

УДК 004.932.2

МЕТОД ПРЕДСКАЗАНИЯ НА ОСНОВЕ АЛГОРИТМИЧЕСКОЙ ВЕРОЯТНОСТИ В ЗАДАЧЕ ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЙ В УТЕРЯННЫХ ОБЛАСТЯХ

© 2013 г. А. С. Потапов*, **, доктор техн. наук; О. В. Щербаков**, аспирант;
И. Н. Жданов**, аспирант

* ОАО “Государственный оптический институт им. С.И. Вавилова”

** Национальный исследовательский университет информационных технологий, механики и оптики, Санкт-Петербург

E-mail: pas.aicv@gmail.com

Рассмотрен критерий алгоритмической вероятности, предоставляющий общее решение задачи экстраполяции символьных строк. Данный критерий расширяет теоретико-информационный подход на основе алгоритмической сложности, широко использующийся при синтезе методов анализа изображений. Даны методические рекомендации о практическом применении критерия алгоритмической вероятности при обработке и анализе изображений. В качестве примера проведена разработка метода восстановления изображений в утраченных областях, вычислимость которого достигнута путем сужения алгоритмически полного пространства моделей изображений до множества их Фурье-образов.

Ключевые слова: алгоритмическая вероятность, восстановление изображений, преобразование Фурье.

Коды OCIS: 150.1135.

Поступила в редакцию 27.05.2013.

Введение

Теория информации является мощным средством для формализации понятий иконоки, видеоинформатики и смежных областей, в которых рассматриваются вопросы формирования, представления, обработки и передачи изображений как информационных объектов. Особый интерес представляет использование теории универсальной индукции на основе алгоритмической теории информации, в которой количество информации вводится через алгоритмическую (колмогоровскую) сложность. Данная теория предоставляет общие принципы, такие как принцип минимальной длины описания и критерии идентификации произвольных (вычислимых) моделей источника информации. В то же время, основная часть задач автоматического анализа изображений

и компьютерного зрения включает идентификацию моделей. Практические методики использования принципа минимальной длины описания нашли широкое распространение при решении задач сегментации изображений [1], построения структурных элементов [2], описания формы границ областей изображений [3], распознавания объектов на изображении [4], распознавания рукописных символов [5], оценки параметров пространственного преобразования между парой изображений [6], оценки поля движения по видеоизображениям [7] и многих других задач.

В то же время, адекватность самого индуктивного вывода, подразумевающего определение и использование одной лучшей модели, ставится под сомнение [8]. Действительно, при индуктивном выводе принципиально нельзя определить достоверно истинную модель источ-

ника информации, а под лучшей моделью понимается лишь наиболее вероятная. В теории универсального предсказания на основе понятия алгоритмической вероятности P . Соломонова [8] постулируется, что для оптимального предсказания необходимо учитывать все возможные модели. Однако различия между индукцией и предсказанием могут оказаться существенными лишь для систем общего искусственного интеллекта, тогда как для узкоспециализированных методов взвешивание по ансамблю моделей лишь незначительно повышает точность предсказания по сравнению с использованием одной лучшей модели. Тем не менее, весьма интересен тот факт, что практические методики использования идей универсальной индукции в анализе изображений весьма распространены, тогда как ни одной попытки разработки аналогичных методик для практического использования идей универсального предсказания, насколько нам известно, нет.

В рамках данной работы делается попытка разработки методики практического использования теории экстраполяции символьных строк на основе алгоритмической вероятности, что достигается за счет сужения алгоритмически полного пространства моделей источников информации до некоторого пространства моделей, задаваемого неуниверсальной опорной машиной, которой соответствует частное представление информации. Возможность подобного применения предсказания по критерию алгоритмической вероятности демонстрируется на примере задачи восстановления изображения в утерянных областях.

Предсказание на основе алгоритмической вероятности

В алгоритмической теории информации А.Н. Колмогорова индивидуальная сложность $K(\alpha)$ некоторой бинарной строки α , или количество содержащейся в ней информации, определяется как длина $l(\mu)$ наикратчайшей программы μ , порождающей данную строку при выполнении на некоторой опорной машине U : $U(\mu) = \alpha$. В рамках проблематики индуктивного вывода данная программа трактуется как наилучшая модель источника строки. Хорошо известна связь между количеством информации I и вероятностью P : $P = 2^{-I}$. Кажется естественным приписать строке α вероятность $P(\mu) = 2^{-K(\alpha)}$, однако более обоснованным [8] является рассмотрение $2^{-l(\mu)}$ в качестве априорной вероятности про-

граммы μ , а с учетом того, что строка α может быть порождена разными программами, ее (алгоритмическую) вероятность P_U следует рассчитывать в соответствии с выражением

$$P_U(\alpha) = \sum_{\mu: U(\mu) = \alpha} 2^{-l(\mu)}, \quad (1)$$

где индекс U указывает на зависимость этого распределения от опорной машины.

