

ИНФОРМАТИКА

УДК 515.142.33

АЛГОРИТМ ПОСТРОЕНИЯ ГРАФОВОЙ МОДЕЛИ БИНАРНЫХ ШТРИХОВЫХ ИЗОБРАЖЕНИЙ

М.В. СТЕРЖАНОВ

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровки, 6, Минск, 220013, Беларусь*

Поступила в редакцию 12 июня 2009

Описывается методика представления штриховых бинарных изображений в виде планарного нагруженного ориентированного псевдографа. Данная форма представления является компактной и сохраняющей топологию объектов. Предлагается гибридный алгоритм построения графовой модели, основанный на кодировании изображения в виде концов серий и представляющий модификацию алгоритма выделения связанных компонент. При этом области соединения сегментов обрабатываются отдельно с помощью аналитического реконструирования. Результат работы может быть использован при векторизации растра, в системах классификации изображений и поиска шаблонов изображения.

Ключевые слова: бинарный растр, RLE-кодирование, планарный псевдограф.

Введение

Под штриховыми изображениями будем понимать бинарные изображения, каждый объект которых представлен совокупностью линий, имеющих относительно одинаковую толщину на всем протяжении [1]. Особенности штриховых изображений являются: большой размер (А0), смешанная графическая и текстовая информация, различные типы объектов (отрезки, символы, дуги кривых). Примерами штриховых изображений являются технические чертежи, планы зданий, электросхемы.

При сканировании широкоформатных чертежей получается растровое изображение большого размера, что накладывает определенные ограничения на применяемые алгоритмы обработки. В частности, необходима структура данных, которая будет обеспечивать компактное хранение изображения, сохраняя его топологию.

Пусть имеется бинарный растр, $F=(f_{ij})_{M \times N}$. При использовании битового или байтового кодирования в отдельных ячейках хранится информация о каждом элементе изображения. Данный способ является неэффективным. Для хранения изображения с размерами 10000·10000 требуется более 95 Мб оперативной памяти при битовом кодировании. При выполнении многих операций над растром (например, выделение связанных компонент, операций математической морфологии) необходимо исследовать все изображение, независимо от размеров объектов изображения. Поэтому логично использовать сжатое представление, хранящее информацию только о точках интереса. Мы используем кодирование концов серий в качестве базового представления изображения. Данная методика даст возможность значительно снизить объем памяти при обработке больших штриховых изображений.

Рассмотрим известные подходы к построению графовых моделей изображений, кодированных длинами серий. Описание многих алгоритмов обработки изображений

упрощается при использовании графа смежности строк [2](LAG). Вершины такого графа соответствуют сериям, а ребра соединяют соседние вершины. Отметим, что в LAG используются только вертикальные либо горизонтальные серии. LAG описывает топологию объектов, и является удобным способом представления изображения, состоящего из большого числа отрезков. Однако, данная структура не хранит информацию о местах соединения, что является серьезным недостатком. В простом вертикальном графе [3](VSG) ребра состоят из нескольких смежных серий. Как и LAG, VSG состоит только из вертикальных серий, что является неэффективным способом представления горизонтальных отрезков. Решением данной проблемы служит граф смешанных серий [4] (MRG), состоящий из горизонтальных и вертикальных серий. MRG является объединением двух VSG, построенных по горизонтальным и вертикальным сериям. Граф серий [5] (RG) может быть построен напрямую по кодированному длинами серий результату сканирования изображения, т.е. служит заменой классическому матричному представлению растра. В данном формате представления узлами являются особые серии, а ребрами — соединяющие их полосы. RG применяется для анализа формы и вычисления характеристик объектов, однако, согласно с исследованиями [6], он неполно описывает топологию штриховых изображений.

В данной работе предлагается представление каждого объекта изображения в виде планарного нагруженного ориентированного псевдографа, в котором все ребра суть прямолинейные отрезки или дуги плоских кривых, а вершины — точки на плоскости, являющиеся концами отрезков или точками сочленения нескольких отрезков.

Данное представление позволяет обеспечить достаточную степень сжатия информации и в то же время обрабатывать изображение напрямую в кодированной форме, что позволяет при обработке изображений большого размера обойтись без разбиения изображения на фрагменты с последующей "сшивкой".

