

СРАВНИТЕЛЬНЫЙ АНАЛИЗ СТРУКТУРНЫХ ПРЕДСТАВЛЕНИЙ ИЗОБРАЖЕНИЙ НА ОСНОВЕ ПРИНЦИПА РЕПРЕЗЕНТАЦИОННОЙ МИНИМАЛЬНОЙ ДЛИНЫ ОПИСАНИЯ

© 2008 г. А. С. Потапов, канд. техн. наук

НПК “Государственный оптический институт им. С.И. Вавилова”, Санкт-Петербург

E-mail: apotapov@mail.wplus.net

На основе ранее предложенного принципа репрезентационной минимальной длины описания (РМДО), предназначенного для количественной оценки степени инвариантности представлений изображений, произведен сравнительный анализ нескольких алгоритмов сегментации, строящих контурные описания изображений, а также нескольких алгоритмов построения структурных описаний изображений с использованием различных алфавитов структурных элементов. Обоснована адекватность использования принципа РМДО при сравнении инвариантности представлений изображений. Установлено, что оптимальные алгоритмы сегментации различаются для разных выборок изображений (например, аэрокосмических или полученных внутри помещений). На основе объективного критерия впервые строго обоснована целесообразность использования в качестве структурных элементов отрезков прямых линий и сегментов кривых второго порядка и низкая эффективность кривых более высоких порядков.

Коды OCIS: 3000.32960.

Поступила в редакцию 02.06.2008.

Введение

Одной из наиболее важных характеристик автоматических методов анализа изображений является степень их инвариантности по отношению к возможным факторам изменчивости изображений, таким как изменение освещения, ракурса съемки или типа сенсора, сезонно-суточной и прочей изменчивости объектов сцены [1]. Степень инвариантности определяет вероятность получения корректного решения задачи и широту сферы применения того или иного метода. В предметных областях с высокой априорной неопределенностью до сих пор не существует методов анализа изображений, которые бы обладали значительной степенью инвариантности. Более того, на настоящий момент не существует подходов в области анализа изображений, которые бы позволяли направленно улучшать данную характеристику методов анализа. Разработка подобного подхода является весьма актуальной проблемой, решение которой может иметь большое практическое значение.

Из-за большого многообразия задач анализа изображений введение общего критерия, на основе которого можно было бы оценивать их степень инвариантности, является проблематичным. Однако любой метод анализа использует некоторое представление изображений, под которым обычно понимается формальная система, содержащая алгоритмы для получения в явном виде описаний объектов заданного класса [2]. Именно степень инвариантности привлекаемых представлений изображений

преимущественно определяет степень инвариантности соответствующих методов анализа.

Ранее автором был предложен принцип репрезентационной минимальной длины описания [3], повышение эффективности представлений на основе которого соответствует повышению степени их инвариантности. Данный принцип обобщает принцип минимальной длины описания [4] и состоит из двух частей:

Моделью, наилучшим образом описывающей некоторое изображение в рамках заданного представления, является модель, минимизирующая сумму длины описания модели и длины описания изображения в рамках модели.

Лучшим представлением для данной выборки изображений является представление, для которого минимизируется сумма длины представления, суммы длин оптимальных описаний изображений в рамках представления.

Критерий РМДО может быть вычислен для некоторого представления по выборке изображений вне зависимости от последующей задачи анализа. Это позволяет сравнивать представления, используемые в различных методах, не опираясь на конечный результат (например, распознавания, измерения, выявления изменений и т. д.), на который также влияют следующие за этапом построения описания изображений этапы анализа.

Поскольку проблема инвариантности наиболее заметно проявляется в задачах с существенной априорной неопределенностью, в настоящей работе производится сравнение структурных представле-

ний, являющихся здесь наиболее предпочтительными [1, 5]. Рассматриваются как алгоритмы сегментации изображений, в которых изображение представляется как совокупность независимо описываемых областей, так и алгоритмы построения структурных элементов на основе контуров. Показывается, что результаты сравнения представлений на основе критерия РМДО не только согласуются с опытом исследователей, но также позволяют получить новую информацию об эффективности различных представлений.

Представления изображений в алгоритмах сегментации

В данной работе рассматриваются следующие представления, составляющие основу одного семейства алгоритмов сегментации изображений.