Данное определение может быть использовано в качестве основы теории универсального предсказания [9], для чего необходимо ввести следующее определение условной вероятности того, что в качестве продолжения строки α будет наблюдаться строка β

$$P_U(\beta | \alpha) = P_U(\alpha\beta) / P_U(\alpha), \quad (2)$$

где $\alpha\beta$ – конкатенация двух строк.

Следует отметить, что при использовании алгоритмической вероятности в контексте предсказания равенство $U(\mu) = \alpha$ следует заменить проверкой того, что программа μ порождает некоторую строку, для которой α является префиксом (т.е. программа печатает α с произвольным продолжением). Такая интерпретация, очевидно, обеспечит выполнение неравенства $P_U(\alpha\beta) \leq P_U(\alpha)$, и вероятность конкатенации $P_U(\alpha\beta)$ будет обладать свойствами классической совместной вероятности, а вероятность продолжения $P_U(\beta | \alpha)$ – свойствами условной вероятности.

Алгоритмическая вероятность является универсальной в силу учета всех возможных вычислимых закономерностей, которые могут содержаться в наблюдаемой последовательности данных. Однако по этой же причине она оказывается и неприменимой на практике. Тем не менее, и колмогоровская сложность является невычислимой (в силу алгоритмической неразрешимости проблемы останова), но методологические следствия из нее как критерия качества модели имеют не только важное теоретическое значение, но и, как отмечалось выше, большую практическую пользу. Такую же пользу можно извлечь и из аналогичного использования алгоритмической вероятности.

Практическая алгоритмическая вероятность

Применение алгоритмической сложности на практике достигается за счет сужения пространства моделей и использования эвристиче-

ских схем кодирования, что более строго можно рассматривать как замену универсальной опорной машины U на неуниверсальную S (не полную по Тьюрингу). Более того, можно ввести дополнительный уровень индуктивного вывода и рассматривать эмуляцию неуниверсальной машины S , играющей роль представления информации, на универсальной машине U , тогда можно разделить две задачи [10]:

- задачу поиска оптимальной модели μ данных α в рамках представления S :

$$\mu^* = \arg \min_{\mu: U(S\mu)=S(\mu)=\alpha} I(\mu).$$

- задачу поиска оптимального представления S для некоторой выборки строк $\alpha_1 \dots \alpha_n$:

$$S^* = \arg \min_S \sum_{i=1}^n \min_{S(\mu)=\alpha_i} I(\mu).$$

Большой интерес представляют методы автоматической оптимизации представлений, что пока выполняется только для узких классов представлений, а также автоматические методы построения эффективных алгоритмов поиска моделей в рамках заданных представлений (т.е. методы метавычислений, примененные к специализации алгоритма универсальной индукции представлениями, что пока на практике, насколько нам известно, не осуществлялось). Однако между идеей создания автоматических методов метаиндукции и возможностью их практического использования в настоящее время слишком большой разрыв, и этот процесс метаиндукции неявно выполняется разработчиками соответствующих прикладных методов. Здесь мы не будем касаться вопроса автоматического синтеза оптимальных специализированных методов, а выполним подобную специализацию вручную для одного наперед заданного представления при решении одной задачи из области обработки изображений.

Восстановление изображения в утерянной области

Одной из задач в области обработки изображений, для которой непосредственным образом можно сформулировать критерий алгоритмической вероятности (2), является задача восстановления изображения в утерянной области.