Используемая терминология

Под серией будем понимать последовательность отсчетов, имеющих одинаковое значение. В зависимости от ориентации серия может быть горизонтальной или вертикальной [5]. Серия однозначно определяется с помощью четырех значений:

d — направление (вертикальное или горизонтальное);

pos — номер столбца(строки) матрицы изображения, которому принадлежит серия;

beg — номер строки(столбца) матрицы изображения, которому принадлежит первый пиксель серии;

end — номер строки(столбца) матрицы изображения, которому принадлежит последний пиксель серии.

Для серии выполняется очевидное условие $beg \leq end$.

Две серии A и B называются смежными, если выполняются условия (1)–(4):

$$A.d = B.d, \quad (1)$$

$$|A.pos - B.pos| = 1, \quad (2)$$

$$A.beg \leq B.end + 1, \quad (3)$$

$$A.end \geq B.beg - 1. \quad (4)$$

Под путем из серии A к серии B будем понимать последовательность серий $A=A_1, A_2, \dots, A_n=B$, таких, что A_i является смежной с A_{i+1} для $1 \leq i \leq n-1$.

Рассмотрим две смежные серии. Они находятся в отношении "родитель-потомок". Серию с меньшим значением pos будем называть родительской. Серию с большим значением pos будем считать дочерней.

Серия будет называться нормальной, если она имеет ровно одного родителя и ровно одного потомка. В противном случае серия является особой.

Под начальной будем понимать серию, не имеющую родителей. Под конечной будем понимать серию, не имеющую потомков. Под серией слияния будем понимать серию, имеющую более одного родителя. Под серией ветвления будем понимать серию, имеющую более одного потомка.

Под полосой будем понимать связанное множество нормальных серий. Полной полосой будем называть полосу, которая не является подмножеством другой полосы. Полоса представляет отдельную "ветвь" 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.

Листинг первым ввел понятие линейного скелета [7], который образуется в результате континуального сжатия области (без изменения топологии) с выделением подмножества пикселей единичной толщины. Средняя ось формируется центрами дисков максимального радиуса, помещенных внутрь области [8].

Скелетной кривой (СКР) в непрерывном пространстве будет являться либо линейный скелет либо средняя ось, сохраняющие топологические или геометрические признаки [9]. Мы задаем СКР множеством из N целочисленных точек $pnt_0, pnt_1, \dots, pnt_{N-1}$. СКР имеет характеристику ширины. На атрибуты СКР задаются следующие ограничения:

$$N \geq 3, \tag{5}$$

$$|pnt.x_{i+1} - pnt.x_i| \leq 1, \quad i = (0, 1, \dots, N-2), \tag{6}$$

$$|pnt.y_{i+2} - pnt.y_i| \leq 4, \quad i = (0, 1, \dots, N-3). \tag{7}$$

Заметим, что условия (6)–(7) являются симметричными относительно координат x и y .

Закрытием полосы является формирование СКР по центральным точкам серий полосы (рис. 1). Серии, образовавшие СКР, удаляются из полосы. СКР представляются только те серии полосы, чьи центральные точки соответствуют условиям (5)–(7).

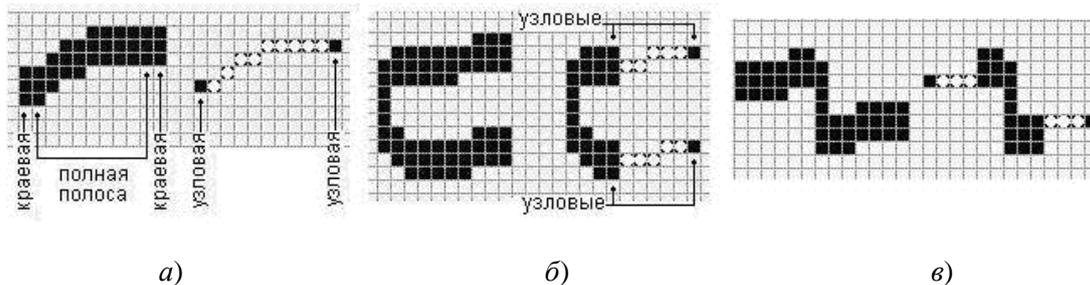


Рис. 1. Пример закрытия полос: *a* — узловые серии являются краевыми; *б* — узловые серии слева отличаются от краевых; *в* — полная полоса не может быть представлена одной СКР из-за нарушения условия

Рассмотрим полную полосу S . Серии, являющиеся родительской и дочерней по отношению к первой и последней сериям S соответственно, назовем краевыми. Обозначим краевые серии через E_1 и E_2 . Узловыми будем называть серии, к которым прикрепляются СКР при закрытии полосы. Если краевая серия является начальной или конечной, то она является узловой. В этом случае узловая серия "сужается" до своей центральной точки. Пусть краевая серия E_1 будет серией слияния или ветвления. Это значит, что серия E_1 принадлежит области соединения. Рассмотрим путь E_1E_2 . Тогда узловой будет серия E_{1+n} . Значение n зависит от толщины и наклона полосы. В данном случае размеры узловой серии не изменяются. Такой выбор узловых серий сохраняет топологию соединения.