1. Представление $S_0^{(1)}$, в котором интенсивности пикселей полагаются отсчетами статистически независимых и одинаково распределенных случайных величин. Длина описания изображения $f(x, y): G \rightarrow R$ в рамках представления $S_0^{(1)}$ может быть оценена как

$$L_{S_0}(f) = \|G\|H(f) + N_{\text{int}} \log_2 N_{\text{int}}, \quad (1)$$

где $\|G\|$ – площадь области G , выражаемая в данном случае как количество пикселей в изображении f , $H(f)$ – энтропия значений интенсивностей в предположении об их статистической независимости, N_{int} – число различных уровней интенсивности. Величина $\|G\|H(f)$ является оценкой математического ожидания длины описания интенсивностей для пикселей области G , зашифрованных с помощью кода Хаффмана. Для расшифровки такого кода требуется таблица перекодировки, которая должна содержаться в описании изображений. Ее объем можно оценить как $N_{\text{int}} \log_2 N_{\text{int}}$ бит.

2. Представление $S_1^{(1)}$, в котором изображение полагается разбитым на совокупность областей G_1, \dots, G_d . В каждой области интенсивности пикселей имеют собственное распределение. Помимо интенсивностей пикселей здесь также необходимо описать границы областей δG_i . Таким образом, длина описания изображения в рамках данного представления может быть оценена как

$$L_{S_1^{(1)}}(f) = \sum_i (\|G_i\|H(f_i) + N_{\text{int}} \log_2 N_{\text{int}} + \|\delta G_i\| \log_2 N_{\text{dir}}), \quad (2)$$

где $f_i(x, y) = f(x, y)|_{G_i}$ – сужение изображения f на область G_i с длиной границы $\|\delta G_i\|$, N_{dir} – число возможных направлений от текущей точки границы на следующую точку (здесь использовано $N_{\text{dir}} = 8$).

Критерий (2) служит для выбора разделения изображения на области (описание алгоритма представлено в работе [6]).

3. Представление $S_2^{(1)}$, в котором изображение также полагается разделенным на совокупность областей, но из значений $f_i(x, y)$ в каждой области вычитена функция $g_i(x, y, \mathbf{w}_i)$, задаваемая в виде

$$g(x, y, \mathbf{w}) = w_0 + w_1 x + w_2 y + w_3 x^2 + w_4 xy + w_5 y^2. \quad (3)$$

Невязки в каждой области $r_i(x, y) = [f_i(x, y) - (x, y, \mathbf{w}_i)]$ полагаются статистически независимыми и обладающими фиксированным распределением в каждой области. Длина описания в рамках представления $S_2^{(1)}$ может быть записана в виде

$$L_{S_2^{(1)}}(f) = \sum_{i=1}^d [\|G_i\|H(r_i) + N_{\text{int}} \log_2 N_{\text{int}} + \|\delta G_i\| \log_2 N_{\text{dir}} + l(\mathbf{w}_i)], \quad (4)$$

где $l(\mathbf{w}_i) = \frac{m_i}{2} \log_2 \|G_i\|$ – оценка длины описания параметров функции $g_i(x, y, \mathbf{w}_i)$ в i -ой области с размерностью вектора параметров $m_i = 6$.

4. Представление $S_3^{(1)}$, которое аналогично представлению $S_2^{(1)}$, но в котором вместо функций вида (3) из значений $f_i(x, y)$ в каждой области вычитается некоторая совокупность функций Габора (более подробно данное представление описано в работе [7]). Длина описания изображения в рамках этого представления также описывается уравнением (4) за тем исключением, что в длине описания параметров

$l(\mathbf{w}_i) = \frac{m_i}{2} \log_2 \|G_i\|$ число параметров m_i зависит от числа функций Габора, вычитаемых из значений $f_i(x, y)$ в области G_i .

