

# Эффективный оператор сжатия динамического диапазона яркостей<sup>1</sup>

Б.Х. Барладян, А.Г. Волобой, В.А. Галактионов, Э.А. Копылов

Институт Прикладной Математики им. М.В. Келдыша РАН

Москва, Россия

Для преобразования большого динамического диапазона яркостей изображений в диапазон яркостей, который может быть воспроизведен на экране монитора, используются специальные операторы сжатия (Tone Mapping Operators). Эта работа представляет эффективный и простой в использовании оператор сжатия, основанный на последних идеях в этой области. Процесс оценки его параметров использует метод выборок. Особое внимание посвящено обеспечению устойчивости алгоритма оценки параметров. Разработанный оператор сжатия обеспечивает хорошее качество изображений и практически не требует ручной настройки параметров.

## 1. ВВЕДЕНИЕ

Изображения с большим динамическим диапазоном яркостей становятся все более распространенными в современной компьютерной графике. Первоначально такие изображения являлись результатом исключительно физически аккуратного моделирования сцен. В настоящее время изображения реальных сцен с большим динамическим диапазоном яркостей могут быть получены путем обработки нескольких фотографий с различной экспозицией [1], панорамного сканирования [2, 3]. Более того, в ближайшем будущем фотоаппараты и видеокамеры будут позволять получать изображения и видео с большим динамическим диапазоном непосредственно [4, 5, 6].

Изображения с большим динамическим диапазоном яркостей имеют много преимуществ по сравнению с обычными изображениями с низким динамическим диапазоном [7]. Также существует ряд приложений, где использование таких

изображений является очень полезным [8, 9]. Однако использование изображений с большим динамическим диапазоном яркостей вызывает большие проблемы при отображении их на стандартных устройствах (графические мониторы, принтеры и т.д.) динамический диапазон которых существенно меньше, чем диапазон яркостей реального мира. Для отображения изображений с большим динамическим диапазоном яркостей на стандартных устройствах отображения этот диапазон должен быть сжат до уровня, который может быть воспроизведен на этих приборах. В литературе такой оператор сжатия называется *tone mapping* или *tone reproduction*. Качество работы таких операторов оценивается главным образом по уровню сохранения деталей изображения, локального контраста и отсутствию артефактов. Существенным фактором является также скорость работы алгоритма и объем требуемой памяти, поскольку в ряде случаев необходимо обеспечить работу в режиме реального времени.

---

<sup>1</sup> Работа была поддержана РФФИ, грант 04-01-00520, а также компанией INTEGRA Inc. (Япония).

Практическая потребность в таких алгоритмах стимулировала появление большого количества работ посвященных созданию операторов сжатия эффективных для изображений с большим динамическим диапазоном яркостей. Ряд работ был представлен на конференциях SIGGRAPH, Eurographics и Graphicon: [9 - 15]. Большинство из алгоритмов, представленных в этих работах, могут с успехом применяться как к реальным изображениям (например, к фотографиям) с большим динамическим диапазоном яркостей, так и к изображениям, полученным в системах физически аккуратного моделирования. В действительности, применение оператора в системах моделирования имеет определенную специфику связанную с технологией сглаживания контуров изображений (anti-aliasing), о которой будет сказано далее.

Операторы сжатия динамического диапазона яркостей могут быть разделены на две большие группы: (1) глобальные (пространственно инвариантные) и (2) локальные (пространственно зависимые) операторы [16, 17, 18, 1, 11].

Глобальные операторы преобразуют яркость каждого пикселя, используя фиксированную кривую, которая строится, используя некоторые усредненные характеристики всего изображения. Наиболее часто в качестве такой характеристики используется средняя логарифмическая яркость. Простота этих алгоритмов очень притягательна, но детали изображений часто теряются в особо ярких или очень темных областях изображения, где динамический диапазон особенно велик.

В локальных операторах преобразование яркости в каждом пикселе зависит от средней яркости окрестности преобразуемого пикселя. Наиболее трудной и вычислительно дорогостоящей проблемой алгоритмов такого типа является правильное определение размера локальной окрестности для каждой точки. Разнообразные артефакты могут возникать в случаях, когда это сделано не

совсем корректно. Примеры таких артефактов приведены в [12].

Использование локальных операторов для преобразования изображений, созданных в физически аккуратных системах моделирования имеет свою дополнительную специфику помимо отмеченных выше проблем с артефактами и большим объемом вычислений. Получение высококачественных изображений требует сглаживания контуров изображений (anti-aliasing). Используемые для этих целей алгоритмы обычно делят пиксел экрана на более мелкие составляющие, что увеличивает размер получаемого изображения более чем на порядок. Затем финальное высококачественное изображение получается путем усреднения составляющих пикселей. Оператор сжатия должен быть применен для каждого составляющего пикселя. В противном случае качество финального изображения может существенно ухудшиться. Применение локальных операторов для изображений такой большой размерности вызывает дополнительные трудности, поскольку требуют большого объема используемой памяти и существенно увеличивают время обработки. Требования к памяти становятся особенно критическими, когда последовательность изображений должна совместно обрабатываться при создании анимационных фильмов. В этом случае оператор сжатия должен обеспечить также гладкое изменение яркости изображений от кадра к кадру.

Принимая во внимание все выше описанные проблемы, нами был разработан оператор сжатия динамического диапазона яркостей и устойчивый алгоритм для оценки его параметров.

## 2. АЛГОРИТМ

Наш оператор сжатия основан на улучшенном операторе Тумблина-Русмеера [18] и дополнительной формуле, предложенной

Рейнхардом и др. [12, 13], которая сжимает главным образом большие яркости.