Пусть имеется полная полоса S , состоящая из n серий R . Общее число отсчетов серий полосы S равняется m . После закрытия она представляется кривой C . Вычисляются следующие морфологические свойства: площадь рассчитывается просто как количество пикселей,

соответствующих полосе; длина вычисляется как евклидово расстояние между центрами узловых серий; ширина W рассчитывается по формуле:

$$W = m / \sum_{i=1}^n \sqrt{1 + \left(\frac{R_i.beg + R_i.end - R_{i-1}.beg + R_{i-1}.end}{2} \right)^2}. \quad (8)$$

Вычисление характеристик объектов

Каждый столбец матрицы изображения представляется упорядоченным по возрастанию координат списком серий. Изображение хранится в виде массива списков серий (МСС). Первый элемент МСС соответствует самому левому столбцу изображения, содержащему черный пиксель. Последний элемент МСС соответствует самому правому столбцу изображения, содержащему черный пиксель. Если столбец матрицы изображений не содержит черных пикселей, то соответствующее значение элемента МСС равно нулю. Кодирование с помощью длин серий является не просто удобным методом расходования памяти для хранения бинарного изображения, оно также хранит семантическую информацию. Так каждый отрезок на изображении представляется последовательностью смежных серий.

Выделим связанные компоненты (СК) изображения. Благодаря RLE-кодированию возможно использование эффективного алгоритма [10]. Его основная идея заключается в том, что метка СК ассоциируется не с отдельным пикселем, а с серией. Однако для больших изображений размер таблицы эквивалентности является фактором, снижающим производительность. Мы применяем методику "разделяй и властвуй" для нахождения СК изображений большого размера. Основываясь на работе [11], разделим изображение на $N \times N$ частей. Для каждой части применим алгоритм [10], затем выполним процедуру слияния. Каждая СК соответствует объекту на изображении и будет представлена псевдографом. Некоторые характеристики объектов могут быть легко вычислены, при использовании представления в виде МСС.

Вычисление характеристик объектов является предметом исследования с начала 1970-х. Правила и принципы поиска характеристик объектов зависят от особенностей предметной области и конкретного приложения (задачи). В данном разделе мы опишем основные параметры объектов штриховых изображений и методики их нахождения.

Отметим, что изначально расчет характеристик объектов ведется в единицах, определяемых системой. Например, площадь и периметр вычисляются в пикселях. Для получения результатов в реальных единицах необходимо выполнить калибровку. Перед началом измерений мы осуществляем линейную калибровку по известным размерам и разрешению изображения.

Геометрические характеристики лежат в основе любого анализа объекта. Для каждой СК вычислим координаты описывающего прямоугольника — минимального по площади многоугольника с осями, параллельными осям координат, внутри которого содержится данная компонента. Площадь характеризует общие размеры объекта. Площадь вычисляется как сумма всех пикселей СК, умноженная на горизонтальный и вертикальный масштабные коэффициенты. В соответствии с формами представления объектов различают несколько способов вычисления площади. Простую площадь иногда называют сетевой (net area). Выпуклая площадь (convex area) — это площадь многоугольника, вершины которого соответствуют выступам объекта. Под полной площадью (filled area) понимают площадь заполненного объекта. Периметр определяется как сумма расстояний между граничными пикселями и характеризует длину контура объекта. Мы вычисляем периметр на основе 8-связных граничных точек. Полученная информация может быть использована при выделении шумов, точечных объектов, линий постоянной ширины и площадных объектов (анализируя отношение площади к периметру).

Основная ось (major axis) соответствует максимальной линии, которую можно провести через объект. В работе [5] доказывается эффективность нахождения основной оси в кодированном концом серий изображении. Производной характеристикой от основной оси

является ортогональная ось, которая определяется как максимальная линия внутри объекта, проходящая перпендикулярно основной оси. Ортогональная и основная оси образуют систему координат, в которой изучаются основные характеристики объекта. Комбинация характеристик осей, например, угол между ними или соотношение их длин, характеризуют топологию объекта.

