

# ИЕРАРХИЧЕСКАЯ АДАПТИВНАЯ СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЯ

О.С. Агаронян

Рассмотрена задача иерархической сегментации изображения путем построения нерегулярной пирамиды. Для моделирования изображения предложен адаптивный подход, предполагающий использование диаграммы Вороного и графа Делоне. Представлена итеративная процедура сегментации для пирамиды с различными уровнями разрешения.

**Ключевые слова:** иерархическая модель изображения, нерегулярная пирамида, сегментация изображения, мозаичное замощение, диаграмма Вороного, граф Делоне.

## ВВЕДЕНИЕ

В машинной графике и обработке изображений широко используются иерархические структуры данных (пирамиды) для моделирования и описания изображений [1]. Необходимость в многоуровневых представлениях обусловлена появлением задач, требующих для своего решения упрощения абстрактных описаний и дискретных представлений сложных структур данных (с сохранением важных топологических свойств). Поэтому при проведении операций над объектами или областями изображения возникает проблема, состоящая в разработке математического аппарата, который позволял бы выполнять действия над частями изображений как над элементарными объектами. Иерархическое представление устраняет необходимость отдельной обработки каждого элемента исходного изображения, а именно: в то время как обычно каждая точка, обрабатываемая ЭВМ, имеет лишь координаты, положение в пространстве, но не имеет формы, пирамидальные структуры позволяют представить в качестве элементарных такие части изображений, которые одновременно имеют и размер, и положение, — это блоки (ячейки) различных уровней. Такой подход позволяет относить к той ли иной области целые фрагменты изображения, соответствующие элементам яркости различных уровней пирамиды. Важно также и взаимное расположение блоков, определяющих форму объекта на изображении. Эту характеристику уже нельзя вывести из «чисто» иерархии — здесь важны «горизонтальные» взаимосвязи блоков различных размеров, вытекающие из топологии объекта.

При обработке изображений чрезвычайно важны алгоритмы просмотра изображений. Например, просмотр дискретного изображения соответствует обходу некоторой дискретной сетки, которую можно интерпретировать как граф: его вершинами служат пиксели, а ребра связывают вершины, соответствующие смежным пикселям. Кроме такого простого отображения, можно построить и другие графы, соответствующие изображению. Развитие методов, основанных на графах, обусловлено тем, что они хорошо адаптируются к выделению формы естественных объектов. Так, в задачах описания и выделения связных регионов обработка изображений упрощается при наличии графа смежности областей. Вершины этого графа соответствуют отдельным областям, а ребра соединяют пары вершин, если соответствующие области смежные. Поэтому целесообразно описывать структуру пирамиды с помощью графа смежности.

Пирамидальные (иерархические) структуры могут использоваться для описания данных, например, в задачах кластеризации и сегментации статических изображений. Под иерархической сегментацией будем понимать последовательность разбиений изображения на вложенные сегменты. Отметим, что составные сегменты на различных уровнях не повторяются. Однако было показано [2], что алгоритмы, основанные на регулярных пирамидах, неинвариантны относительно сдвига (т. е. сегментация изображения и незначительно смещенная версия изображения могут значительно отличаться). Это послужило одним из основных обстоятельств для введения нерегулярных пирамид [3], которые действуют на общую структуру графа вместо регулярных окрестностей (как в слу-



чае регулярных пирамид). По этой причине они предоставляют больше возможностей в процессе построения пирамиды и могут адаптироваться к контексту изображения. Например, использование иерархий сегментов для представления изображения позволяет обеспечивать лучшую аппроксимацию объектов минимальным числом ячеек адекватной формы и размеров. Кроме того, в пирамидах изображений, основанных на регулярных решетках, всегда присутствуют артефакты, обусловленные жесткостью дискретной структуры. С другой стороны, пирамидальное изображение, определенное с помощью нерегулярной дискретной пирамиды, может быть сформировано на основе структуры входного изображения.