Пусть дано изображение $f(x, y)$, которое должно быть реконструировано по некоторой области R на основе яркостей в остальной части изображения R' . Будем обозначать изображения в соответствующих областях через $f|_R$ и $f|_{R'}$. Через $f|_{R'} g|_R$ будем обозначать “конкатенацию” двух изображений:

$$(f|_{R'} g|_R)(x, y) = \begin{cases} f(x, y), & (x, y) \notin R, \\ g(x, y), & (x, y) \in R. \end{cases}$$

Тогда задача определения наилучшего заполнения изображения в области R может быть сведена к задаче оптимального предсказания:

$$g^*|_R = \arg \max_g P_U(g|_R | f|_{R'}) = \arg \max_g P_U(f|_{R'} g|_R). \quad (3)$$

Хотя принцип минимальной длины описания здесь напрямую не применим, причина этого заключается не в том, что в индукции строится одна лучшая модель, а в том, что в ней модель строится только по имеющимся данным, и эта модель может не осуществлять никакого предсказания. Предсказание также можно пробовать выполнять на основе одной лучшей модели, но при этом нужно будет найти такое заполнение области R , при котором длина описания всего изображения в совокупности с реконструированной областью окажется минимальной.

Может показаться сомнительным, что заполнение изображения в утерянной области какой-то “содержательной” информацией может уменьшить длину описания всего изображения по сравнению, скажем, с его заполнением просто однородным фоном. Однако это действительно так, если мы не ограничиваемся какими-то простыми способами сжатия изображений. Поясним это на простом примере. На рис. 1 представлено исходное изображение, составленное из наложенных друг на друга прямоугольников разной формы, и это же изображение с удаленной областью. Очевидно, самое простое (то есть самое вероятное) описание данного изображения должно состоять просто из указания координат углов и яркостей накладывающихся прямоугольников, например, отсортированных по “дальности”. Рассмотрим изображение с вырезанной областью. Действительное содержание этой области не вполне тривиально и включает пять подобра-

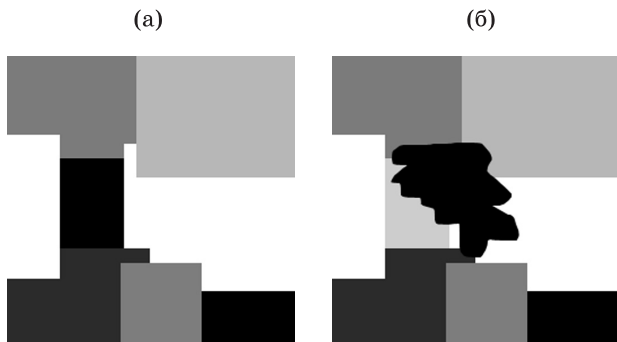


Рис. 1. Исходное изображение и изображение с вырезанной областью.

стей с разной яркостью. Однако если оставить вырезанную область заполненной пикселями постоянной (например, нулевой) яркости, то полученное изображение будет иметь более сложное описание. Если же заполнить ее пикселями разной яркости так, чтобы полное изображение представлялось совокупностью наложенных прямоугольников, длина описания полного изображения сократится.

Для практического применения критерия (3), как отмечалось выше, мы должны сузить пространство моделей, заменив универсальную машину некоторым ограниченным представлением изображений. Естественно, чем богаче представление изображений, чем больше типов регулярностей (текстурных, структурных, семантических) оно использует, тем более качественным может быть восстановление. В частности, для реализации реконструкции из приведенного выше примера необходимо использование структурных представлений. И в практике реконструкции изображений большое внимание уделяют статистическим (текстурным) и структурным представлениям [11, 12] и возможности их совместного использования. Однако здесь для демонстрации перспективности подхода на основе алгоритмической вероятности мы воспользуемся простейшим функциональным представлением. Основное требование, которое следует предъявить к такому представлению, заключается в том, чтобы оно было распределенным, а не локальным, то есть каждый элемент описания изображения зависел бы от яркостей различных пикселей. Этому требованию вполне удовлетворяет представление изображения в форме пространственного спектра.

Реконструкция изображения с использованием спектрального представления

Рассмотрим дискретное преобразование Фурье от изображения $f(x, y)$

$$F(u, v) = \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} f(x, y) e^{-\frac{2\pi i}{N}(ux + vy)},$$

где N – линейный размер изображения (для упрощения записи будем рассматривать квадратные изображения). При неизвестном (произвольном) содержании изображения в области R это преобразование примет вид

$$F(u, v, g|_R) = \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} (f|_R g|_R)(x, y) e^{-\frac{2\pi i}{N}(ux + vy)}.$$

Интересно, что попытка поиска заполнения $g|_R$ с использованием критерия минимизации энергии спектра $F(u, v, g|_R)$ не приводит к положительному результату, поскольку “наилучшим” по такому критерию оказывается заполнение $g|_R(x, y) = 0$.

