

Эмпирический подход к оценке алгоритмов выделения границ

П.П. Кольцов

Аннотация. Эмпирический подход к оценке работы алгоритмов выделения границ основан на использовании специального набора искусственно синтезированных тестовых изображений. Рассмотрен ряд типов границ объектов, сгенерированных с помощью редактора изображений. В работе приведены результаты сравнительного исследования качества работы известных детекторов границ Heitger, Canny, Rothwell, Black, Iverson, Smith, полученные на тестовых изображениях, подвергаемых искажениям.

Ключевые слова: детектор границ, ground truth, чувствительность, специфичность.

Введение

Задача выделения границ объектов на цифровом изображении является одной из самых активно разрабатываемых среди задач цифровой обработки изображений. За последние полвека количество алгоритмов, предлагаемых для решения этой задачи (детекторов границ, edge detectors), достигло нескольких сотен. В этой связи становится актуальной проблема выбора алгоритма, наиболее адекватного конкретным внешним условиям его применения. Под условиями применения в работе подразумеваются различные искажения, изначально присутствующие на изображении и связанные с особенностями получения оцифровки, а также типы граничных точек, присутствующих на обрабатываемом изображении.

Очевидно, что для решения проблемы выбора адекватного детектора границ прежде всего необходимо определить способ получения унифицированной оценки качества работы для множества таких детекторов. Существуют различные подходы к получению такой оценки [1-5]. Известны как аналитические методы [6,7], так и эмпирические. Среди последних можно выделить методы, основанные на использовании изображений с априори заданным истин-

ным (ground truth) решением задачи выделения границ [8, 9], так и в отсутствие такой информации [10, 11]. Использование аналитических методов, как правило, подразумевает выполнение достаточно строгих условий, существенно сужающих круг методов решения задачи выделения границ, для которых такой подход можно использовать. Это обстоятельство определило выбор эмпирического метода в работе. Далее, как отмечено в [10], оценка результатов, полученных при работе алгоритма, реализующего тот или иной метод, зависит как собственно от выбранного метода оценки, так и от самого алгоритма, включая его параметры, а также от изображения, с которым работает алгоритм. Таким образом, в случае, когда имеется фиксированное изображение или набор таких изображений, оценку качества работы того или иного детектора границ можно будет получить, соотнося результаты работы с исходным изображением. Очевидно, что лишь только это соотношение будет выполнено с использованием одной и той же меры близости, оценки качества различных детекторов границ станут сопоставимыми. При этом знание ground truth, то есть истинного решения, позволит считать такие оценки достаточно надежно обоснованными. Все приведенные выше содержательные об-

стоятельства определяют выбор эмпирического подхода к построению унифицированной оценки качества работы различных детекторов границ с использованием ground truth относительно границ на изображении [12]. Ниже в работе будет рассмотрен метод, позволяющий провести сравнение различных алгоритмов выделения границ на основе тестирования поведения детекторов границ на специальном наборе изображений, подвергаемых контролируемым искажениям. Приведены примеры сравнения ряда детекторов границ и полученные на их основе выводы, позволяющие определить детекторы, наиболее адекватные тем или иным внешним условиям.

1. Используемый метод

В соответствии с вышеизложенным, для получения оценки качества работы детекторов границ предлагается использовать набор искусственных полутоновых изображений. Такой набор тестовых изображений должен включать в себя примеры типичных ситуаций, с которыми сталкиваются детекторы границ, включая и те, которые являются для них в некотором смысле трудными. При этом набор должен охватывать значительное количество ситуаций, возникающих при обработке реальных изображений детекторами, то есть обладать определенной полнотой. Как известно [13], имеются следующие типы граничных точек.

- *Уступ (Step)* – граничная точка, по разные стороны от которой интенсивности существенно различаются.
- *Горб (Ridge)* – короткий отрезок, такой, что интенсивность внутри него существенно отличается от интенсивности вне его (два прилегающих Уступа).
- *Излом (Roof)* – граничная точка, в которой производная интенсивности имеет Уступ.

