

ПРИМЕНЕНИЕ МАСШТАБИРОВАННОЙ КАРТЫ ГЛУБИНЫ ДЛЯ ПЕРВОЙ ИТЕРАЦИИ ВЫЧИСЛЕНИЯ ЦЕН СОПОСТАВЛЕНИЯ НМІ

А.В. Аргу́тин

В статье рассматривается алгоритм вычисления цен сопоставления на основе взаимной информации между пикселями стереоизображений. Во второй части приводится математическое выражение величины взаимной информации, основывающееся на величинах энтропии изображений. Далее описывается способ применения и оптимизации величины взаимной информации при вычислении карты диспаратностей. В параграфе «Особенности алгоритма» анализируются возможные условия работы алгоритма, а также ошибки в его работе и причины их появления.

Ключевые слова: стереозрение, взаимная информация, цена сопоставления, распределение вероятностей.

Введение

Одним из важнейших аспектов работы любого алгоритма стереосопоставления является способ вычисления степени взаимного соответствия пикселей исходных изображений. Среди множества способов вычисления цен взаимного соответствия между пикселями, описанных в [1, 2], стоит выделить величину Биршфилда и Томаси [3] как нечувствительную к явлению сэмплинга (особой форме дискретизации изображения), а также цену взаимного сопоставления, основывающуюся на взаимной информации [2, 3], нечувствительную к изменениям освещенности и условиям записи. Свойства цены взаимной информации неоченимы при работе с реальными изображениями, например, при работе с видеопотоками со стереопары или с веб-камер, где уровень освещенности одного изображения колеблется относительно уровня освещенности другого изображения.

Методы вычисления взаимной информации

Величина, определяющая взаимную информацию двух изображений, вычисляется на основе величин энтропии каждого изображения в отдельности за вычетом их общей энтропии [4].

$$MI_{I_1, I_2} = H_{I_1} + H_{I_2} - H_{I_1, I_2} . \quad (1)$$

Для пары качественно снятых стереоизображений одной сцены величина взаимной энтропии мала, поскольку одно изображение может быть в значительной мере предсказано его стереопарой, что, в свою очередь, повышает значение взаимной информации для двух изображений. Значение взаимной информации для левого и правого изображений будет максимальным, если выполнить сдвиг пикселей одного из двух стереоизображений с помощью карты диспаратностей, поскольку в результате данного действия пиксели левого и правого изображений будут находиться по одним индексам по вертикали и горизонтали [5]. Максимальное значение взаимной информации для двух стереоизображений является доказательством того факта, что найдена максимально верная карта диспаратностей для данной пары стереоизображений.

Однако выражение (1) применимо лишь к изображению в целом и не может быть использовано как величина взаимного сопоставления пикселей, не говоря уже о том, что требует для вычисления уже предопределенной карты глубины (что подразумевает решение задачи более высокого уровня).

Применение взаимной информации в качестве цены сопоставления

Ким и др. в [6] преобразовал вычисление взаимной энтропии H_{I_1, I_2} к виду суммы термов, используя ряд Тейлора. Каждый терм суммы зависит от соответствующих интенсивностей на стереоизображениях и вычисляется независимо для каждого пикселя p [7].

$$H_{I_1, I_2} = \sum_p h_{I_1, I_2}(I_{1p}, I_{2p}). \quad (2)$$

Терм h_{I_1, I_2} вычисляется на основе взаимного распределения вероятностей P_{I_1, I_2} , соответствующих значений интенсивности.

$$h_{I_1, I_2}(i, k) = -\frac{1}{n} \log(P_{I_1, I_2}(i, k) \circ g(i, k)) \circ g(i, k). \quad (3)$$

Величина взаимного распределения вероятности определяется оператором T , равным 1, если аргумент является истинным и 0, если аргумент является ложным.

$$P_{I_1, I_2}(i, k) = \frac{1}{n} \sum_p T[(i, k) = (I_{1p}, I_{2p})] \quad (4)$$