Для построения иерархической модели в настоящей статье предлагается использовать в качестве исходного представления изображения не регулярную решетку пикселей, а нерегулярное мозаичное замощение. В качестве замощения применяется диаграмма Г.Ф. Вороного и двойственный ей граф Б.Н. Делоне. Такой подход позволяет получать пирамиды с меньшим числом уровней (по сравнению с инициализацией иерархии от исходного изображения). Поскольку граф Делоне может быть к тому же вычислен параллельно [4], то и весь алгоритм удобен для параллельной обработки.

## 1. МОДЕЛИРОВАНИЕ ИЕРАРХИЧЕСКОГО ИЗОБРАЖЕНИЯ

Пирамидальная структура представляет собой упорядоченную последовательность изображений изменяющегося разрешения, располагаемых обычно одно над другим (рис. 1). Изображение разбивается на определенные блоки, затем процедура

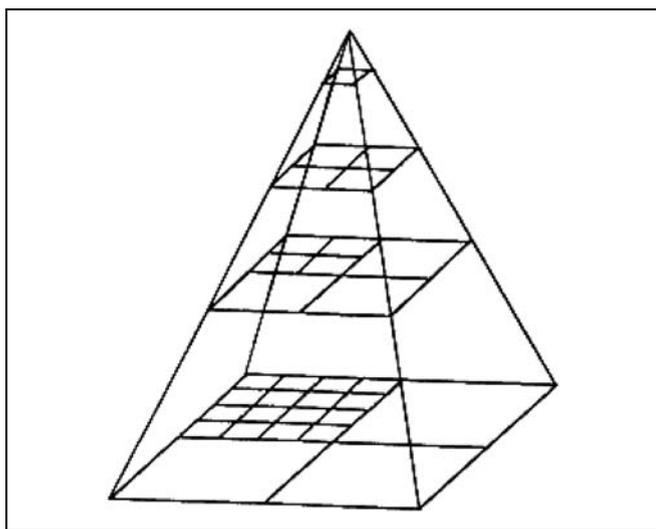


Рис. 1. Пирамида с несколькими уровнями разрешения

разбиения повторяется для каждого блока до тех пор, пока его размер не станет равным размеру элемента исходного изображения. Каждому блоку присваивается значение яркости или цвета. По завершении процесса получается набор изображений, каждое из которых состоит из блоков определенного размера. Эти изображения последовательно уточняют друг друга и сходятся к исходному. Построение пирамиды может идти и обратным путем — от низших уровней к высшим [1].

Пирамидальное описание предполагает некоторый закон перехода от одного уровня пирамиды к другому. Существуют три важные характеристики для описания пирамиды:

- структура, например, соседи, отношения предков-потомков между уровнями;
- содержимое ячейки, например, уровень серого для пикселя;
- обработка, производимая над ячейкой, например, фильтрация.

В настоящей работе рассматривается только структура пирамиды. Обсуждение содержания ячеек и производимой над ними обработки можно найти в работе [5].

### 1.1. Структура пирамиды

Структура пирамиды изображения определяется отношениями соседства внутри одного уровня пирамиды и отношениями предков-потомков между смежными уровнями. Базовый уровень пирамиды является исходным изображением. Структура пирамиды может быть также представлена графом. В этом случае каждый уровень пирамиды описывается с помощью графа соседства (смежности)  $G_i = \langle V_i, A_i \rangle$ . Здесь множество вершин  $V_i$  соответствует областям уровня  $i$ , и  $A_i \subseteq V_i \times V_i$  — отношения соседства пикселей. Две вершины  $p, q \in V_i$  связаны ребром в графе  $G_i$ , если они являются соседними в структуре (т. е.  $p, q \in A_i$ ).

Любая пирамида с  $n$  уровнями может быть описана посредством  $n$  графов соседства и  $n - 1$  вертикальных графов (описывающих связность между уровнями). Необходимо различать регулярные структуры и нерегулярные, зависящие от того, одинаковы ли структурные отношения для всех ячеек пирамиды (кроме граничных) или они могут изменяться от ячейки к ячейке.