Топологические характеристики описывают форму объекта. Мы вычисляем следующие: число Эйлера, характеризующее вложенность объектов внутри измеряемого; количество дыр (holes), или отверстий, внутри объекта.

С помощью *RLE*-кодирования могут быть легко вычислены текстурные характеристики. Несмотря на широкое применение термина "текстура", для него нет точного определения [12]. Часто под текстурой понимают повторяемость фрагментов изображения. Следовательно, наличие какой-либо периодичности является текстурной характеристикой объекта. Мы вычисляем горизонтальный и вертикальный интерцепты.

Полученные характеристики описывают объект в целом. В процессе построения графовой модели будут вычислены характеристики, относящиеся к элементам СК.

Построение графовой модели бинарного растра

Следующим действием является нахождение осевых линий объектов изображения. При использовании классических алгоритмов утоньшения возникают две проблемы: чувствительность к мелким неровностям границ объектов и погрешности в обработке соединений. Важным фактором также является производительность алгоритма. Нами предлагается использование быстрого алгоритма частичной скелетизации, основной задачей которого является нахождение осевых линий объектов простой формы. Более сложные фрагменты будут обработаны отдельно в дальнейшем. Выполним частичную скелетизацию СК, скан-проход которой заключается в следующем. Изображение сканируется по вертикали, анализируется связность смежных серий и выделяются полосы. Найденные полосы закрываются.

Очевидно, что в результате скан-прохода будут найдены осевые линии горизонтальных "ветвей" изображения. Чтобы обработать вертикальные объекты, выполняется следующее. После вертикального сканирования изображение поворачивается на 90° , снова выполняется скан-проход. Поворот изображения на 90° осуществляется без потерь, и является простой операцией. Затем изображение поворачивается в исходное положение (рис. 2). Заметим, что алгоритм частичной скелетизации инвариантен по отношению к поворотам объектов изображения на угол, кратный 90° .

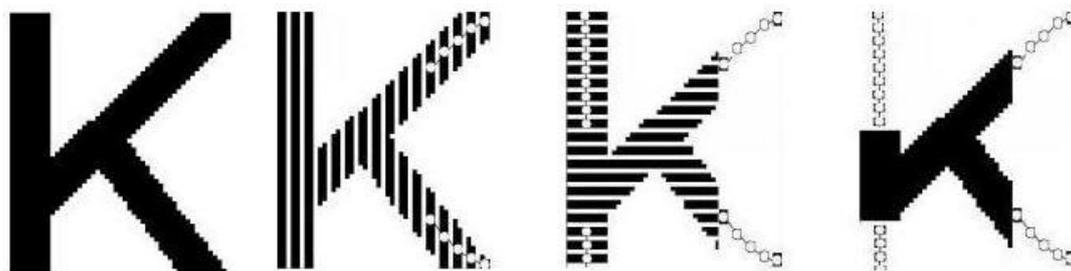


Рис. 2. Пример работы алгоритма частичной скелетизации

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. Найдем точку пересечения векторов направлений ОС и соединим ее отрезками с начальными точками СКР. Пометим точки растра, через которые проходят эти отрезки. Затем применим параллельный алгоритм утоньшения для ОС, который не будет удалять помеченные пиксели. Таким способом обеспечивается корректная обработка соединений.