Каждая из этих типов точек характеризуется соответствующими параметрами, как-то: высота Уступа, угол Излома, ширина Горба. Для максимально полного охвата ситуаций, встречающихся при работе детекторов границ, в работе предлагается использовать в качестве ground truth кривые на плоскости. Приведенная выше классификация рассматривает граничные

точки изолированно. Использование кривых как совокупности граничных точек потребует расширения исходной классификации, учитывающего вид границ. Отметим, что для различных граничных точек кривой необходимо предусмотреть варьирование значений характеризующих их параметров, что позволит при обработке изображения получать результат работы детектора границ сразу для широкого спектра значений таких параметров. Кроме того, это позволит выполнить исследование поведения детекторов границ в таких сложных ситуациях, когда варьируется кривизна границы, а также имеются пересечения, то есть при наличии общей граничной точки для нескольких граничных кривых.

В качестве тестовых изображений, предназначенных для оценки качества работы алгоритмов выделения границ, предлагается использовать следующие изображения. Для тестирования алгоритмов на прямолинейных участках границ – изображения *Исчезающий уступ*, *Изменяющийся горб I* (Рис. 1), *Изменяющийся горб II* и *Исчезающий излом* (Рис.2) с параметрами граничных точек, изменяющимися по заданному закону. Для тестирования детекторов границ в случае пересечений и при



Рис 1. Исчезающий уступ (слева) и Изменяющийся горб I (справа)

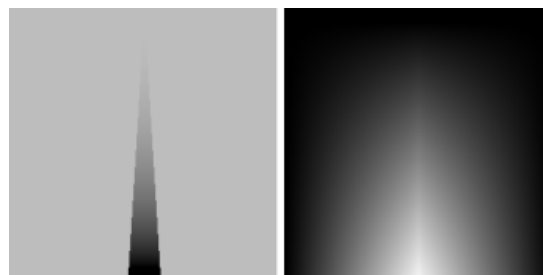


Рис. 2. Изменяющийся горб II (слева) и Исчезающий излом (справа)

разной кривизне границ – изображения *Улитка* и *Узел* (Рис. 3). Параметры граничных точек меняются так же, как и в предыдущем случае.

Тестовые изображения таковы, что на одном изображении представлен широкий спектр параметров граничных точек, относящихся к данному типу изображения. Такой подход позволяет использовать небольшой набор тестовых изображений и представлять результаты тестирования в компактной форме, выявляя особенности тестируемых алгоритмов выделения границ и разумные границы их применимости. Кроме того, учтены сложные для детекторов границ ситуации, связанные с контрастом, кривизной, пересечением.

Отметим, что возможны и иные реализации тестовых изображений, обладающих вышеперечисленными свойствами, повторяющих или дополняющих ситуации, на которых выполняется сравнительное исследование детекторов границ. Соответствующие примеры на приведены на Рис. 4. То есть набор тестовых изображений является открытым.

Для соотнесения результатов работы детектора границ с исходным тестовым изображением возможно использование различных мер близости между полученным результатом и ground truth. В качестве таких мер можно использовать как обычные метрики (Евклидова, Хаусдорфова и т.п.) так и статистические (ошибки первого и второго рода и т.д.). Тем самым можно провести всесторонний анализ поведения детекторов границ на тестовых изображениях.

Контролируемые искажения, налагаемые на тестовые изображения, позволяют моделировать реальные условия получения изображений, предназначенных для обработки детекторами границ. Как правило, в качестве таких искажений рассматриваются размытие или аддитивный шум, поскольку это типичные искажения, встречающиеся на практике. Пример тестового изображения, подвергнутого искажению, приведен на Рис. 5. Исследование динамики поведения детекторов границ при контролируемых искажениях позволяет выявлять детекторы, наиболее адекватные тем или иным внешним условиям, а также определять границы их применимости.

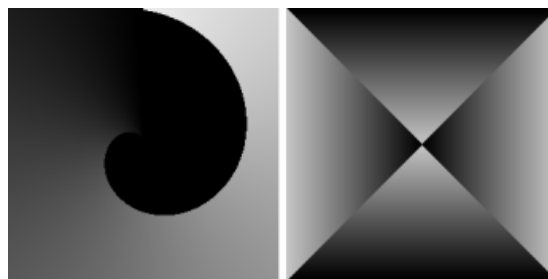


Рис. 3. Улитка (слева) и Узел (справа)