В регулярной (классической) пирамиде изображения число пикселей на некотором уровне  $i$  в  $\lambda$  раз больше, чем число пикселей на уровне  $i + 1$ . При этом так называемый коэффициент изменения разрешения  $\lambda > 1$  и одинаков для всех уровней [6]. Такое пирамидальное представление предполагает один и тот же закон перехода от одного

уровня пирамиды к другому. Для построения пирамиды задается лишь правило преобразования группы элементов изображений, которое затем распространяется «вширь» — на остальные элементы, и «вглубь» — на другие уровни. Это правило описывает и схему преобразования исходных данных, и результирующую структуру данных; наполнение этой структуры происходит при задании конкретного изображения в качестве входной структуры.

В нерегулярной пирамиде каждый ее уровень представляет собой разбиение множества пикселей на ячейки, т. е. связанные подмножества пикселей. На базовом уровне нерегулярной пирамиды ячейками служат отдельные пиксели исходного изображения, а окрестности ячеек определяются заданной связностью пикселей. На уровне  $i + 1$  каждая ячейка (предок) представляет собой объединение соседних ячеек на уровне  $i$  (потомков) [3]. Значения каждого предка вычисляются независимо от других ячеек того же уровня, что позволяет проводить параллельные вычисления при обработке таких изображений. Каждый последующий уровень получается из предыдущего с помощью функции снижения разрешения (характеризующей отношение между уровнями).

Структура нерегулярной пирамиды представляет собой не только пучок графов. Она может также рассматриваться как дерево от корня (вершины пирамиды) до листьев (пикселей исходного изображения). Использование древовидной структуры в сегментации изображения означает, что заданный регион любого уровня может последовательно расщепляться на подобласти на следующем, более низком уровне. Просмотр дерева в направлении сверху вниз обеспечивает получение приближенного решения задачи сегментации после обработки  $i + 1$ -го уровня дерева, которое является уточнением решения после обработки  $i$ -го уровня.

В настоящей работе рассматриваются нерегулярные пирамиды. Известны два пути построения нерегулярной пирамиды: сжатие параллельного графа и снижение разрешения в графе соседства (смежности).

Полный формализм сжатия графа описан в работе [7]. Процесс снижения разрешения делит ячейки некоторого уровня пирамиды на два класса: ячейки, которые сохраняются на следующем уровне, и ячейки, которые не появляются на уменьшенном уровне (несохраняемые). В пирамиде с жесткой структурой априори известно, какие ячейки сохраняются. В случае нерегулярной пирамиды ее архитектура считается полностью определенной, если заданы как способ построения нового уровня, так и связь между предком и его потом-

ками. Тогда одинаковый процесс может быть применен для построения всей пирамиды (уровень за уровнем).

## 1.2. Нерегулярные пирамиды

Среди существующих подходов к построению нерегулярных пирамид наиболее интересные стохастический и адаптивный [6]. Для построения стохастической пирамиды используется как граф смежности, так и граф подобия. Процесс отбора сохраняемых вершин (или процесс децимации), определяющий процедуру снижения разрешения, задается следующими двумя правилами [3, 8]:

1) две соседние ячейки уровня  $i$  не могут быть одновременно сохранены на следующем уровне  $i + 1$ ;

2) для каждой несохраняемой ячейки существует, по крайней мере, одна сохраняемая ячейка в ее окрестности.

Процесс отбора вершин, удовлетворяющий этим условиям, называется «правильным». В теории графов задача определения сохраняемого множества вершин известна как проблема нахождения максимально независимого множества вершин графа [9]. Это означает, что выполнение правил 1 и 2 эквивалентно утверждению, что вершины  $V_{i+1}$  графа  $G_{i+1}$  на уровне  $i + 1$  определяют максимально независимое множество (вершин) для графа  $G_i = \langle V_i, A_i \rangle$ .