Расчет энтропии левого и правого изображений производится аналогично, учитывая, что распределения вероятностей интенсивностей левого и правого изображения могут быть рассчитаны на основе взаимного распределения вероятностей P_{I_1, I_2} :

$$P_{I_1}(i) = \sum_k P_{I_1, I_2}(i, k), \quad (5)$$

$$H_I = \sum_p h_I(I_p), \quad h_{I_1}(i) = -\frac{1}{n} \log(P_{I_1}(i) \circ g(i)) \circ g(i). \quad (6)$$

Таким образом, можно выразить

$$MI_{I_1, I_2} = \sum_p h_{I_1}(i) + h_{I_2}(k) - h_{I_1, I_2}(i, k). \quad (7)$$

Величина MI_{I_1, I_2} (7) является мерой взаимной информации, сосредоточенной в двух стереоизображениях, а каждый терм под знаком суммы является искомой величиной цены сопоставления двух пикселей. Таким образом:

$$C_{MI}(p, d) = -\left\langle h_{I_1}[I_{base}(p)] + h_{I_2}[I_{matching}(q)] - h_{I_1, I_2}[I_{base}(p), I_{matching}(q)] \right\rangle, \quad (8)$$

где $q = [p_x - d, p_y]$.

Особенности алгоритма

Необходимость в карте глубины а priori на данном этапе остается не решенной проблемой. В работе Кима и др. [6] предлагается использовать итеративный подход, на первой итерации которого карта глубины имеет случайные значения. Далее результирующие цены сопоставления используются для поиска карты глубины, используемой в следующей итерации. Количество итераций весьма мало (в частности, 3). Это объясняется тем, что даже ошибочные значения глубин позволяют получить значения взаимного распределения вероятностей, близкие к реальным.

В работе Hirsh Muller [7] описывается развитие этой идеи, где предлагается использовать карту глубины, полученную при расчете уменьшенного изображения, растянутую до размеров требуемой карты глубины. Расчеты показывают, что если трижды провести подобную операцию над изображением, уменьшенным в 16 раз, а затем итеративно подавать на вход алгоритма 1/8, затем 1/4, 1/2 и исходного размера, то количество операций, необходимое для выполнения данного алгоритма, возрастет лишь в 1,14 относительно количества операций, необходимых для расчета последней итерации над изображением исходного размера, что увеличивает производительность в 2,6 раза по сравнению с тремя итерациями алгоритма над изображениями исходного размера.

Однако если реализовать метод, описанный выше, результирующая карта глубины будет содержать артефакты, свидетельствующие об ошибках в работе алгоритма (см. рис. 1). Ниже приведен анализ причин этого явления.