Представления изображений в алгоритмах построения структурных элементов

Структурные элементы могут быть построены на основе контуров, выделенных, как границы областей, полученных в рамках рассмотренных выше представлений в алгоритмах сегментации, или выделенных каким-либо другим способом. Рассмотрим следующие представления для описания единственного контура: $\delta G = \{(x_i, y_i)\}_{i=1}^N$, где $N = \|\delta G\|$ – общее число точек на данном контуре. В простейшем представлении контура, которое обозначим через $S_0^{(2)}$, используется цепное кодирование, а длина описания контура может быть оценена как

$$L_{S_0^{(2)}}(\delta G) = \|\delta G\| \log_2 N_{\text{dir}}. \quad (5)$$

Отметим, что именно это упрощенное представление использовалось для оценки длины описания

границ областей в уравнениях (2) и (4). Привлечение более точных оценок, приведенных ниже, подразумевает необходимость построения структурных элементов в процессе сегментации изображений (о проблеме построения иерархических описаний см. [8]).

Обозначим через $S_1^{(2)}$ представление, в котором контур описывается как совокупность сегментов, для каждого из которых проводится отрезок прямой с минимальной энтропией невязок (отклонений отрезка от точек контура). Через $S_2^{(2)}$ обозначим представление, в котором помимо отрезков прямых используются в качестве структурных элементов дуги окружностей и кривые второго порядка, а через $S_3^{(2)}$ – представление, в котором используются также и кривые третьего порядка. Для всех этих представлений критерий РМДО можно представить единообразно.

Рассмотрим вначале один k -й сегмент. Пусть он состоит из $N^{(k)}$ точек и описывается кривой с $m_p^{(k)}$ параметрами, дающей невязки $r_i^{(k)}$, $i = 1, \dots, N^{(k)}$. Оценка длины описания сегмента состоит из оценок длин описания модели (структурного элемента, который обозначим через s_k) и данных в рамках модели (невязок, с которыми структурный элемент описывает сегмент контура). Длина описания структурного элемента состоит из

числа битов (которое обозначено через b), необходимых для указания типа структурного элемента, длины описания параметров структурного элемента $\frac{m_p^{(k)}}{2} \log_2 N^{(k)}$.

Итак, длина описания одного структурного элемента может быть оценена как $b + \frac{m_p^{(k)}}{2} \log_2 N^{(k)}$.

Длина описания данных в рамках модели соответствует длине описания невязок $r_i^{(k)}$ – значений отклонения структурного элемента от точек контура. Считая невязки независимыми отсчетами некоторой случайной величины r_k , эту длину описания невязок можно оценить как длину кода Хаффмана $N^{(k)} H(r_k)$, где $H(r_k) = -\sum_r P_k(r) \log_2 P_k(r)$ – оценка энтропии невязок по их гистограмме $P_k(r)$. Помимо передачи невязок для декодирования необходима таблица перекодировки (или гистограмма $P_k(r)$), длину которой можно грубо оценить как $n_r \log_2 n_r$, где n_r – количество различных значений невязок для данного сегмента.

Длина описания одного сегмента составит

$$\begin{aligned} L_k &= K_{S_0^{(3)}}(s_k) + K_{S_n^{(2)}}(\delta G | s_k) = \\ &= b + \frac{m_p^{(k)}}{2} \log_2 N^{(k)} + N^{(k)} H(r_k) + n_r \log_2 n_r. \end{aligned}$$

Тогда общая целевая функция будет записываться в форме

$$L_{S_n^{(2)}}(\delta G) = \sum_{k=1}^d L_k, \quad (6)$$

где от номера представления n зависит величина b , а также выбор конкретного числа сегментов d и типов структурных элементов для каждого из них. Алгоритм сегментации контуров, минимизирующий длину описания (6), представлен в работе [9].

Сравнение эффективности представлений

Рассмотрим три предметные области с высокой априорной неопределенностью, которые характеризуются выборками изображений F_1 , F_2 и F_3 , представленными на рис. 1 и содержащими аэрокосмические, радиолокационные изображения и изображения, полученные внутри помещений соответственно. Сравним эффективности приведенных выше представлений на каждой из трех предметных областей.

В табл. 1 приведены отношения средних длин описания изображений разных выборок в рамках представлений $S_n^{(1)}$, используемых при сегментации, а также доля изображений выборок, для которых происходит уменьшение длины описания при смене представления.

