«Оптика атмосферы и океана», **16,** № 9 (2003)

МЕТОДЫ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

УДК 551.222.11

В.Т. Калайда

Применение методов морфологического анализа для задач идентификации полутоновых изображений

Томский государственный университет

Поступила в редакцию 17.07.2003 г.

Предлагаются модификация метода морфологического анализа изображений для идентификации полутоновых изображений и программно-техническая система распознавания, основанная на предлагаемом алгоритме.

Обработка и идентификация изображений традиционно основывались на принципах и теории линейных систем и на преобразовании Фурье (или на других подобных преобразованиях) [1, 2]. Во многих приложениях такой подход оказался достаточно эффективным, однако для изображений он не предоставляет возможности численно описать их форму или геометрическую структуру. Альтернативным подходом в анализе изображений является математическая морфология. Она позволяет дать количественное описание особенностей геометрической структуры.

Морфологический анализ формы изображений, базирующийся на теории множеств, интегральной геометрии, анализе выпуклых функций, стереологии и геометрической теории вероятностей, был разработан J. Serra и Ю.П. Пытьевым в 60-е гг. XX в. [3, 4]. Активное применение методов морфологического анализа обусловлено непрерывным развитием и усовершенствованием архитектур ЭВМ, используемых для реализации морфологических преобразований сигнала. Но несмотря на достаточно эффективную математическую структуру морфологического анализа, эти методы еще не получили широкого признания.

Основная идея морфологического анализа может быть представлена следующим образом.

Пусть R и Z представляют собой множества действительных и целых чисел, а E – d-мерное непрерывное пространство R^d (d = 1, 2, 3, ...) или дискретное пространство Z^d. Тогда d-мерный сигнал можно представить как функцию области R^d (непрерывной) или Z^d (дискретной), диапазон которой составляет либо R – при непрерывном изменении амплитуды, либо Z – при квантованном изменении амплитуды.

Двоичные сигналы могут быть представлены с помощью множеств. Двоичные изображения часто получают в результате пороговой селекции полутоновых изображений, и пороговая селекция часто используется для того, чтобы представить полутоновые изображения через двоичные сигналы, т.е. с помощью множеств. J. Serra [3] использует представление действительной d-мерной функции $f(x)$, где x означает d-мерный вектор, с помощью ансамбля его d-мерных пороговых множеств, определяемых как

$$T_a(f) = \{x : f(x) \geq a\}, \ -\infty < a < \infty,$$

где амплитуда a полностью перекрывает R или Z, в зависимости от того, имеет ли сигнал f непрерывный или квантованный диапазон. У пороговых множеств есть два важных свойства. Они линейно упорядочены, поскольку $a < b \Rightarrow T_a(f) \supseteq T_b(f)$, и позволяют однозначно восстановить сигнал, так как

$$f(x) = \max\{a : x \in T_a(f)\}, \ \forall x .$$

Следовательно, $f(x)$ может быть восстановлена по ансамблю $f_a(x)$, поскольку

$$f(x) = \max\{a : f_a = 1\}, \ \forall x .$$

Преобразования сигнала в математической морфологии представляют собой нелинейные операторы, локально модифицирующие геометрические характеристики многомерных сигналов. Пусть $X \subseteq E$ есть множественное представление двоичного входного сигнала и пусть $B \subseteq E$ есть компактное множество малого размера и простой формы (например, d-мерная сфера). Множество B называется структурирующим элементом. Пусть $X \pm b = \{x \pm b : x \in X\}$ выражает векторный перенос X на $\pm b \in E$. Основными морфологическими операторами для множеств являются наращение (dilation) $\oplus$ и эрозия (erosion) $\circ X$ с помощью B, они определяются следующим образом:

$$X \oplus B = \bigcup_{b \in B} X + b = \{x + b : \ x \in X \ \text{ и } \ b \in B\},$$

$$X \quad B = \bigcap_{b \in B} X - b = \{z : \ (B + z) \subseteq X\}.$$

Исходя из этих определений, можно показать, что выход оператора наращения представляет собой множество перенесенных точек, такое, что перенос отраженного структурирующего элемента $\breve{B} =$

$= \{-b : b \in B\}$ образует непустое пересечение с входным множеством, т.е. $X \oplus B = \left\{z : (\breve{B} + z) \bigcap X \neq \varnothing\right\}$. Аналогично выход оператора эрозии представляет собой множество перенесенных точек, такое, что перенесенный структурирующий элемент содержится во входном множестве.