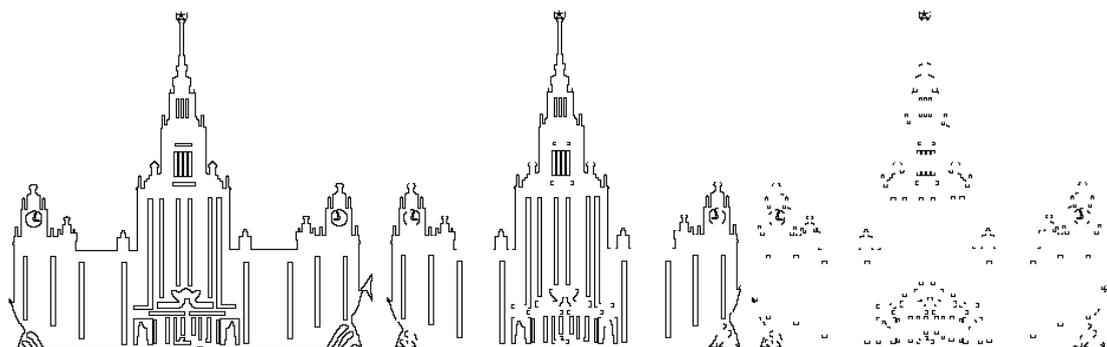


Рис. 3. Пример работы алгоритма частичной скелетизации

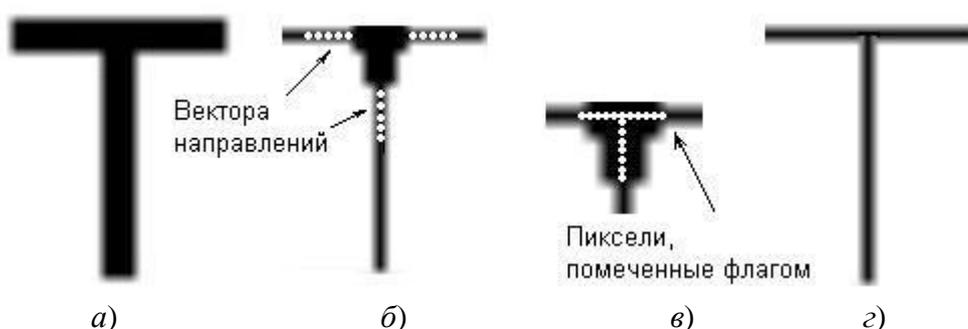


Рис. 4. Обработка соединений: *a* — исходное изображение; *б* — результат частичной скелетизации; *в* — ОС, белым показаны точки соединения; *г* — результат скелетизации

На следующем шаге анализом последовательности серий единичной ширины выделим группы серий, удовлетворяющие условиям (5)–(7). Заменяем эти серии с помощью СКР. В данном случае СКР будет являться альтернативным способом представления последовательности пикселей единичной толщины. Перейдем к построению графовой модели. По координатам узловых серий (т.е. серий, которые указывают на СКР) сформируем множество вершин V . По точкам СКР построим ребра. Каждая СКР была прикреплена к двум узловым сериям, которые преобразовались в вершины, следовательно, можно найти вершины, из которых исходят ребра. Рассмотрим два смежных ребра. Если они являются коллинеарными (допускается погрешность в 5°), то они могут быть заменены одним ребром.

В результате вышеописанных действий каждая СК изображения представляется нагруженным ориентированным планарным псевдографом, вершинам которого соответствуют концевые и узловые точки отрезков СК, а ребрам — сами отрезки СК, представленные в форме СКР. Независимые подходы к описанию и построению графовых моделей были предложены в [2, 4, 6, 13–15].

Граф G задается парой $G=(V, E)$, где V — множество вершин. Описание каждой вершины содержит координаты точки, порядок вершины и ее тип (концевая или узловая); E — мультимножество ребер, каждое из которых соединяет две вершины из V , причем изображения ребер из E на плоскости не пересекаются, поэтому (V, E) представляет собой планарный граф. Каждое ребро имеет важные характеристики (например: длина, ширина, элонгация), которые могут быть использованы при последующем создании векторной модели. Т.е. выделенные на этапе векторизации объекты могут быть снабжены соответствующими атрибутами.

Построенная графовая модель обладает важными свойствами. Графовая модель является компактной формой представления СК изображения. Она описывает топологию СК, связи между отрезками СК (ОСК) и позволяет осуществлять эффективное нахождение графических примитивов. Выделенные на этапе векторизации отрезки могут иметь атрибуты, соответствующие атрибутам ребер графа.

Под мультиточкой будем понимать точку пересечения как минимум трех отрезков [5]. Нахождение мультиточек является важной задачей при распознавании образов. Вершины графа, из которых исходит более трех ребер, являются мультиточками.

По графовой модели могут быть получены следующие характеристики:

а) отдельной СК: выделение петли на изображении СК; количество ОСК; длина каждого ОСК; общая длина всех ОСК; средняя длина всех ОСК; максимальная и минимальная длина ОСК; средняя элонгация всех ОСК; максимальная и минимальная элонгация ОСК; средняя ширина всех ОСК; максимальная и минимальная ширина ОСК.

б) изображения: количество объектов; количество ОСК всех объектов; суммарное, среднее, максимальное и минимальное значение параметров СК.

Граф может быть преобразован в более компактную форму гиперграфа, гиперребра которого состоят из ребер, соединяющих вершины степени 1 и 2 исходного графа (рис. 5).