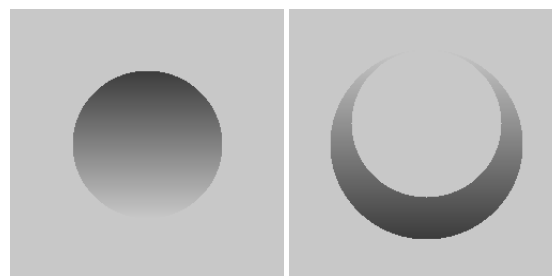


Рис. 4. Возможные варианты тестовых изображений

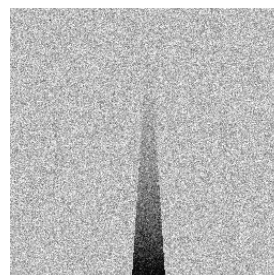


Рис. 5. Изменяющийся горб II (зашумление)

2. Используемый инструментарий

Для выполнения сравнительного исследования алгоритмов выделения границ в соответствии с описанным эмпирическим подходом в качестве инструментального средства была использована программная система PICASSO [14], состоящая из следующих блоков.

1. Редактор изображений, позволяющий формировать искусственные полутоновые изображения с заданными параметрами для границ, а также истинные границы.

2. Блок искажений, позволяющий добавлять к текущему изображению искажения заданного типа с заданными параметрами.

3. Блок вызова детекторов границ, обеспечивающий применение детектора границ к текуще-

му изображению. Вызов детектора границ выполняется внутри цикла, в процессе которого могут изменяться как параметры искажения, так и параметры применяемого детектора границ.

4. Блок оценки качества результатов работы детекторов границ, использующий ground truth, сформированную редактором изображений, а также выбранную для выполнения исследования детектора границ меру близости.

Таким образом, система позволяет генерировать тестовые изображения, накладывая на них искажения, применять к ним детекторы границ и оценивать качество работы последних. Для получения оценки качества работы алгоритмов выделения границ в качестве тестовых были использованы изображения, приведенные на Рис. 1 – Рис.3. Как контролируемые искажения были использованы стандартное размытие, представляющее собой осреднение значений каждого пиксела в некотором окне, а также обычное аддитивное зашумление, представляющее собой изменение каждого пиксела на случайную величину с некоторой функцией распределения. Качество работы детекторов границ измерялось в терминах чувствительности Se и специфичности Sp :

$$Se = \frac{n(B \cap A)}{n(A)}, \quad Sp = \frac{n(B \cap A)}{n(B)},$$

где A – ground truth пиксела, B – пиксела, полученные детектором границ, $n(S)$ – число пикселей в S . Очевидно, что Se показывает, какая часть истинных точек границы найдена детектором, а Sp – какова доля истинных точек границы среди всех найденных.

3. Результаты сравнительного исследования

В работе было выполнено тестирование следующих шести известных детекторов границ.

1. Алгоритм Heitger [15], являющийся реализацией оператора, построенного на основе энергетического фильтра с подавлением шума и улучшением (реставрацией) изображения.

2. Алгоритм Canny [13] – классический детектор границ на основе линейного фильтра с Гауссовым ядром для подавления шума и последующим вычислением направления и мощности границы.

3. Алгоритм Rothwell [16], использующий динамический выбор порога и уточнение границы при постобработке результатов.

4. Алгоритм Black [17], построенный на базе робастной анизотропной диффузии и итеративного сглаживания.

5. Алгоритм Iverson [18], основанный на применении логических операторов, проверяющих для улучшения качества линейного фильтра существование границы.

6. Алгоритм Smith [19], применяющий для определения связанных частей изображения нелинейную фильтрацию.

Более подробное описание этих детекторов границ можно найти в [20]. Перечисленные детекторы границ имеют внешние управляющие параметры. В работе были взяты по умолчанию их значения, использованные авторами. На Рис. 6 и Рис.7, на примере изображения Исчезающий уступ, приведены полученные зависимости значений чувствительности и специфичности по каждому из тестируемых алгоритмов от окна осреднения ω и дисперсии шума $\sigma_{\text{шн}}$ соответственно. Каждый алгоритм применялся к тестовому изображению, подвергнутому размытию с окном осреднения ω в диапазоне от 1 до 10 и зашумлению белым шумом с дисперсией $\sigma_{\text{шн}}$ в диапазоне от 1 до 10, значения Se и Sp приведены в процентах.