Другие операторы определяются как комбинации эрозии и наращения. Например, два дополнительных оператора — размыкание (opening) $\circ$ и замыкание (closing) $\bullet X$ с помощью B определяются как

$$X \circ B = (X \quad B) \oplus B, \; X \bullet B = (X \oplus B) \quad B.$$

Описанный набор операторов может быть обобщен на многоуровневые (т.е. недвоичные) сигналы, представляемые действительно-значимыми функциями. J. Serra использовал представление d-мерной функции $f(x)$ набором ее пороговых множеств. При этом операция наращения всех пороговых множеств функции f с помощью одного и того же компактного множества B дает множества, $T_a(f) \oplus B$, которые являются пороговыми множествами новой функции $f \oplus B$, называемой наращением функции f с помощью B. Эта новая функция может быть вычислена либо как $(f \oplus B)(x) = = \max\{a : x \in T_a(f) \oplus B\}$, или из прямой эквивалентной формулы

$$(f \oplus B)(x) = \max_{y \in B}\{f(x - y)\},$$

$$(f \quad B)(x) = \min_{y \in B}\{f(x + y)\}.$$

Размыкание ($\circ$) и замыкание ($\bullet$) функции f с помощью B определяются как $f \circ B = (f \quad B) \oplus B$ и $f \bullet B = (f \oplus B) \quad B$.

Эрозия функции f с помощью малого выпуклого множества B уменьшает число пиков и повышает минимумы функции. Наращение функции f с помощью B увеличивает долины и удлиняет максимумы функции. Размыкание с помощью B сглаживает график функции f снизу путем срезания пиков, а замыкание сглаживает его сверху путем заполнения долин функции f.

Еще одно расширение состава морфологических операторов для функций было произведено Штернбергом [8], который использовал представление d-мерной функции $f(x)$ посредством $(d + 1)$-мерного множества. Соответствующий теневой оператор выражается в виде

$$U(f) = \{(x, a) : \; a \leq f(x)\},$$

т.е. тень (umbra) есть множество точек, лежащих ниже поверхности, выражаемой $f(x)$. В общем, множество, выражающее тень, расширяется до $a = -\infty$. Функция может быть восстановлена по ее тени, поскольку

$$f(x) = \max\{a : (x, a) \in U(f)\}, \;\; \forall x.$$

Операции наращения или эрозии тени от f с помощью тени от g дают тени от новых функций — от наращения или эрозии f с помощью g. Эти новые функции могут быть вычислены по прямым формулам

$$(f \oplus g)(x) = \max_{y}\{f(y) + g(x - y)\},$$

$$(f \quad g)(x) = \min_{y}\{f(y) - g(y - x)\},$$

где диапазоны x и y определяются как пересечение «опоры» (support) функции f и (смещенной на x) опоры функции g. Под опорой функции f понимается множество значений x, на которых $f(x) \neq -\infty$. Предполагается, что функция g имеет компактную опору и играет роль структурирующего элемента. Размыкание и замыкание f посредством g выражаются соответственно функциями

$$f \circ g = (f \quad g) \oplus g \quad \text{и} \; f \bullet g = (f \oplus g) \quad g.$$

Морфологические фильтры широко применяются при обработке и анализе изображений. Если W есть малый симметричный двумерный двоичный структурирующий элемент, тогда разность множеств $X \setminus (X \quad W)$ дает границу двоичного изображения X, а алгебраическая разность

$$EG(f) = f - (f \quad W). \tag{1}$$

Аналогичным улучшающим контуры оператором является градиент наращения

$$DG(f) = (f \oplus W) - f. \tag{2}$$

Комбинируя операторы (1) и (2), можно получить новые контурные операторы, которые обеспечивают более симметричную обработку изображения и его фона.

Устойчивость обнаружения контуров этими морфологическими контурными операторами может быть повышена путем предварительного сглаживания входного сигнала изображения f с помощью линейного смазывания либо с помощью фильтров с автоподстройкой.

Одной из фундаментальных проблем анализа изображений является создание адекватного математического описания изображений, передающего их содержание, смысл. Иными словами, это описание должно отражать лишь существенные (с точки зрения решаемой задачи) особенности изображения и не зависеть от несущественных деталей. В морфологическом анализе такими несущественными характеристиками являются условия регистрации изображений объекта или сцены (условия освещения, изменение оптических свойств объектов, изменение углов наблюдения) и параметры регистрирующей аппаратуры.

