

Во многих организациях и конструкторских бюро Республики Беларусь работа ведется с документами в бумажной форме. Безусловно, такое представление наглядно и удобно для человека, однако оно далеко от идеала. Путем сканирования можно получить растровое представление документа. У растра есть недостатки: достаточный большой объем данных, потеря семантической информации, трудность редактирования объектов изображения. Под векторной формой понимается цифровое представление точечных, линейных и полигональных пространственных объектов в виде набора координатных пар, с описанием только геометрии объектов. Как следствие, векторная форма хранения характеризуется компактностью, простотой обращения к объектам, с ее помощью несложно осуществлять трехмерное моделирование объектов. Векторизацией называется перевод изображения из растрового формата в векторный.

Под серией будем понимать последовательность отсчетов, имеющих одинаковое значение [1]. Серия описывается только координатами граничных элементов. Каждый столбец матрицы изображения представляется упорядоченным по возрастанию координат списком серий. Изображение хранится в виде массива списков серий (МСС). Выделим связанные компоненты (СК) изображения. Затем выполним частичную скелетизацию СК, которая заключается в следующем. Осуществим скан-проход по МСС, проанализируем связность смежных серий и выделяются полосы (в англоязычной литературе используется термин *stroke*). Полоса представляет отдельную «ветвь» СК изображения. Полоса может содержать строго одну серию из столбца изображения, т.е. в полосе отсутствуют случаи ветвления и слияния серий. Под длиной полосы будем понимать количество серий, которые образуют полосу. Под весом полосы будем понимать суммарное число отсчетов ее серий. Если квадрат длины полосы больше ее веса, то полоса закрывается.

Скелетной кривой (СКР) в непрерывном пространстве будет являться либо линейный скелет либо средняя ось, сохраняющие топологические или геометрические признаки. Мы задаем СКР с помощью множества из  $N$  целочисленных точек  $pnt_0, pnt_1, pnt_{N-1}$ . СКР имеет характеристику ширины. На атрибуты СКР задаются следующие ограничения:

$$N \geq 3 \quad (1)$$

$$|pnt_{i+1}.x - pnt_i.x| \leq 1, \quad i = (0, 1, \dots, N-2) \quad (2)$$

$$|pnt_{i+2}.y - pnt_i.y| \leq 4, \quad i = (0, 1, \dots, N-3) \quad (3)$$

Закрываем полосы является формирование СКР по центральным точкам серий полосы. Серии, образовавшие СКР, удаляются из полосы. СКР представляются только те серии полосы, чьи центральные точки соответствуют условиям(1)–(3). После окончания вертикального скан-прохода изображение поворачивается на  $90^\circ$ , и выполняется скан-проход. Затем изображение поворачивается в исходное положение.

В результате двух скан-проходов прямолинейные отрезки СК заменяются скелетными скривыми (СКР). Группы серий, которые не были заменены СКР на процедуре частичной скелетизации, представляют собой области соединений (например, X, T, Y -типа). Из области соединения (ОС) исходят СКР, аппроксимирующие относительно прямолинейные участки. Для каждой СКР, исходящей из ОС получим вектор направления, построенный по ее начальным точкам. Найдём точку пересечения векторов направлений ОС и соединим ее отрезками с начальными точками СКР. Пометим точки растра, через которые проходят эти отрезки флагом. Затем применим параллельный алгоритм утоньшения для ОС, который не будет удалять помеченные флагом пиксели. Таким способом обеспечивается корректная обработка соединений.

В результате вышеописанных действий каждая СК изображения представляется планарным графом, вершинам которого соответствуют концевые и узловые точки отрезков СК, а ребрам – сами отрезки СК, представленные в форме СКР. Сохраним информацию об узловых и концевых точках в отдельные массивы [2](МУТ, МКТ). Следующим шагом является построение векторной модели на основании имеющегося графа. Применяется метод генерализации Дугласа-Пекера [3]. Пусть имеется цепочка вершин  $V = \{V_0, V_1, \dots, V_n\}$ . Рассмотрим процедуру аппроксимации. Построим отрезок  $V_0V_n$ . Пусть  $V_k$  – самая удалённая от отрезка  $V_0V_n$  вершина. Если расстояние от  $V_k$  до отрезка  $V_0V_n$  меньше заданного порога, то  $V_0V_n$  аппроксимирует цепочку. В противном случае разобьем цепочку  $V$  на 2 части:  $V_1 = \{V_0, \dots, V_k\}$  и  $V_2 = \{V_k, \dots, V_n\}$ . Для каждой цепочки рекурсивно применяется процедура аппроксимации.

В качестве постобработки предлагается стыковка отрезков в местах соединений. Это осуществляется анализом МУТ и МКТ.

Достоинством предлагаемого подхода является простота реализации и эффективность. Благодаря кодированию серий выделение СК и построение скелета осуществляется быстрее, чем при использовании классических методов попиксельного анализа. Также решается проблема обработки соединений. Строится графовая модель, которая может быть использована в задачах поиска шаблона изображения в базе изображений. Алгоритм может быть успешно применен для векторизации планов зданий, технических чертежей.

#### Литература

Абламейко С. В., Лагуновский Д. М. Обработка изображений: технология, методы, применения. Учебное пособие. – Мн.: Амалфея, 2000. – 304 с.

Семенков О.И., Абламейко С.В., Берейшик В.И., Старовойтов В.В. Обработка и отображение информации в растровых графических системах. - Минск: Наука и техника. - 1989 - 183с.

Douglas, David H.; Peucker, Thomas K.: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. In: The Canadian Cartographer 10 (1973), No. 2, P. 112–122.