Анализ результатов, приведенных на Рис. 6, показывает, что алгоритмы Heitger и Canny в рассматриваемом случае обладают наилучшей чувствительностью, алгоритмы Black и Rothwell имеют несколько худшие показатели. Кроме того, нетрудно видеть, что чувствительность алгоритмов Heitger и Black снижается с увеличением ω не так резко, как у остальных алгоритмов. В отношении специфичности можно говорить, что алгоритмы Canny и Rothwell очевидным образом являются наилучшими, а все остальные имеют слишком плохую специфичность, поскольку на каждую верно определенную точку границы приходится 10 ложных. Аналогично, изучив результаты, приведенные на Рис. 7, можно утверждать, что в рассматриваемом случае все тестируемые алгоритмы обладают чувствительностью, устойчивой к белому шуму, то есть практически не теряют истинных граничных точек. Но, как и выше,

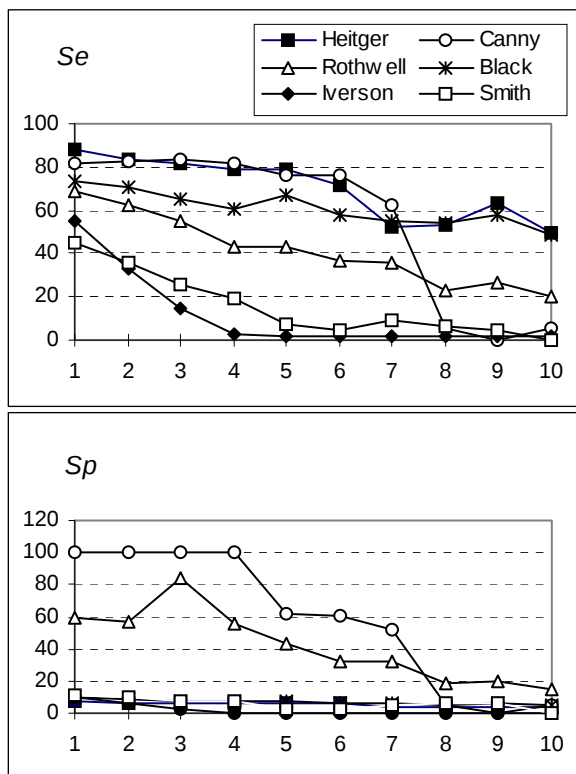


Рис. 6. Ось абсцисс – ширина окна размытия ω , ось ординат – чувствительность Se (вверху) и специфичность Sp (внизу)

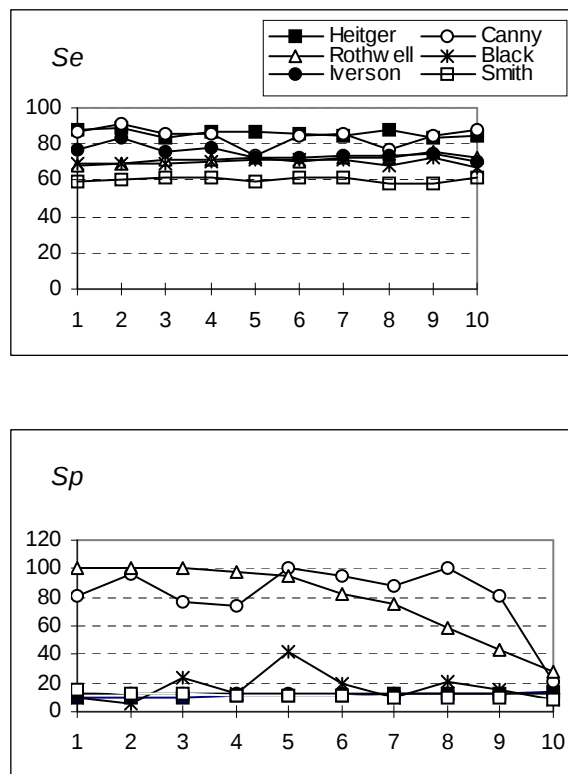


Рис. 7. Ось абсцисс – дисперсия белого шума σ_{wn} , ось ординат – чувствительность Se (вверху) и специфичность Sp (внизу)

алгоритмы Canny и Rothwell обладают лучшей специфичностью, в то время как все остальные алгоритмы показывают примерно одинаковые плохие результаты.