При построении стохастической пирамиды (снизу вверх) вначале всем ячейкам присваиваются случайные числа, а затем отбираются те, в которых достигается локальный максимум. Эти ячейки являются сохраняемыми, а остальные ставятся им в соответствие (используя граф подобия) и формируют поля (называемые рецептивными), которые определяют отношения соседства (рис. 2).

В работе [10] приведен алгоритм стохастического определения сохраняемых ячеек (для однородного случая).

При построении адаптивной пирамиды нет необходимости в использовании графа подобия для

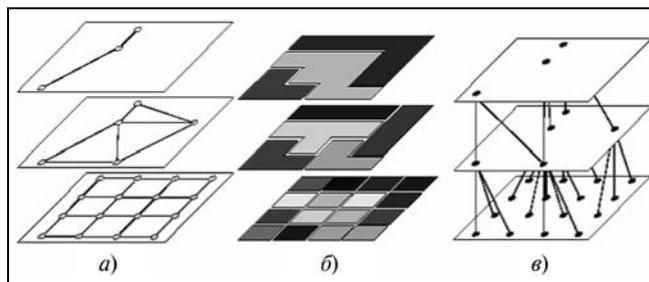


Рис. 2. Соответствие между областями графов смежности (а), рецептивными полями пирамиды (б) и структурой дерева (в)



процесса децимации. Сохраняемые ячейки отбираются в соответствии с графом смежности и с применением так называемой «интересующей характеристики», в качестве которой часто используется однородность областей изображения. В этом случае процедура определения сохраняемой информации состоит в выборе ячеек, представляющих локальные максимумы однородности. Тогда несохраняемые ячейки присоединяются к их наиболее сходным соседям. Таким образом, в случае адаптивной пирамиды выбор сохраняемых ячеек связывается с информацией, содержащейся в изображении. Условия генерации для адаптивной пирамиды аналогичны правилам для стохастической пирамиды.

Основные применения пирамид в обработке изображений встречаются в задачах сегментации на основе замощений, обнаружения объектов и выделения признаков [11—13]. Интерес к использованию нерегулярной пирамиды в задачах анализа изображения объясняется тем, что подобная иерархическая структура представляет собой множество разбиений исходного изображения, в котором форма областей является контекстно-зависимой. Это означает, что структура данных такой пирамиды позволяет выделять детали изображения, для чего каждое из разбиений, т. е. каждый уровень пирамиды, представляется как графом, так и набором рецептивных полей (см. рис. 2). Использование этого графа и свойств областей дает возможность проводить их объединение с помощью итеративного процесса отбора.

В настоящей работе для реализации иерархической сегментации изображения предлагается подход, основанный на построении начального описания изображения с помощью замощения Вороного, которое потом адаптивно улучшается. В следующем параграфе будут кратко изложены некоторые определения и результаты, используемые для построения иерархического изображения на основе многоугольников Вороного.

## 2. ИЕРАРХИЧЕСКИЕ СТРУКТУРЫ НА ОСНОВЕ МОЗАИЧНЫХ МОДЕЛЕЙ ИЗОБРАЖЕНИЯ

Применение вычислительной геометрии к анализу изображений касается, в основном, проблем определения структур данных. Одно из классических направлений в вычислительной геометрии связано с понятием диаграммы Вороного [14, 15].

Рассмотрим общую процедуру построения так называемых мозаичных моделей, к которым относится и диаграмма Вороного. Процесс моделирования случайной мозаики состоит из двух этапов: заполнения (замощения) области  $A$  ячейками с не-

которой структурой и независимого присваивания каждой ячейке одного из цветов  $c_1, \dots, c_m$  в соответствии с фиксированным набором вероятностей  $p_1, \dots, p_m$ , в сумме равных 1. С помощью такого процесса можно осуществить разбиение области  $A$  на подобласти  $A_1, \dots, A_m$  с различными геометрическими характеристиками, где  $A_j$  — объединение всех ячеек цвета  $j$ . Следовательно, разбиение  $A$  представляет собой реализацию случайного геометрического процесса.

