2(i))+abs(y\_manual-y2(i));  dist\_II\_vector(i)=dist\_II;  dist\_II\_sum=dist\_II\_sum+dist\_II;    dist\_III=abs(x\_manual-x3(i))+abs(y\_manual-y3(i));  dist\_III\_vector(i)=dist\_III;  dist\_III\_sum=dist\_III\_sum+dist\_III;    end    [dist\_I\_avg,ind\_I]=min(dist\_I\_vector);  disp(dist\_I\_vector, "Distances to all class A objects: ");  dist\_I\_vector(ind\_I)=max(dist\_I\_vector);  [dist\_I\_I,ind\_I\_I]=min(dist\_I\_vector);  dist\_I\_avg=dist\_I\_avg+min(dist\_I\_vector);  disp(dist\_I\_avg, "Summary distance to two closest neighbours: ");    [dist\_II\_avg,ind\_II]=min(dist\_II\_vector);  disp(dist\_II\_vector, "Distances to all class B objects: ");  dist\_II\_vector(ind\_II)=max(dist\_II\_vector);  [dist\_II\_I,ind\_II\_I]=min(dist\_II\_vector);  dist\_II\_avg=dist\_II\_avg+min(dist\_II\_vector);  disp(dist\_II\_avg, "Summary distance to two closest neighbours: ");    [dist\_III\_avg,ind\_III]=min(dist\_III\_vector);  disp(dist\_III\_vector, "Distances to all class C objects: ");  dist\_III\_vector(ind\_III)=max(dist\_III\_vector);  [dist\_III\_I,ind\_III\_I]=min(dist\_III\_vector);  dist\_III\_avg=dist\_III\_avg+min(dist\_III\_vector);  disp(dist\_III\_avg, "Summary distance to two closest neighbours: ");    [class\_det,index]=min([dist\_I\_avg,dist\_II\_avg,dist\_III\_avg]);    if index==1 then  plot(x\_manual,y\_manual,'r.')  plot([x\_manual, x1(ind\_I)],[y\_manual, y1(ind\_I)], 'r--')  plot([x\_manual, x1(ind\_I\_I)],[y\_manual, y1(ind\_I\_I)], 'r--')  elseif index ==2 then  plot(x\_manual,y\_manual,'go')  plot([x\_manual, x2(ind\_II)],[y\_manual, y2(ind\_II)], 'g--')  plot([x\_manual, x2(ind\_II\_I)],[y\_manual, y2(ind\_II\_I)], 'g--')  elseif index==3 then  plot(x\_manual,y\_manual,'bd')  plot([x\_manual, x3(ind\_III)],[y\_manual, y3(ind\_III)], 'b--')  plot([x\_manual, x3(ind\_III\_I)],[y\_manual, y3(ind\_III\_I)], 'b--')  else  plot(x\_manual,y\_manual,'m\*')  end    trigger=input("Continue?(0/1) ");    end |