Графики, приведенные на Рис. 6 и Рис. 7, иллюстрируют возможность выполнить сравнение качества работы различных детекторов границ на изображениях с одним и тем же ground truth расположением границ. Для дальнейшего анализа качества работы тестируемых алгоритмов выделения границ уже по всему набору тестовых изображений были выполнены исследования, результаты которых приведены на Рис. 8 и Рис. 9.

Приводимые результаты были получены путем осреднения значений Se и Sp по всему диапазону значений параметров искажений для каждого тестового изображения. На Рис. 8 приведены значения Se и Sp (в процентах), осредненные по параметру размытия ω , на Рис. 9 – аналогичные значения Se и Sp , осредненные по параметру зашумления σ_{wn} . Для более полного

представления относительно характера поведения детекторов границ в различных сложных ситуациях, на Рис. 8 и Рис. 9 приведены также значения дисперсий для Se и Sp .

Анализ результатов, приведенных на Рис. 8, позволяет заключить, что изображение Исчезающий излом является наиболее трудным для всех детекторов границ – никто из них практически не выделил границу. Изображение Изменяющийся горб II является для исследуемых детекторов границ более трудным, чем Изменяющийся горб I. Изображения Улитка и Узел обрабатываются детекторами лучше, чем другие изображения. Алгоритмы Heitger, Canny, Rothwell и Black имеют хорошую чувствительность на всех изображениях, за исключением изображения Исчезающего излома. Алгоритм Smith имеет достаточно хорошую чувствительность на изображениях Улитка и Узел и заметно худшие результаты на изображениях Исчезающий излом, Изменяющийся горб I и Изменяющийся горб II. Алгоритм Iverson имеет стабильно