### 2.1. Диаграмма Вороного и граф Делоне

Наибольшее применение при решении многочисленных и разносторонних практических задач приобрела случайная мозаичная модель Вороного [16—18]. Пусть задан стационарный пуассоновский процесс  $\Phi$  в пространстве  $R^2$  с некоторой интенсивностью. Разобьем пространство  $R^2$  на ячейки, внутри которых содержится по одной точке процесса  $\Phi$ , по следующему правилу: точке из процесса  $\Phi$  ставится в соответствие множество тех точек пространства  $R^2$ , которые являются ближайшими (в смысле евклидовой метрики) к заданной точке, называемой ядром. Таким способом построенные ячейки, называемые многоугольниками Вороного, представляют собой ограниченные выпуклые многоугольники (пересечения полуплоскостей). Полученное замощение плоскости называют мозаикой или диаграммой Вороного.

Пусть точки  $s_i \in S$  — ядра многоугольников Вороного. Тогда диаграммой Вороного  $V(S)$  множества  $S$  является разбиение плоскости на  $N$  многоугольников Вороного, вершины которых называются вершинами диаграммы Вороного. Методы построения диаграммы  $V(S)$  множества точек достаточно широко освещены в литературе. Наиболее распространен из них двойственный метод построения через триангуляцию Делоне [19]. Поэтому применение диаграммы  $V(S)$  для разработки вычислительных алгоритмов связано с вопросом построения триангуляции Делоне  $D(S)$  множества  $S$ , где  $D(S)$  — плоский граф, двойственный диаграмме Вороного (рис. 3). Он получается в результате соединения отрезками точек множества  $S$ , многоугольники Вороного которых имеют общее ребро [15].

Модели  $V(S)$  и  $D(S)$  в последние годы приобретают все больший интерес, поскольку обладают весьма полезными свойствами. Формы многоугольников Вороного отражают свойства локальных пространственных распределений точек. Кроме того, замощение Вороного присваивает каждую область плоскости в качестве окрестности одной и

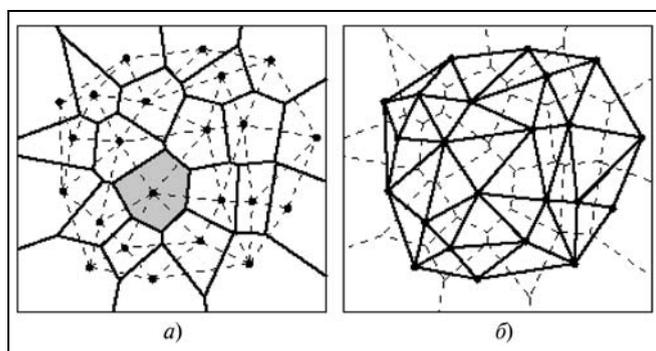


Рис. 3. Диаграмма Вороного (а) и триангуляция Делоне точек на плоскости (б)

только одной точке. Поэтому замощение плоскости дает однозначно порожаемое множество непересекающихся выпуклых областей.

Далее предлагается способ применения замощения Вороного в иерархической структуре для решения задачи сегментации полутонового изображения.

### 3. АДАПТИВНАЯ СЕГМЕНТАЦИЯ

Предположим, что в качестве исходного замощения изображения при построении иерархической модели применяется диаграмма Вороного. Для описания структуры данных используется представление в виде графа, что позволяет присваивать каждому многоугольнику его генератор (ядро), а также точки смежных многоугольников. При таком подходе связи между предками и потомками в пирамидальной конструкции определяются следующим образом. Предки (т. е. ядра многоугольников Вороного) выбираются при помощи случайного процесса. Тогда ребра определяются двойственной к замощению Вороного триангуляцией Делоне [17]. Соответственно, потомками являются ячейки покрытия, ассоциированного с предками.

Отметим, что пирамида строится сверху вниз. Ее применение позволяет генерировать различные сегментации изображения при различных разрешениях посредством использования древовидной структуры (представленной иерархией графов смежности). В такой структуре любая заданная область некоторого уровня может последовательно расщепляться на следующем, более низком уровне. При этом дерево используется для управления процессом обхода.