Как видно из табл. 1, только для первой выборки (аэрокосмические изображения) представление $S_3^{(1)}$ превосходит представление $S_1^{(1)}$ на всех изображениях, и только для третьей выборки (изображения внутри помещений) представление $S_2^{(1)}$ превосходит представление $S_1^{(1)}$ на всех изображениях. В то же время описание изображений в виде совокупности областей оказывается всегда эффективнее его описания без такого разделения. Результат, представленный в табл. 1, согласуется с эвристическими оценками содержания изображений, а именно:

Сцены внутри помещений содержат большое число областей с плавными переходами яркости (рис. 2), которые при использовании представления $S_1^{(1)}$ разделяются на множество областей по уровням освещенности. В рамках представления $S_2^{(1)}$ в качестве дополнительных областей выделяются лишь области теней, а не плавного изменения освещенности. В то же время аэрокосмические изображения могут также содержать плавные переходы яркости, однако для них это является менее характерным, чем для изображений, полученных внутри помещений.

Искусственные текстуры, характерные для изображений, полученных внутри помещений, недостаточно эффективно описываются функциями Габо-



Рис. 1. Примеры фрагментов изображений из выборок а) аэрокосмических изображений, б) радиолокационных изображений, в) изображений, полученных внутри помещений.

ра в рамках представления $S_3^{(1)}$, тогда как для естественных текстур, присутствующих на аэрокосмических изображениях, данное представление оказывается более эффективным.

Радиолокационные изображения, как правило, не содержат ни плавных переходов, ни текстур, хорошо описываемых функциями Габо́ра.

Представление изображений $S_1^{(1)}$, основанное на моделях сегментации, позволяет осуществлять выделение однородных областей на изображении под управлением информационной целевой функции. Эти области соответствуют границам видимых поверхностей, однако из-за сильного упрощения представления такое соответствие может нарушаться.

Таблица 1. Сравнение эффективности представлений $S_n^{(1)}$

Отношение длин описания	Выборка		
	F_1	F_2	F_3
$\overline{L_{S_1^{(1)}}(f)} / \overline{L_{S_0^{(1)}}(f)}$	0,837	0,968	0,713
$\overline{L_{S_2^{(1)}}(f)} / \overline{L_{S_1^{(1)}}(f)}$	0,985	0,999	0,921
$\overline{L_{S_3^{(1)}}(f)} / \overline{L_{S_1^{(1)}}(f)}$	0,946	0,988	0,996
$P(L_{S_1^{(1)}}(f) < L_{S_0^{(1)}}(f))$	100%	100%	100%
$P(L_{S_2^{(1)}}(f) < L_{S_1^{(1)}}(f))$	57%	1%	100%
$P(L_{S_3^{(1)}}(f) < L_{S_1^{(1)}}(f))$	100%	63%	28%

Усложнение представления позволяет улучшить результаты сегментации, однако не любое усложнение приводит к повышению качества сегментации, что находит непосредственное отражение в средней длине описания изображений выбранной предметной области. Таким образом, критерий РМДО может быть использован для оценки эффективности представлений в задаче сегментации.

Сравним эффективности приведенных выше представлений контуров (границ областей) с помощью структурных элементов. В табл. 2 приведены средние длины описаний границ областей δG в рам-

Таблица 2. Сравнение эффективности представлений $S_n^{(2)}$

Отношение длин описания	Выборка		
	F_1	F_2	F_3
$\overline{L_{S_1^{(2)}}(\delta G)} / \overline{L_{S_0^{(2)}}(\delta G)}$	0,809	0,812	0,679
$\overline{L_{S_2^{(2)}}(\delta G)} / \overline{L_{S_1^{(2)}}(\delta G)}$	0,831	0,845	0,791
$\overline{L_{S_3^{(2)}}(\delta G)} / \overline{L_{S_2^{(2)}}(\delta G)}$	1,007	1,007	1,006

ках представлений $S_n^{(2)}$ на разных выборках изображений.

Из табл. 2 видно, что при введении в представление контуров в качестве структурных элементов, как прямых линий, так и сегментов кривых второго порядка происходит заметное уменьшение средней длины описания. При этом представление $S_1^{(2)}$ оказывается эффективнее представления $S_0^{(2)}$ на всех изображениях выборок, а представление $S_2^{(2)}$ эффективнее представления $S_1^{(2)}$. Иными словами, в качестве структурных элементов обязательно должны использоваться как отрезки прямых линий, так и сегменты кривых второго порядка. Напротив, введение структурных элементов, соответствующих кривым более высоких степеней, не улучшает критерий РМДО ни для одной из выбранных предметных областей.