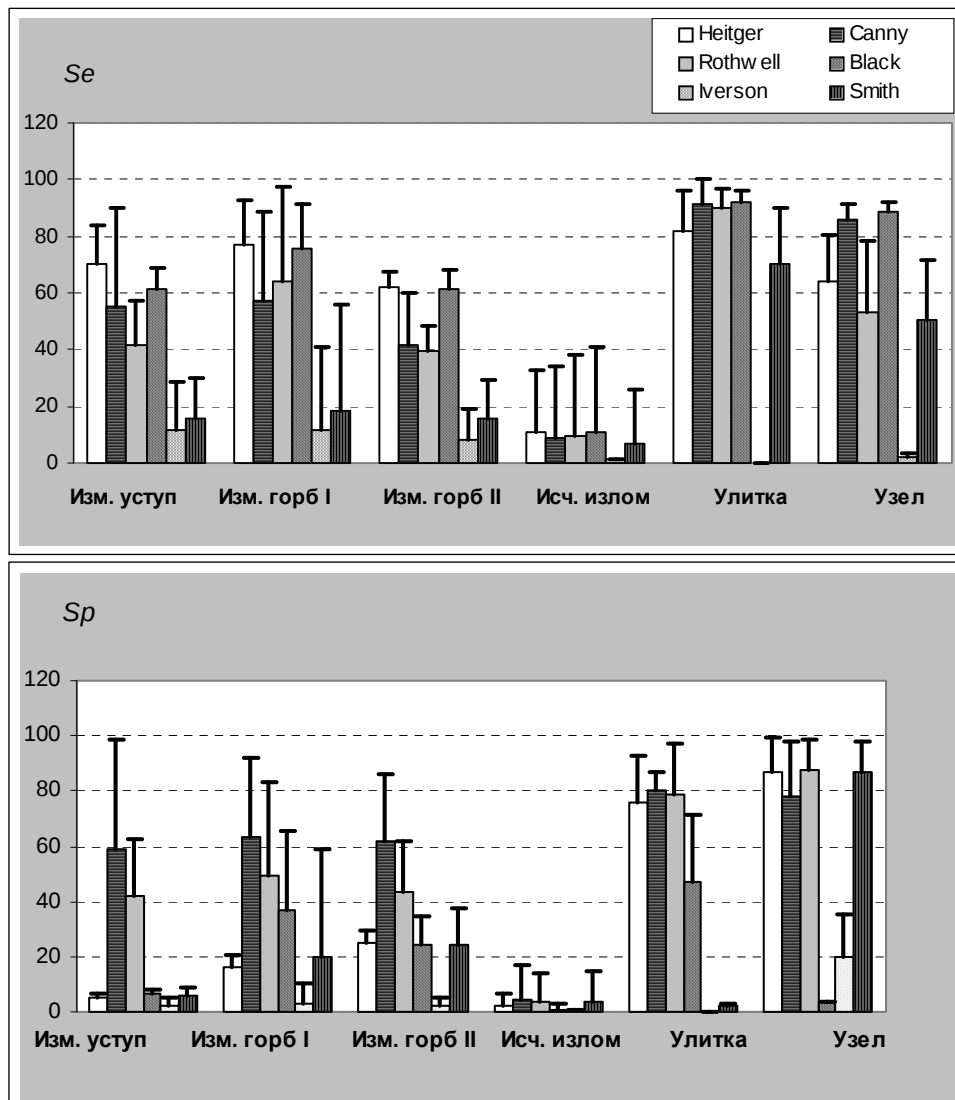


Рис 8. Ось абсцисс - детекторы границ, сгруппированные по тестовым изображениям. Ось ординат - осредненные по параметру зашумления значения (столбцы) чувствительности Se (вверху) и специфичности Sp (внизу). Значения дисперсии показаны в виде вертикальных отрезков над соответствующими столбцами

плохое качество на всех изображениях. Как и ранее, только алгоритмы Canny и Rothwell имеют хорошую специфичность на всех тестовых изображениях, за исключением изображения Исчезающий излом. Алгоритм Heitger имеет очень хорошую специфичность, но только на изображениях Улитка и Узел. Алгоритм Iverson имеет плохую специфичность на всех тестовых изображениях. Аналогичный анализ результатов, приведенных на Рис. 9, позволяет заключить, что в отношении устойчивости к белому шуму картина совершенно иная: все алгоритмы, включая алгоритм Iverson, имеют практически один и тот

же показатель чувствительности. Алгоритмы Canny и Rothwell снова являются наилучшими и имеют хорошую специфичность на всех тестовых изображениях, за исключением изображения Исчезающий излом. Алгоритм Heitger имеет хорошую специфичность, а алгоритм Black обладает удовлетворительной специфичностью только на изображениях Изменяющийся горб I и Узел. Алгоритм Smith показывает очень хорошую специфичность на изображении Изменяющийся горб I. Алгоритм Iverson, как и в предыдущих случаях, имеет плохие результаты на всех изображениях.