Рассмотрим сужение критерия (3) на данное представление. Генеративным представлением S , которое позволяет по описанию изображения F породить само изображение $S(F) = f$, является обратное преобразование Фурье. То есть обратное преобразование Фурье является “машиной”, для которой спектр изображения является “программой”, и “исполнение” этой программы на выходе дает исходное изображение. Очевидно, такая машина является неполной по Тьюрингу. Удобной особенностью этого представления является то, что каждому изображению в нем соответствует лишь одно описание, которое можно получить, используя прямое преобразование Фурье, которое мы обозначим через S^{-1} : $S^{-1}(f) = F$, поскольку оно является обратным к S . Перепишем общий критерий (3) для такого представления S :

$$\begin{aligned} g^*|_R &= \arg \max_g P_S(f|_R g|_R) = \arg \max_g \sum_{F: S(F)=f|_R} 2^{-l(F)} = \\ &= \arg \max_g 2^{-l(S^{-1}(f|_R g|_R))} = \arg \min_g l(S^{-1}(f|_R g|_R)). \end{aligned} \quad (4)$$

Отметим, что данный результат верен для любой обратимой программы S (а не только про-

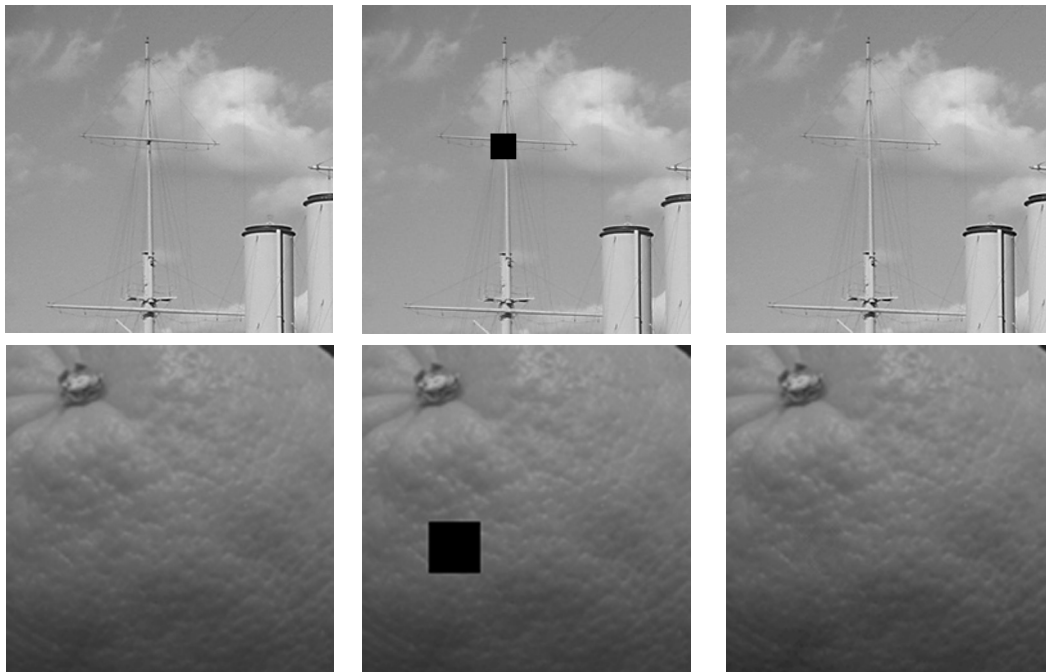


Рис. 2. Примеры исходных изображений (слева), изображений с удаленной областью (посередине) и результаты реконструкции (справа).

граммы, которая восстанавливает изображение по его спектру). При этом если обращение программы S является задачей класса P , то возможен синтез эффективного практического метода.

Таким образом, нам необходимо найти такое заполнение $g|_R(x, y)$ области S , при котором сложность спектра минимальна. Критерий сложности спектра требует некоторого уточнения. Поскольку сам спектр является “моделью” источника изображения, к нему не должны применяться дополнительные методы поиска закономерностей. Здесь мы ограничиваемся энтропийной оценкой для величины $l(F) = H(F)$:

$$H(F) = - \sum_y p(y | F(u, v) = y) \log_2 p(y | F(u, v) = y),$$

где суммирование ведется по всем возможным значениям (амплитудам и фазам) гармоник.