Важным классом задач, предшествующих решению задачи сравнения по форме предъявленных изображений, являются задачи совмещения и выделения заданных фрагментов. Под формой изображения понимается максимальный инвариант преобразований изображения, которым оно подвергается при изменении условий наблюдения, изменении параметров регистрирующей аппаратуры и др. [5].

Форма, таким образом, определяется не только исследуемым объектом или сценой, но и тесно связана с моделью регистрации изображения объекта или сцены. Один из способов построения формы состоит в задании областей постоянной яркости по физическим свойствам объекта, т.е. по расположению однородно светящихся или отражающих граней или границ относительно наблюдателя. Приписывая этим областям всевозможные яркости, получим форму изображения как множество изображений. Если же мы не имеем столь подробных сведений об объекте исследования, то можно построить форму по какому-либо одному изображению, зная, к каким преобразованиям яркости данного изображения могут привести изменившиеся условия наблюдения. Любое изображение, форма которого не сложнее f, можно получить подбором яркостей C_i:

$$f(x,y) = \sum_{i=0}^{n} C_i \chi_i(x,y).$$

Здесь C_i — яркость i-й точки; $\chi_i(x, y)$ индикаторная функция, которая определяется следующим образом:

$$\chi_i(x,y) = \begin{cases} 0 \notin A_i \\ 1 \in A_i \end{cases},$$

где A_i — множество точек с данной интенсивностью.

Таким образом, построение формы изображения заключается в определении индикаторных функций $\chi_i(x,y)$. В простейшем случае можно диапазон изменения яркостей изображения f разбить на n одинаковых интервалов.

На основании этих предположений можно предложить следующий алгоритм идентификации изображений.

Первоначально изображение преобразовывается в полутоновое с n-битной градацией яркости в диапазоне $(\overline{0,n})$. Путем последовательного применения морфологических операций осуществляются выделение заданного фрагмента изображения (сюжетной части) и приведение его к «универсальному» размеру. Для выбранного класса изображений по яркостям одного из изображений («эталон») вычисляются индикаторные функции путем попиксельного анализа всего изображения. В результате такой операции формируются индикаторные функции формы изображения для соответствующих яркостей.

Затем по индикаторным функциям изображения компенсируется разность яркостных характеристик, без нарушения формы анализируемого изображения («претендента»). Для этого для каждой индикаторной функции анализируемого изображения вычисляется корректирующий коэффициент, так как сравнение может производиться лишь тогда, когда условия получения были максимально идентичны.

Такого результата можно достичь применением метода наименьших квадратов (МНК). Для этого яркость анализируемого изображения приводится к интенсивности «эталона» так, чтобы сумма квадратов отклонения для каждой индикаторной функции была минимальна:

$$\sum_{i=0}^{m} (C_i - K_j C_i^*)^2 \to \min,\ j = (\overline{0,n}),$$

где C_i — яркость эталона; C_i^* — яркость i-й точки претендента; K_j — корректирующий коэффициент; m — количество точек для каждой индикаторной функции.

В результате получается приведенное изображение без искажения формы, которое можем сравнивать с базовым изображением путем попиксельного вычитания в каждой точке i, k:

$$\left| C_{ik} - C_{ik}^{**} \right| = A_{ik},$$

где A_{ik} — разница между изображениями; C_{ik} — яркость эталонного изображения; C_{ik}^{**} — яркость приведенного изображения.

В случае идентичных изображений разница между ними будет равна нулю. При сравнении реальных изображений разница может быть не нулевая. Для оценки остаточного изображения используется моментный анализ [6], основанный на фундаментальной теореме: «Бесконечная последовательность моментов изображения $\{m_{\alpha\beta}\}$ или $\{\mu_{\alpha\beta}\}$ однозначно определяется функцией $f(x, y)$, и, наоборот, функция $f(x, y)$ однозначно определяется последовательностью моментов изображения $\{m_{\alpha\beta}\}$ или $\{\mu_{\alpha\beta}\}$ $(\alpha, \beta) = 0, 1, \ldots$.» [7]:

$$m_{\alpha\beta} = \iint_{\infty} f(x,y) x^{\alpha} y^{\beta} dxdy$$

— собственный момент;

$$\mu_{\alpha\beta} = \iint_{\infty} f(x,y)(x - x_0)^{\alpha}(y - y_0)^{\beta} dxdy$$

— центральный момент.