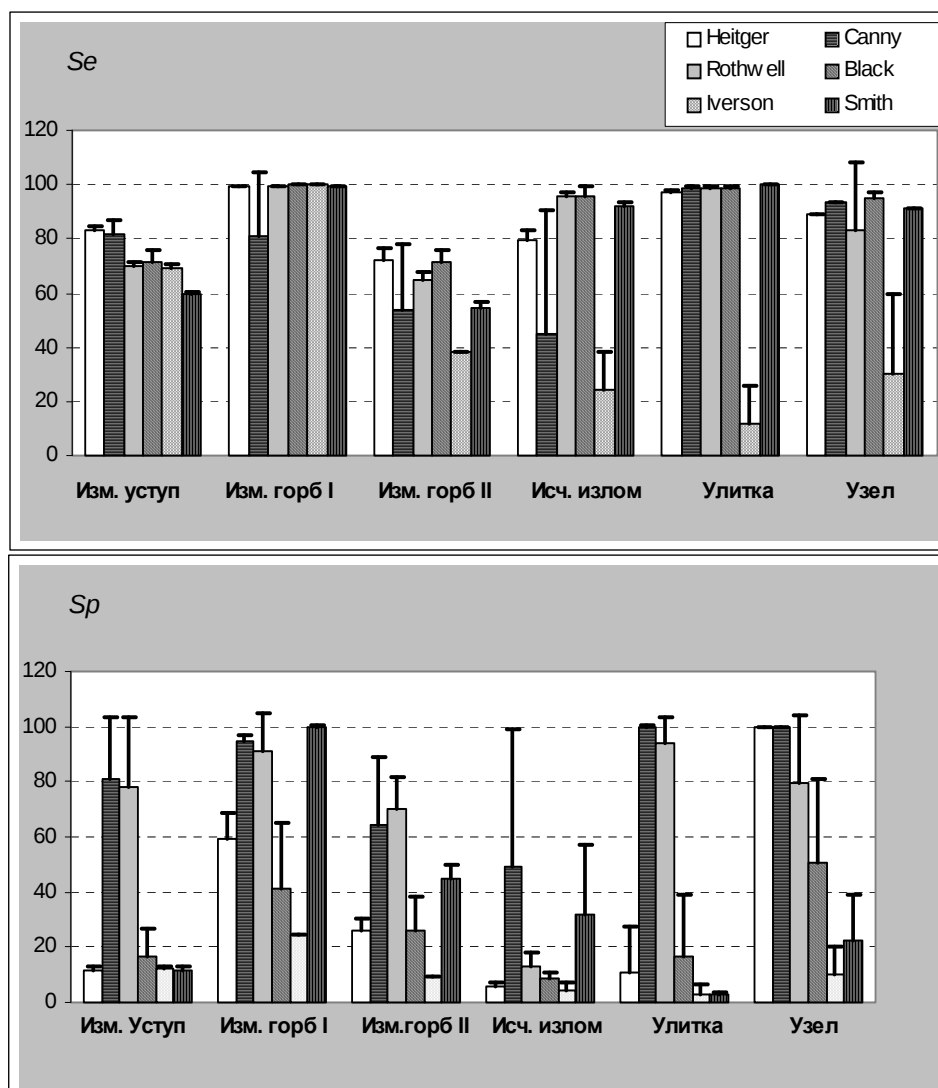


Рис 9. Ось абсцисс - детекторы границ, сгруппированные по тестовыми изображениям. Ось ординат - осредненные по параметру размытия значения (столбцы) чувствительности Se (вверху) и специфичности Sp (внизу). Значения дисперсии показаны в виде вертикальных отрезков над соответствующими столбцами

Заключение

Описанный в работе эмпирический подход к оценке качества алгоритмов выделения границ, основанный на использовании специализированного набора тестовых изображений, позволяет получить унифицированную оценку для различных алгоритмов, решающих задачу нахождения границ объектов на изображении. Собственно выбор набора тестовых изображений является содержательной задачей, относящейся к так называемым плохо формализуемым задачам. Это определяется, в первую

очередь, требованием включения в эту совокупность изображений достаточно полного набора ситуаций, типичных для решаемой задачи, а также ситуаций, в некотором смысле сложных для тестируемых алгоритмов. В работе был использован достаточно компактный набор тестовых изображений, учитывающий сложности, возникающие при работе детекторов границ по уровню контраста на границах объекта, а также по кривизне границ и наличию пересечений. Внесение контролируемого искажения в тестовые изображения позволило более полно выявлять качество работы различных детекторов

границ в условиях, приближенных к реальным. Выбор в качестве мер близости чувствительности и специфичности показал качество работы различных детекторов границ с точки зрения ошибок необнаружения истинных граничных точек и отнесения к граничным ложных точек. В статье для ряда известных детекторов границ были получены выводы по результатам сравнения качества их работы на тестовых изображениях. Таким образом, для этих алгоритмов и для использованной меры близости, на изображениях с объектами, обладающими границами соответствующих типов, можно указать тот, который наиболее адекватен обрабатываемому изображению.

Очевидно, что описанный в работе эмпирический подход может быть использован и для других алгоритмов цифровой обработки изображений. Предлагаемый в статье набор тестовых изображений является открытым и поэтому может быть пополнен изображениями, содержащими ситуации, являющиеся типичными и сложными для других цифровых алгоритмов. При этом, поскольку различные меры близости имеют различные содержательные интерпретации, выбор меры для оценки качества работы тестируемого алгоритма должен определяться конкретными требованиями к решаемой задаче.