Относительное качество представлений $S_1^{(2)}$ и $S_2^{(2)}$ близко для каждой из выборок, однако для изобра-

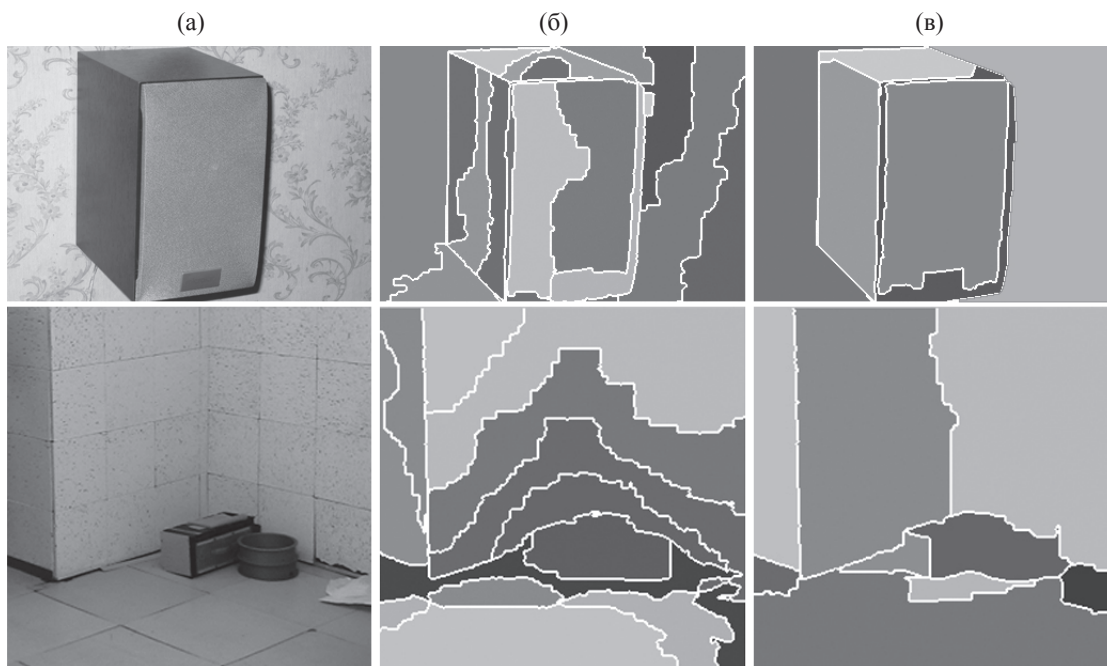


Рис. 2. Примеры сегментации фрагментов изображений (а), полученных внутри помещений, с помощью представлений $S_1^{(1)}$ (б) и $S_2^{(1)}$ (в).

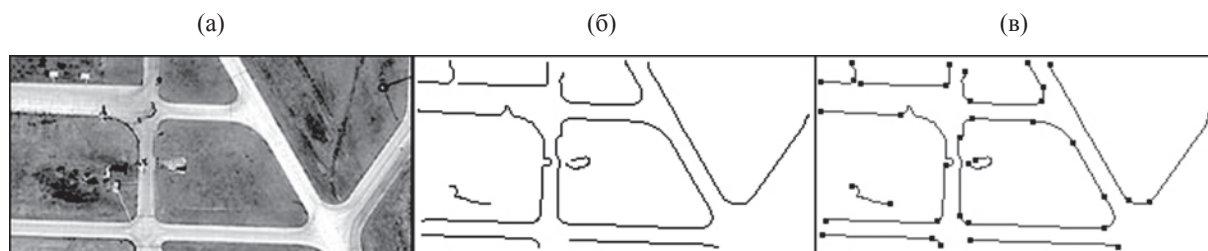


Рис. 3. Сегментация контуров, полученных с помощью фильтрации Дерриша при некотором уровне сглаживания.

жений, полученных внутри помещений, длина описания контуров в рамках этих представлений уменьшается сильнее, то есть контуры на изображениях этого типа обладают большей регулярностью.