Рассмотрим кратко процедуру иерархической сегментации изображения, моделируемого замощением Вороного. Общая идея состоит в следующем: стартуя из случайной сегментации, изображение

итеративно улучшается и сходится к оптимальному разбиению (при заданном статистическом критерии). На первом этапе осуществляется инициализация на основе небольшого числа случайных генераторов, полученных с использованием пуассоновского процесса, что позволяет вычислять диаграмму Вороного с помощью известных методов [19]. На втором этапе в целях адаптации процесса используется принцип «расщепления и слияния» [20].

Пусть каждая точка изображения описывается значением ее уровня серого, а многоугольник Вороного характеризуется средним значением его пикселей  $m(P)$ . Стартуя с вершины структуры, критерий однородности вычисляется для каждой области: например, стандартное отклонение в рассматриваемом многоугольнике  $P$  сравнивается с пороговым значением  $\sigma(P)$  для принятия решения о необходимости дальнейшего расщепления этого региона. Итеративный процесс состоит во введении новых генераторов в неоднородные области и последующей модификации диаграммы Вороного. Процесс продолжается до тех пор, пока все многоугольники можно считать однородными.

Процедура «слияния» областей состоит в удалении «ненужного» генератора. Считаем, что если разность между средними значениями в соседних однородных многоугольниках меньше допустимого значения, то соответствующий генератор можно удалить. В результате смежные многоугольники увеличатся за счет удаленной области. Итеративный процесс, основанный на изменении статистического распределения геометрических областей, продолжается до достижения требуемой сегментации. При этом каждый регион исходного изображения выделяется на том уровне, где его представление — оптимально.

Вопрос о сходимости описанной процедуры сегментации решается по существу эвристически, поскольку сходимость итерационных процессов «расщепления и слияния» определяется на основе заданных статистических параметров и достигается при выполнении соответствующих условий.

Заметим, что помимо указанных статистических характеристик могут использоваться и другие параметры, особенно, если есть априорная информация о текстурированном или зашумленном изображении.

### ЗАКЛЮЧЕНИЕ

Предложен иерархический подход к моделированию изображений. Показано, что иерархическая декомпозиция изображения, обеспечиваемая нерегулярной пирамидой, позволяет получать раз-



личные сегментации на различных уровнях иерархии. Применение нескольких уровней разрешения дает возможность адаптации к характерным свойствам изображения: более гладкие области могут быть представлены с помощью меньшего числа точек, чем те, в которых необходима большая детализация. При этом настройка параметров дает возможность достигать желаемого уровня детализации объектов. Рассмотрены возможности применения диаграммы Вороного в качестве нерегулярного замощения. Предлагаемая инициализация пирамиды посредством графа Делоне приводит к уменьшению вычислений при построении пирамиды, поскольку число ячеек в базовом уровне уже уменьшено. Адаптация алгоритма реализуется на основе итерационной процедуры, позволяющей в интерактивном режиме выделять объект при наилучшем разрешении для заданных статистических характеристик.