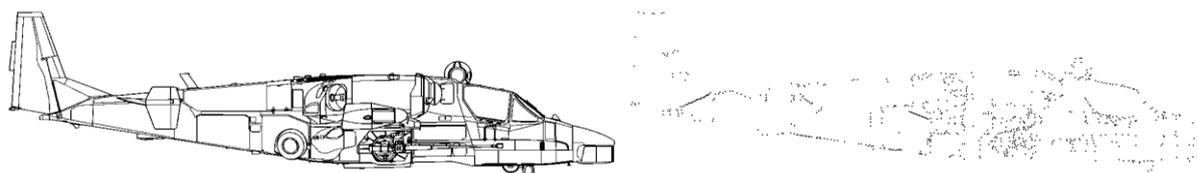


Рис. 5. Исходное изображение и вершины гиперграфа

Заключение

Эффективный метод представления растрового изображения должен обеспечивать достаточную степень сжатия информации и в то же время позволять обрабатывать изображение напрямую в кодированной форме. Например, векторизация часто осуществляется для изображений большого размера, и разбиение изображения на фрагменты с последующей "сшивкой" является источником ошибок.

Мы предлагаем использование графовой модели как компактного средства хранения и описания структуры раstra. Достоинством предлагаемого подхода является простота реализации и достаточно высокое быстродействие. Нами предлагается модификация алгоритма нахождения СК, созданная специально для обработки изображений большого размера. Благодаря кодированию концов серий построение скелета осуществляется быстрее, чем при использовании классических методов попиксельного анализа. Также решается проблема обработки соединений. Кроме того, в процессе построения графовой модели вычисляются важные геометрические, топологические и текстурные характеристики объектов. Исходная СК может быть реконструирована анализом графа. Недостатком подхода является неполное описание площадных объектов. Полученная графовая модель может быть успешно применена при векторизации планов зданий, технических чертежей [15].

GRAPH MODEL CONSTRUCTION ALGORITHM FOR BINARY LINE DRAWINGS

M. V. STERJANOV

Abstract

The method for the description of binary line drawings in a form of planar attributed oriented pseudograph is presented. This form of compression preserves the topology. We are extending further the work of Pavlidis, Di Zenzo, Monagan, amongst others. Hybrid algorithm for graph model construction is suggested. The algorithm is based on run length encoding of the image. Junctions are analysed separately using a specific procedure for analytical reconstruction. The result of the research can be used in the vectorization task, in image classification and image retrieval systems.

Литература

1. *Ablameyko S., Pridmore T.* Machine interpretation of line-drawing images, Springer. 2000.
2. *Pavlidis T.* // IEEE Trans. Systems, Man, and Cybernetic, 1978. № 8 (1). P. 66–69.
3. *Boatto L., Consorti L., et. al.* // IEEE Computer, 1992. № 25 (7). P. 25–33.
4. *Burge M.* The Representation and Analysis of Document Images, Vol. 100 of Computer Vision and Graphics Dissertations. Austrian Computer Society, Vienna, 1998.
5. *Di Zenzo S., Cinque L., Leviardi S.R.* // IEEE Transactions on PAMI, 1996. Vol. 18, No. 1. P. 83–89.
6. *Kropatsch W., Burge M., Glantz R.* Graphs in image analysis // Digital image analysis: selected techniques and applications. Springer-Verlag, 2001. P. 179–197.
7. *Listing J.* Dercensusraeumlicher complexe oder verallgemeinerungen des euler'schen satzes von den polyeadern., Abhandlungen der Mathematischen Classe der Koeniglichen Gesellschaft der Wissenschaften zu Goettingen. 1861.
8. *Blum H.* A transformation for extracting new descriptors of shape, Models for the Perception of Speech and Visual Form (W.Wathen-Dunn, ed.), MIT Press, P. 362–380, 1967.
9. *Klette G.* Topologic, Geometric, or Graph-Theoretic Properties of. Skeletal Curves, PhD dissertation, Groningen Univ., 2007. 122 p.
10. *Shapiro L.* Connected Component Labeling and Adjacency Graph Construction. Topological algorithms for digital image processing. Elsevier Science B.V. 1996. P. 1–31.
11. *Jung-Me P., Carl G. Looney, Hui-Chuan C.* // CATA 2000. Conference on Computers and Their Applications, 2000. P. 373–376.
12. *Старовойтов В.В.* Локальные геометрические методы цифровой обработки и анализа изображений. Минск, 1997.
13. *Monagan. G., Roosli M.* // Document Analysis and Recognition: Proc. Int. Conf., Japan. 1993. P. 623–626.
14. *Костюк Ю.Л., Новиков Ю.Л.* // Вестник ТГУ. 2002. № 275. С. 153–160.
15. *Стержанов М.В.* // GraphiCon'2009: Труды конф. 2009. С. 194–197.