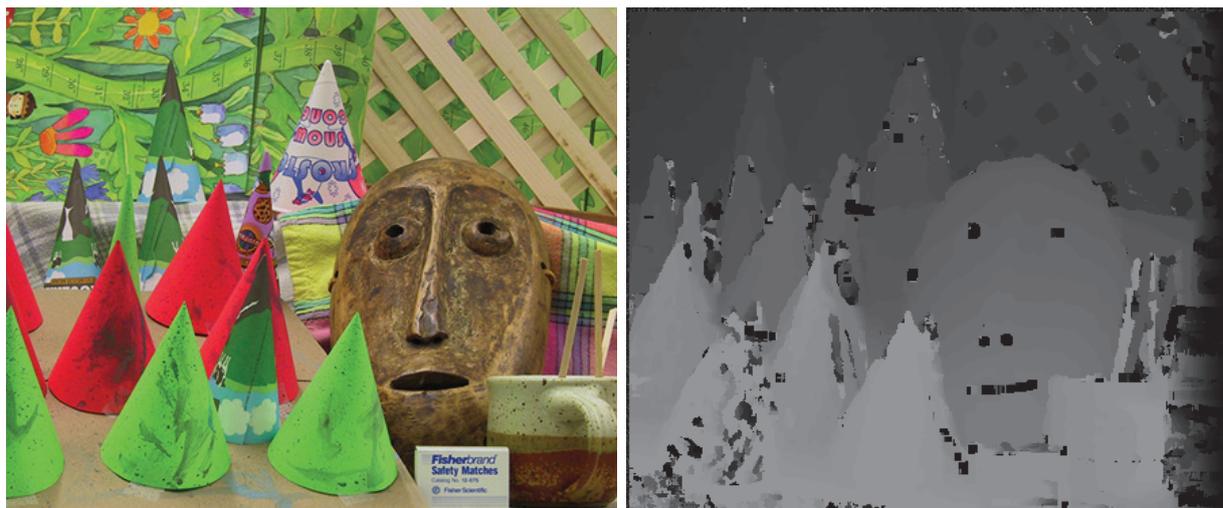


Рис. 1. Исходное изображение (слева) и карта диспаратностей, рассчитанная на основе цен сопоставления, описываемых взаимной информацией

При вычислении термов h_{I_1, I_2} по формуле (3), описанной в [7], возможным результатом является значение бесконечности со знаком плюс. Такой результат получается при расчете термина энтропии для нулевого значения взаимного распределения вероятностей (см. рис. 2) пикселей левого и правого изображений.

С алгоритмической точки зрения данное значение, будучи математически верным, не является удобным в использовании: причина тому заключается в том, что значение бесконечности, как цена сопоставления, при суммировании с другими значениями в случае агрегирования цен сопоставления смазывает картину – суммарная цена сопоставления всегда будет равняться бесконечности.

Заключение

Полученный результат показывает, что в таком случае невозможно оценить, какой из двух участков с бесконечной агрегированной ценой сопоставления соответствует исходному участку с большей точностью. Подобное явление можно наблюдать в виде артефактов черного цвета на карте диспаратностей на рис. 1. Это означает, что для стерео-алгоритма, использующего агрегирование цен стерео-сопоставления, все пиксели-кандидаты на сопоставление с базовым пикселем, в рассматриваемой области которых лежит пиксель с нулевым значением $P_{I_1, I_1}(i, k)$, являются равноценно невероятными. Негативным последствием этого является тот факт, что при отсутствии других кандидатов будет выбран первый пиксель среди рассматриваемых, и это означает, что алгоритм не выполняет анализа для пикселей такого рода.

Литература

1. Scharstein, D. A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms / D. Scharstein, R. Szeliski // *IJCV*. – 2002. – 47(1/2/3) – P. 7–42.
2. Hirsh Muller, H. Evaluation of Cost Functions for Stereo Matching / H. Hirsh Muller // *CVPR*. – 2007. – 17–2(6). – P. 1–8.

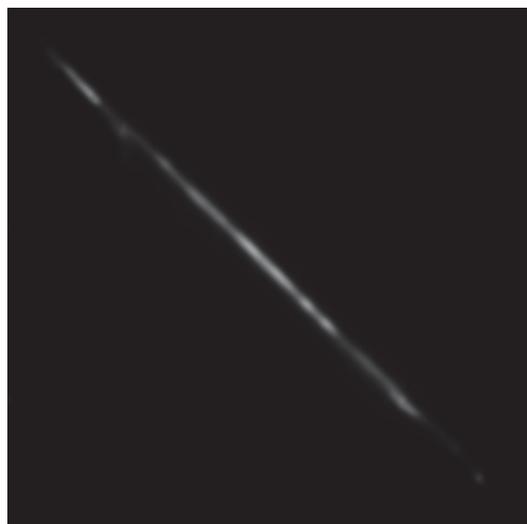


Рис. 2. Карта взаимного распределения вероятностей для интенсивностей пикселей левого и сдвинутого правого изображений

3. Birchfield, S. *Depth Discontinuities by Pixel-to-pixel Stereo* / S. Birchfield, C. Tomasi // *IJCV*. – 1999. – 35(3) – P. 1–25.