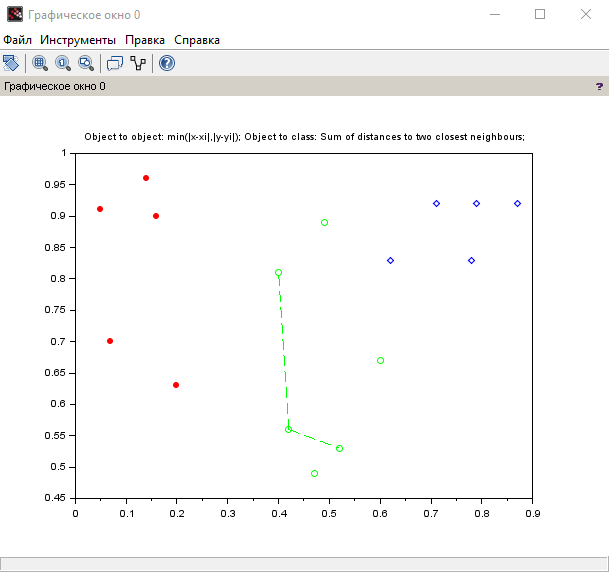


Рис. 3 Классификация методами 4/5

|  |
| --- |
| Код программы (методы 4/5) |
| x1=[0.05,0.14,0.16,0.07,0.20];  y1=[0.91,0.96,0.90,0.70,0.63];  x2=[0.49,0.40,0.60,0.47,0.52];  y2=[0.89,0.81,0.67,0.49,0.53];  x3=[0.62,0.79,0.71,0.78,0.87];  y3=[0.83,0.92,0.92,0.83,0.92];  trigger=1;  plot(x1,y1,'r.',x2,y2,'go',x3,y3,'bd')  while trigger==1    dist\_I=0;  dist\_I\_vector=zeros(x1);  dist\_II=0;  dist\_II\_vector=zeros(x2);  dist\_III=0;  dist\_III\_vector=zeros(x3);  x\_manual=input("Input the x-coordinate of the point: ");  y\_manual=input("Input the y-coordinate of the point: ");  for i=1:length(x1)    dist\_I=min([abs(x\_manual-x1(i)),abs(y\_manual-y1(i))]);  dist\_I\_vector(i)=dist\_I;    dist\_II=min([abs(x\_manual-x2(i)),abs(y\_manual-y2(i))]);  dist\_II\_vector(i)=dist\_II;    dist\_III=min([abs(x\_manual-x3(i)),abs(y\_manual-y3(i))]);  dist\_III\_vector(i)=dist\_III;    end    [dist\_I\_avg,ind\_I]=min(dist\_I\_vector);  dist\_I\_vector(ind\_I)=max(dist\_I\_vector);  dist\_I\_avg=dist\_I\_avg+min(dist\_I\_vector);    [dist\_II\_avg,ind\_II]=min(dist\_II\_vector);  dist\_II\_vector(ind\_II)=max(dist\_II\_vector);  dist\_II\_avg=dist\_II\_avg+min(dist\_II\_vector);    [dist\_III\_avg,ind\_III]=min(dist\_III\_vector);  dist\_III\_vector(ind\_III)=max(dist\_III\_vector);  dist\_III\_avg=dist\_III\_avg+min(dist\_III\_vector);    [class\_det,index]=min([dist\_I\_avg,dist\_II\_avg,dist\_III\_avg]);    if index==1 then  plot(x\_manual,y\_manual,'r.')  elseif index ==2 then  plot(x\_manual,y\_manual,'go')  elseif index==3 then  plot(x\_manual,y\_manual,'bd')  else  plot(x\_manual,y\_manual,'m\*')  end |