Литература

1. Courtney P., Thacker N.A. Performance Characterizations in Computer Vision: The Role of Statistics in Testing and Design, In *Imaging and Vision Systems: Theory, Assessment and Applications*, Blanc-Talon J. and Popescu D. (Eds.), 2001, NOVA Science Books.
2. Haralick R.M. Performance Characterization in Computer Vision, *CVGIP: Image Understanding*, vol. 60, № 2, 1994, pp. 245-249.
3. Heath M., Sarkar S., Sanocki T., Bowyer K. A Robust Visual Method for Assessing the Relative Performance of Edge Detection Algorithms, *IEEE Trans. PAMI*, vol. 19, № 12, 1997, pp. 1338-1359.
4. Kanungo T., Jaisimha M.Y., Palmer J., Haralick R.M. A Methodology for Quantitative Performance Evaluation of Detection Algorithms, *IEEE Trans. Image Processing*, vol. 4, № 12, 1995, pp. 1667-1674.
5. Coleman S.A., Scotney B.W., Herron M.G. A Validated Edge Model Technique for the Empirical Performance Evaluation of Discrete Zero-crossing Methods, *Image and Vision Computing*, vol. 25, № 8, 2007, pp. 1315-1328.
6. Ramesh V., Haralick R.M. Performance Characterization of Edge Detectors, *SPIE*, vol. 1, *Applications of Artificial Intelligence X: Machine Vision and Robotics*, 1992, pp. 252-266.
7. Steger C. Analytical and Empirical Performance Evaluation of Subpixel Line and Edge Detection, *Empirical Evaluation Techniques in Computer Vision*, Bowyer K.W. and Phillips P.J. (Eds.), 1998, IEEE Computer Press.
8. Pavlidis T. Why Progress in Machine Vision Is So Slow, *Pattern Recognition Letters*, vol. 13, № 4, 1992, pp. 221-225.
9. Salotti, M., Bellet, F., Garbay, C. Evaluation of Edge Detectors: Critics and Proposal, *Workshop on Performance Characteristics of Vision Algorithms*, 1996, Cambridge.
10. Abdou I.E., Pratt W.K. Quantitative Design and Evaluation of Enhancement/Thresholding Edge Detectors, *Proc. IEEE*, vol. 67, № 5, 1979, pp. 753-763.
11. Jiang X.Y., Hoover A., Jean-Baptiste G., Goldgof D., Bowyer K., Bunke H. A Methodology for Evaluating Edge Detection Techniques for Range Images, *Proc. Asian Conf. Computer Vision*, 1995, pp. 415-419.
12. Fernández-García N.L., Carmona-Poyato A., Medina-Carnicer R. Madrid-Cuevas F.J. Automatic Generation of Consensus Ground Truth for the Comparison of Edge Detection Techniques, *Image and Vision Computing*, vol. 26, № 4, 2008, pp. 496-511.
13. Canny J. A Computational Approach to Edge Detection, *IEEE Trans. PAMI*, vol. 8, № 16, 1986, pp. 679-697.
14. Gribkov I.V., Koltsov P.P., Kotovich N.V., Kravchenko A.A., Koutsaeve A.S., Osipov A.S., Zakharov A.V. PICASSO 2 – a System for Performance Evaluation of Image Processing Methods. *Proc. 5th WSEAS Int. Conf. on Signal Processing, Computational Geometry & Artificial Vision*, 2005, pp. 88-93.
15. Heitger F., Rosenthaler L., von der Heydt R., Peterhans E., Kubler O. Simulation of neural contour mechanisms: From simple to end-stopped cells, *Vision Research*, № 32, 1992, pp. 963-981.
16. Rothwell C.A., Mundy J.L., Hoffman W., Nguyen V.-D., *Driving Vision by Topology*, *Int. Symp. Computer Vision*, 1995, pp. 395-400.
17. Black M., Sapiro G., Marimont D., Heeger D., Robust Anisotropic Diffusion, *IEEE Trans. Image Processing*, vol. 7, № 3, 1998, pp. 421-432.
18. Iverson L.A., Zucker S.W. Logical/Linear Operators for Image Curves, *IEEE Trans. PAMI*, vol. 17, № 10, 1995, pp. 982-996.
19. Smith S.M. Flexible Filter Neighbourhood Designation, *Proc. 13th Int. Conf. on Pattern Recognition*, 1996, vol. 1, pp. 206-212.
20. http://marathon.csee.usf.edu/edge/edgecompare_main.html, Edge Detector Performance. Evaluation Study @ USF.

Кольцов Петр Петрович. Заместитель директора Научно-исследовательского института системных исследований РАН. Окончил Московский физико-технический институт в 1971 году. Кандидат физико-математических наук, доцент. Автор автор более 80 печатных работ. Область научных интересов - цифровая обработка изображений, распознавание образов. E-mail: koltsov@niisi.msk.ru.