4. Egnal, G. *Mutual Information as a Stereo Correspondence Measure* / G. Egnal // *CIS Technical Reports*. – 2009. – 18(2). – P. 124–132.

5. Viola, P. *Alignment by maximization of mutual information* / P. Viola, W. M. Wells // *IJCV*, 24(2). – 1997. – P. 137–154.

6. Kim, J. *Visual Correspondence Using Energy Minimization and Mutual Information* / J. Kim, V. Kolmogorov, R. Zabih // *ICCV*. – Vol. 2. – 2003. – P. 1033–1040.

7. Hirschmuller, H. *Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information* // *CVPR*. – 2005. – 20–26(6). – Vol. 2. – P. 807–814.

Аргутин Александр Вячеславович, аспирант кафедры ЭВМ, Южно-Уральский государственный университет (г. Челябинск), alex.argutin@gmail.com; телефон +7 912 407 50 29.

Bulletin of the South Ural State University
Series “Computer Technologies, Automatic Control, Radio Electronics”
2013, vol. 13, no. 2, pp. 118–121

USING OF THE SCALED DISPARITY MAP FOR THE FIRST ITERATION OF THE MATCHING COST HMI CALCULATION PROCESS

A.V. Argutin, South Ural State University, Chelyabinsk, Russian Federation,
alex.argutin@gmail.com

The article describes the algorithm of the matching cost calculation based on mutual information between pixels of stereo-images. The second part explains mathematical formulation of the mutual information based on image entropies. Mutual information application and optimization approaches in disparity calculation are described next. The «Algorithm peculiarities» paragraph analyses possible algorithm conditions, errors and their reasons.

Keywords: stereo-vision, mutual information, matching cost, probability distribution.

References

1. Scharstein D.A., Szeliski R. Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms. *International Journal of Computer Vision*, 2002, vol. 47, pp. 7–42, available at: <http://vision.middlebury.edu/stereo/taxonomy-IJCV.pdf>

2. Hirshmuller H. Evaluation of Cost Functions for Stereo Matching. *Computer Vision and Pattern Recognition*, 2007, vol. 0, pp. 1-8, available at: http://vision.middlebury.edu/~schar/papers/evalCosts_cvpr07.pdf

3. Birchfield S., Tomasi C. Depth Discontinuities by Pixel-to-pixel Stereo. *International Journal of Computer Vision*, 1999, vol. 3, pp. 1–25, available at: http://www.ces.clemson.edu/~stb/publications/p2p_iccv1998.pdf

4. Egnal G. Mutual Information as a Stereo Correspondence Measure. *CIS Technical Reports*, 2009, vol. 2, pp. 124–132, available at: http://repository.upenn.edu/cgi/viewcontent.cgi?article=1115&context=cis_reports

5. Viola P., Wells W. M. Alignment by Maximization of Mutual Information. *International Journal of Computer Vision*, 1997, vol. 2, pp. 137–154, available at: <http://research.microsoft.com/en-us/um/people/viola/Pubs/Detect/IJCV-97.pdf>

6. Kim J., Kolmogorov V., Zabih R. Visual Correspondence Using Energy Minimization and Mutual Information. *International Conference of Computer Vision*, 2003, vol. 2, pp. 1033–1040, available at: <http://www.cs.cornell.edu/~rdz/papers/kkz-iccv03.pdf>

7. Hirschmuller H. Accurate and Efficient Stereo Processing by Semi-Global Matching and Mutual Information. *Computer Vision and Pattern Recognition*, 2005, vol. 2, pp. 807–814, available at: <http://www.dlr.de/rm/en/PortalData/3/Resources/papers/modeler/cvpr05hh.pdf>

Поступила в редакцию 28 февраля 2013 г.