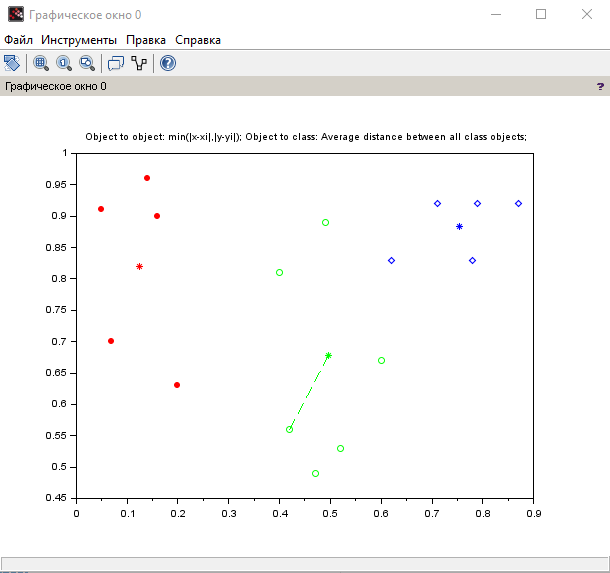


Рис. 3 Классификация методами 4/3

|  |
| --- |
| Код программы (методы 4/3) |
| x1=[0.05,0.14,0.16,0.07,0.20];  y1=[0.91,0.96,0.90,0.70,0.63];  x2=[0.49,0.40,0.60,0.47,0.52];  y2=[0.89,0.81,0.67,0.49,0.53];  x3=[0.62,0.79,0.71,0.78,0.87];  y3=[0.83,0.92,0.92,0.83,0.92];  trigger=1;  plot(x1,y1,'r.',x2,y2,'go',x3,y3,'bd')  while trigger==1    title('Object to object: min(|x-xi|,|y-yi|); Object to class: Average distance between all class objects;');    dist\_I=0;  dist\_I\_vector=zeros(x1);  dist\_II=0;  dist\_II\_vector=zeros(x2);  dist\_III=0;  dist\_III\_vector=zeros(x3);  x\_manual=input("Input the x-coordinate of the point: ");  y\_manual=input("Input the y-coordinate of the point: ");  for i=1:length(x1)    dist\_I=min(abs(x\_manual-x1(i)), abs(y\_manual-y1(i)));  dist\_I\_vector(i)=dist\_I;  dist\_I\_sum=dist\_I\_sum+dist\_I;    dist\_II=min(abs(x\_manual-x2(i)), abs(y\_manual-y2(i)));  dist\_II\_vector(i)=dist\_II;  dist\_II\_sum=dist\_II\_sum+dist\_II;    dist\_III=min(abs(x\_manual-x3(i)), abs(y\_manual-y3(i)));  dist\_III\_vector(i)=dist\_III;  dist\_III\_sum=dist\_III\_sum+dist\_III;    end    dist\_I\_avg=dist\_I\_sum/length(x1);  disp(dist\_I\_vector, "Distances to all class A objects: ");  disp(dist\_I\_avg, "Average distance: ");  dist\_II\_avg=dist\_II\_sum/length(x2);  disp(dist\_II\_vector, "Distances to all class B objects: ");  disp(dist\_II\_avg, "Average distance: ");  dist\_III\_avg=dist\_III\_sum/length(x3);  disp(dist\_III\_vector, "Distances to all class C objects: ");  disp(dist\_III\_avg, "Average distance: ");    [class\_det,index]=min([dist\_I\_avg,dist\_II\_avg,dist\_III\_avg]);    x1\_avg = sum(x1)/length(x1);  y1\_avg = sum(y1)/length(y1);    x2\_avg = sum(x2)/length(x2);  y2\_avg = sum(y2)/length(y2);    x3\_avg = sum(x3)/length(x3);  y3\_avg = sum(y3)/length(y3);    plot(x1\_avg, y1\_avg, 'r\*')  plot(x2\_avg, y2\_avg, 'g\*')  plot(x3\_avg, y3\_avg, 'b\*')    if index==1 then  plot(x\_manual,y\_manual,'r.')    plot([x\_manual, x1\_avg], [y\_manual, y1\_avg], 'r--');    elseif index ==2 then  plot(x\_manual,y\_manual,'go')    plot([x\_manual, x2\_avg], [y\_manual, y2\_avg], 'g--');    elseif index==3 then  plot(x\_manual,y\_manual,'bd')    plot([x\_manual, x3\_avg], [y\_manual, y3\_avg], 'b--');  else  plot(x\_manual,y\_manual,'m\*')  end    trigger=input("Continue?(0/1) ");    end |

Среди использованных методов определения расстояния между объектами большей точностью обладает метод суммы модулей разностей признаков. Данный метод обладает лучшим приближением к Эвклидовому расстоянию, чем метод минимального модуля разности признаков, при этом за счет исключения необходимости вычисления квадратного корня обладает сравнительно высокой скоростью работы. При использовании метода определения расстояния по минимальному модулю разности признаков высока вероятность ошибочной классификации при расположении классов близко по одной из осей признаков.

Из двух методов определения расстояния от объекта до класса предпочтительней использовать метод среднего расстояния ко всем эталонам класса, поскольку он лучше показывает себя в неоднозначных ситуациях. С точки зрения скорости работы метод среднего расстояния также превосходит метод двух ближайших соседей, поскольку требует лишь одного итеративного прохода по эталонам каждого класса.

**Выводы:** В данной лабораторной работе мы разработали алгоритм решения задачи классификации в детерминированной системе распознавания, выполнили программную реализацию алгоритма, отладили программу с помощью распознавания неизвестных объектов по векторам их признаков.