В результате вычислений получается последовательность моментных характеристик. Для набора характеристик всех эталонов строится опорная гиперплоскость

$$\sum_{i=1}^{n} a_i m_i + a_{n+1} = 0,$$

где m_i — моментные характеристики; a_i — коэффициенты гиперплоскости;

$$m_k = \sum_{i=1 (i \neq k)}^{n} b_i m_i + b_{n+1},$$

$$b_i = \frac{a_i}{a_k}, \quad i = 1,\ldots,n+1\,.$$

Для вычисления коэффициентов опорной гиперплоскости данного класса (сюжета) изображений потребуем, чтобы моментные характеристики, вычисленные через коэффициенты гиперплоскости, минимально отклонялись от реально вычисленных моментных характеристик для всех изображений данного класса:

$$L(\mathbf{b}) = \sum_{j=1}^{m}\left[m_k{}^j - \left(\sum_{i=1(i\neq k)}^{n} b_i m_i^j + b_{n+1}\right)\right]^2 \to \min,$$

где $\mathbf{b} = (b_1,\ldots,b_{n+1})$, $i \neq k$.

Для оценки этих коэффициентов гиперплоскости строится система линейных уравнений

$$\frac{\partial L(\mathbf{b})}{\partial b_m} = 0, \quad m = 1,\ldots,n(m \neq k),$$

из решения которой находим b_i для данного класса изображений.

На основании этих вычислений строится база данных векторов опорной гиперплоскости для различных классов изображений.

Для изображения, пришедшего с видеодатчика (претендента), последовательно выбираются изображения из базы данных. Для этих двух изображений вычисляются их приведенная разность и моментные характеристики. По имеющимся в базе данных коэффициентам опорной гиперплоскости оценивается разность между реально вычисленной характеристикой и полученной по коэффициентам гиперплоскости:

$$m_{k_{re}}^l = \sum_{i=1}^{n} b_i m_i^l + b_{n+1}(i \neq k), \quad l = 1,\ldots,N,$$

где $m_{k_{re}}^l$ — реально вычисленная моментная характеристика для l-го эталона; N — количество эталонов в базе; l — текущий эталон.

Фиксируется тот эталон, у которого разница с текущим изображением была минимальна:

$$\left| m_{k_{re}} - m_{k_{gip}} \right| = \Delta,$$

где $m_{k_{gip}}$ — моментная характеристика, вычисленная по коэффициентам гиперплоскости.

Полученное значение Δ сравнивается с допустимым отклонением для данного класса изображений. Если оно меньше допустимого отклонения, то претендент принадлежит к данному классу. В противном случае претендент не принадлежит ни к одному из классов (*ЧУЖОЙ*). В качестве допустимого значения можно выбрать оценку 3σ, где σ — среднеквадратическое отклонение для данного класса изображений.

Предложенный метод позволяет сравнивать изображения, классифицировать по определенным признакам.

Основное достоинство предложенного метода — выполнение коррекции изображения без потери информации о форме. С помощью данного метода могут быть построены технические устройства, применимые в различных областях техники, в медицине, науке и др.

1. *Brigham E.O.* The Fast Fourier Transform. NY: Englewood Cliffs, 1974. 62 c.
2. *Gonzalez R.C. and Wintz P.* Digital Image Processing. Mass.: Addison-Wesley, 1987. 14 c.
3. *Serra J.* Image Analysis and Mathematical Morphology. Academic Press. ISBN: 0-12-637240-3.
4. *Пытьев Ю.П.* Морфологический анализ изображений // Докл. АН СССР. 1983. Т. 269. № 5. С. 1061—1064.
5. *Пытьев Ю.П.* Задачи морфологического анализа изображений // Математические методы исследования природных ресурсов Земли из космоса / Под ред. В.Г. Золотухина. М.: Наука, 1984. С. 41—83.
6. *Kalaida V.T., Belov V.V., Esipova V.A., Klimkin V.M.* Physical and Mathematical Methods for the Visualization and Identification of Watermarks // Solanus. Published by the School of Slavonic and East European Studies (University of London). Typeset in Plantin and Times Cyrillic at Oxford University Computing Service. 1999. № 13. C. 80 — 92.
7. *Анисимов Б.В., Курганов В.Д., Злобин В.К.* Распознавание и цифровая обработка изображений. М.: Высш. шк., 1983. 295 с.
8. *Маракос П., Шафер Р.В.,* Морфологические системы для обработки многомерных сигналов // ТИИЭР. 1990. Т. 78. № 4.

V.T. Kalaida. **Application of morphological analysis method to identification of gray scale images.**

The paper proposes a modified method of morphological analysis for identification of gray scale images and a hardware/software image recognition system based on the proposed algorithm.