## ЛИТЕРАТУРА

1. Александров В.В., Горский Н.Д. Представление и обработка изображений. Рекурсивный подход. — Л.: Наука, 1985. — 192 с.
2. Bister M., Cornelis J., Rosenfeld A. A critical view of pyramid segmentation algorithms // Pattern Recognition Lett. — 1990. — Vol. 11, N 9. — P. 605—617.
3. Meer P. Stochastic Image Pyramids // Computer Vision, Graphics Image Process. — 1989. — Vol. 45. — P. 269—294.
4. Saxena S., Bhatt C.P., Prasad V.C. Efficient VLSI parallel algorithm for Delaunay triangulation on orthogonal tree network in two and three dimensions // IEEE Trans. Comput. — 1990. — Vol. 38, N 3. — P. 400—404.
5. Bischof H., Kropatsch W.G. Neural networks versus image pyramids // Artificial Neural Networks and Genetic Algorithms / Ed. by R.F. Albrecht, C.R. Reeves. — Berlin New York: Springer-Verlag. — 1993. — P. 137—144.
6. Агаронян О.С. Идентификация статических объектов с использованием методов обработки иерархических изображений // Тр. VII междунар. конф. «Идентификация систем и задачи управления» (SICPRO'08). — М.: Институт проблем управления, 2008. — С. 347—357.
7. Kropatsch W.G., Leonardis A., Bischof H. Hierarchical, adaptive and robust methods for image understanding // Surveys on Mathematics for Industry. — 1999. — N 9. — P. 1—47.
8. Jolion J.M. Stochastic pyramid revisited // Pattern Recognition Letters. — 2003. — Vol. 24. — P. 1035—1042.
9. Luby. M. A simple parallel algorithm for the maximal independent set problem // SIAM Journ. of Computing. — 1986. — Vol. 15, N 4. — P. 1036—1053.
10. Агаронян О.С. Об одном методе моделирования адаптивного изображения // Тр. V междунар. конф. «Идентификация систем и задачи управления» (SICPRO'06). — М.: Институт проблем управления, 2006. — С. 306—315.
11. Burt P., Hong T., Rosenfeld A. Segmentation and estimation of image region properties through cooperative hierarchical computation // IEEE Trans. Syst. Man. Cybern. — 1981. — Vol. 11, N 12. — P. 802—809.
12. Montanvert A., Meer P., Bertolino P. Optimal hierarchical shape analysis in grey level images // Proc. of the NATO Advanced Workshop «Shape in Picture». — Driebergen, The Netherlands, September 7—11. — Springer Verlag, — 1992. — Vol. 126. — P. 13—34.
13. Peleg S., Federbusch O., Hummel R. Custom-made pyramids // Parallel Computer Vision. — 1998. — P. 125—146.
14. Вороной Г.Ф. Исследования о примитивных параллелоэдрах. Собр. соч. / Под ред. И.М. Виноградова. — Киев: Изд-во АН УССР, 1952. — Т. 2. — С. 293—369.
15. Препарата Ф., Шеймос М. Вычислительная геометрия: Введение. — М.: Мир, 1989. — 478 с.
16. Ласло М. Вычислительная геометрия и компьютерная графика на C++. — М.: Бином, 1997. — 304 с.
17. Агаронян О.С. Статистические аспекты применения диаграмм Вороного в задачах обработки изображений // Автоматика и телемеханика. — 1999. — № 4. — С. 58—64.
18. Rom H., Peleg S. Image representation using Voronoi tessellation: adaptive and secure // IEEE Comput. Vision Pattern Recognition. — 1988. — P. 282—285.
19. Watson D.F. Computing the n-dimensional Delaunay tessellation with application to Voronoi politopes // Comput. J. — 1981. — Vol. 24, N 2. — P. 167—172.
20. Павлидис Т. Алгоритмы машинной графики и обработки изображений. — М.: Радио и связь, 1986. — 400 с.

Статья представлена к публикации членом редколлегии Л.П. Боровских.

**Агаронян Ольга Сергеевна** — канд. техн. наук, ст. науч. сотрудник, Институт проблем управления РАН им. В.А. Трапезникова, г. Москва, ☎(495) 334-91-61, ✉ kirillium@mail.ru.

## Новая книга

**Александров А.Г. Методы построения систем автоматического управления.** — М.: Физматлит, 2008. — 232 с.

Изложены методы синтеза регуляторов и адаптивного управления, позволяющие построить алгоритмы управления объектами по их линейным моделям. Данные методы опираются на процедуры  $LQ$ - и  $H_\infty$ -оптимизации и конечно-частотную идентификацию, а алгоритмы обеспечивают требуемую точность регулирования при ограниченных неизвестных внешних возмущениях, действующих на объект. Описана доступная программная система ГАММА, являющаяся программным обеспечением части излагаемых методов и позволяющая оценить их эффективность при построении реальных систем.

Для инженеров-разработчиков современных систем автоматического управления и студентов соответствующих специальностей.