Даже в рамках выбранного упрощенного представления задача поиска минимума критерия (4) не решается в явном виде. Можно предложить следующий итерационный алгоритм (который, однако, не гарантирует сходимости к лучшему решению):

1. Восстанавливаемая область заполняется начальными значениями путем, например, линейной интерполяции.

2. Итерационно выполняется: для каждого пикселя в восстанавливаемой области подбирается такое значение яркости, что энтропия спектра всего изображения (вычисленная при фиксированных текущих значениях яркостей всех остальных пикселей в области) минимальна.

На рис. 2 представлены примеры изображений с восстановлением информации в удаленной области. При этом видно, что восстанавливаются как структурные, так и текстурные особенности изображения, хотя в спектре изображения в явном виде они не представлены.

Заключение

Рассмотрена возможность практического использования теории универсального предсказания на основе алгоритмической вероятности в анализе изображений. Ранее такая возможность не рассматривалась, несмотря на то, что родственная теория универсальной индукции на основе алгоритмической сложности широко и успешно используется в иконике и видеоинформатике.

Синтез эффективных методов экстраполяции данных возможен при замене универсальной опорной машины, обеспечивающей алгоритмическую полноту пространства моделей источников данных, некоторой неуниверсаль-

ной машиной, которая может нести априорную информацию о предметной области, и обращение которой может быть задачей класса P .

Указанная возможность продемонстрирована на примере задачи восстановления изображений в утерянных областях. Предложен метод решения этой задачи, использующий представление изображения в виде пространственного спектра. Несмотря на ограниченность такого представления, с его помощью получены визу-

ально приемлемые результаты восстановления. Использование подхода на основе алгоритмической вероятности совместно с более выразительными представлениями изображений может улучшить качество восстановления.

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации и Совета по грантам Президента Российской Федерации (грант МД-1072.2013.9).

* * * * *

ЛИТЕРАТУРА

1. Luo Q., Khoshgoftaar T.M. Unsupervised Multiscale Color Image Segmentation Based on MDL Principle // IEEE Trans. on Image Processing. 2006. V. 15. № 9. P. 2755–2761.
2. Lindeberg T., Li M.-X. Segmentation and classification of edges using minimum description length approximation and complementary junction cues // Computer Vision and Image Understanding. 1997. V. 67. № 1. P. 88–98.
3. Ward A., Hamarneh Gh. Statistical Shape Modeling using MDL Incorporating Shape, Appearance, and Expert Knowledge // In Lecture Notes in Computer Science, Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2007. P. 278–285.
4. Lanterman A. Minimum Description Length understanding of infrared scenes // Proc. SPIE. 1998. V. 3371. P. 375–386.
5. Li M., Gao Q. and Vitanyi P.M.B. Recognizing on-line handwritten characters using MDL // Proc. IEEE Information Theory Workshop. 1993. P. 24–25.
6. Yvan G. et al. Self-Consistency and MDL: A Paradigm for evaluating point-correspondence algorithms, and its application to detecting changes in surface elevation // Int. J. of Computer Vision. 2003. V. 51. № 1. P. 63–83.
7. Mansouri A.-R. and Konrad J. Minimum description length region tracking with level sets // Proc. SPIE Image and Video Communications and Process. 2000. V. 3974. P. 515–525.
8. Solomonoff R. Does Algorithmic Probability Solve the Problem of Induction? // Oxbridge Research, P.O.B. 391887, Cambridge, Mass. 02139. 1997.
9. Solomonoff R. Algorithmic Probability, Heuristic Programming and AGI. / Baum, E., Hutter, M., Kitzelmann, E. (eds). Advances in Intelligent Systems Research. 2010. V. 10. P. 151–157.
10. Потапов А.С. Выбор представлений изображений на основе минимизации репрезентационной длины их описания // Изв. вузов. Приборостроение. 2008. Т. 51. № 7. С. 3–7.
11. Arias P., Caselles V., Sapiro G. A Variational Framework for Non-local Image Inpainting // Proc. EMMCVPR'09. 2009. P. 345–358.
12. Shibata T., Iketani A., Senda Sh. Fast and Structure-preserving Inpainting Based on Probabilistic Structure Estimation // MVA2011 IAPR Conference on Machine Vision Applications, Nara, JAPAN. 2011. P. 22–25.