Вывод о том, что для реальных изображений использование в качестве структурных элементов кривых третьей и более высоких степеней является неэффективным, относится только к алгоритму сегментации изображений, используемому рассмотренные выше представления. Чтобы распространить этот вывод на структурные элементы, строящиеся на основе контуров, получаемых с помощью других алгоритмов, были определены средние длины описаний контуров, выделенных на изображениях выборки $F_1 \cup F_2 \cup F_3$ с помощью фильтра Дерриша [10] (пример контуров, построенных с помощью этого метода, и результаты сегментации этих контуров приведены на рис. 3). При этом получены следующие значения:

$$\frac{\overline{L_{S_1^{(2)}}(\delta G)}}{\overline{L_{S_0^{(2)}}(\delta G)}} = 0,833,$$

$$\frac{\overline{L_{S_2^{(2)}}(\delta G)}}{\overline{L_{S_1^{(2)}}(\delta G)}} = 0,798,$$

$$\frac{\overline{L_{S_3^{(2)}}(\delta G)}}{\overline{L_{S_2^{(2)}}(\delta G)}} = 1,008.$$

Таким образом, относительные длины описания получились достаточно близкими для контуров, выделенных с помощью фильтра Дерриша и с помощью алгоритма сегментации с использованием описанных выше представлений. Однако, как оказывается, при применении фильтра Дерриша несколько большее предпочтение отдается кривым второго порядка, что может быть объяснено особенностями применяемого в нем фильтра, осуществляющего сглаживание перепадов яркости.

Заключение

Предложенный принцип репрезентационной минимальной длины описания позволил произвести сравнение нескольких представлений изображе-

ний на выборках изображений с большой априорной неопределенностью. В результате сравнения обнаружено, что представление, учитывающее возможность плавного изменения яркости пикселей области, более эффективно для изображений, полученных внутри помещений. В то же время представление, в котором привлекаются функции Габора для описания содержимого областей, более эффективно для аэрокосмических изображений.

Обнаружена также высокая эффективность отрезков прямых линий и сегментов кривых второго порядка в качестве структурных элементов, аппроксимирующих контуры на изображениях трех рассмотренных выборок. Кривые более высоких порядков оказываются малоэффективными, то есть не должны использоваться при построении инвариантных структурных описаний изображений.

Таким образом, принцип РМДО может быть использован для обоснованного выбора представлений изображений. В работе рассмотрены несколько частных представлений, которые требуют дальнейшего совершенствования, что может быть выполнено в рамках предложенного подхода.

ЛИТЕРАТУРА

1. Lutsiv V.R., Malyshev I.A., Pepelka V., Potapov A.S. Target independent algorithms for description and structural matching of aerospace photographs // Proc. SPIE. 2002. V. 4741. P. 351–362.
2. Март Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов: М.: Радио и связь, 1987. 36 с.
3. Потанов А.С. Исследование представлений изображений на основе принципа репрезентационной длины описания // Изв. вузов. Приборостроение. 2008. Т. 51. № 7. С. 3–7.
4. Vitanyi P.M.B., Li M. Minimum description length induction, Bayesianism, and Kolmogorov complexity // IEEE Transactions on Information Theory. 2000. V. 46. № 2. P. 446–464.

5. *Rares A., Reinders M.J.T., Hendriks E.A.* Image Interpretation Systems // Technical Report (MCCWS 2.1.1.3.C), MCCWS project, Information and Communication Theory Group. TU Delft. 1999. 32 p.
 6. *Potapov A.S.* Image matching with the use of the minimum description length approach // Proc. SPIE. 2004. V. 5426. P. 164–175.
 7. *Потанов А.С.* Теоретико-информационный подход к введению обратных связей в многоуровневых системах компьютерного зрения // Оптический журнал. Т. 74. № 10. 2007. С. 59–65.
 8. *Потанов А.С., Гуров И.П., Аверкин А.Н.* Анализ модели клеток зрительной коры на основе принципа репрезентационной минимальной длины описания // Изв. вузов. Приборостроение. 2008. (в печати).
 9. *Potapov A.S., Gamayunova O.S.* Information criterion for constructing the hierarchical structural representations of images // Proc. SPIE. 2005. V. 5807. P. 443–454.
 10. *Deriche R.* Optimal edge detection using recursive filtering // Proc. 1st Int. Conf. Computer Vision. 1987. P. 501–505.
-