Оператор сжатия, первоначально предложенный Тумблином и Русмеером [19] использует модель сохранения яркости, основанной на математической модели человеческого зрения Стивенса [20]. Целью этой модели является обеспечение постоянного отношения яркости изображения воспринимаемого с экрана монитора к яркости соответствующей части реальной сцены при любых условиях освещения. Наш оператор основан на улучшенном операторе Тумблина-Русмеера, который уменьшает контраст изображения на мониторе для очень темных сцен, что предотвращает преувеличение и обращение контраста.

Улучшенный оператор сжатия определяется следующей формулой [17, 18]:

$$L_d = m(L_{wa}) \cdot L_{da} \cdot \left( \frac{L_w}{L_{wa}} \right)^{\left( \frac{\gamma_w}{\gamma_d} \right)} \quad (1)$$

где

$L_{da}$  - адаптационная яркость экрана. Типичное значение порядка  $10\text{-}30 \text{ cd/m}^2$ ,

$L_{wa}$  - адаптационная яркость сцены, найденная из яркостей сцены (физических яркостей пикселей  $L_w$ ) используя формулу

$$\log(L_{wa}) = \text{mean} \left\{ \log(L_w + 2.3 \cdot 10^{-5} \text{ cd/m}^2) \right\} \quad (2)$$

$\gamma_d$  и  $\gamma_w$  определяются, используя формулу Стивенса для чувствительности контраста:

$$\gamma(L_a) = \begin{cases} 2.655 & \text{for } L_a > 100 \text{ cd/m}^2 \\ 1.855 + 0.4 \log_{10}(L_a + 2.3 \cdot 10^{-5}) & \text{otherwise,} \end{cases} \quad (3)$$

для человека адаптированного для экрана монитора и сцены соответственно -  $\gamma_d = \gamma(L_{da})$

и  $\gamma_w = \gamma(L_{wa})$ ,

$m(L_{wa})$  - коэффициент, зависящий от адаптационной яркости сцены, предотвращающий anomalously серые ночные изображения:

$$m(L_{wa}) = \left( \sqrt{C_{\max}} \right)^{\gamma_{wd}^{-1}} \quad (4)$$

где  $C_{\max}$  есть максимально допустимый контраст монитора (типично от 30 до 100),

$$\gamma_{wd} = \left( \frac{\gamma_w}{1.855 + 0.4 \log(L_{da})} \right).$$

Коэффициент  $m$  монотонно увеличивает яркость монитора с ростом адаптационной яркости сцены  $L_{wa}$  вплоть до верхней границы зрения.

Мы применяли эти формулы к яркостям, вычисленным из RGB троек используя формулу

$$L_w(x,y) = 0.252R(x,y) + 0.664G(x,y) + 0.084B(x,y). \quad (5)$$

Вычисление адаптационной яркости сцены по всем пикселям требует чрезмерно больших вычислительных затрат и большого объема памяти. Поэтому мы использовали для вычисления адаптационной яркости сцены выборку пикселей, сделанную по некоторой, относительно редкой сетке. Только небольшая часть пикселей (обычно  $\sim 1\%$ ) использовалась для этих вычислений. Некоторые детали этих вычислений и обеспечение устойчивости оценки параметров, обсуждается в следующем разделе.

В общем случае, в результате использования улучшенного оператора сжатия Тумблина-Русмеера, яркость некоторых пикселей монитора, соответствующих наиболее яркой части сцены, может превышать максимальную яркость монитора и соответствующая часть изображения будет перенасыщена. Для преодоления этой проблемы наш оператор использует формулу, сжимающую большие яркости, предложенную Рейнхардом и др. [12]:

$$L_{df}(x,y) = \frac{L_d(x,y) \left( 1 + \frac{L_d(x,y)}{L_{white}^2} \right)}{1 + L_d(x,y)}, \quad (6)$$

где  $L_{white}$  есть минимальная яркость, отображаемая в чисто белый цвет. Для изображений с очень большим динамическим диапазоном яркостей уровень белого  $L_{white}$  может быть выбран как почти любая достаточно большая величина (обычно близкая к максимальной яркости сцены). Выбор этого параметра становится критическим для изображений с низким и

средним динамическим диапазоном яркостей. Используемый в работе способ оценки этого параметра обсуждается в следующем разделе.

### 3. ОЦЕНКА ПАРАМЕТРОВ.

В предыдущем разделе было показано, что предложенный оператор сжатия зависит от двух параметров – адапционной яркости сцены  $L_{wa}$  и уровня белого  $L_{white}$ . Выборка по редкой сетке используется для оценки обоих этих параметров. Оценка производится в два этапа.

Первоначально адапционная яркость сцены оценивается, используя уравнение (2) и уровень белого выбирается так, чтобы ~1% пикселей сцены имело яркость, превосходящую эту величину.

Этот подход работает удовлетворительно для большинства сцен, но дает плохие результаты для изображений с большими темными областями – значительная, наиболее интересная часть изображения, становится перенасыщенной. Предложенная в [18] константа  $2.3 \cdot 10^{-5} cd/m^2$  позволяет защититься от появления нулевых величин при вычислении логарифма, но является слишком малой для получения удовлетворительной величины адапционной яркости для таких сцен. Для преодоления этой проблемы мы исключаем слишком темные пиксели из вычислений при определении адапционной яркости сцены по формуле (2). Следующая эвристическая формула была предложена и использовалась для определения уровня отсечения:

$$L_{trsh} = \text{Min}(L_{wa}/20, L_{white}/100)$$

Дальнейшее улучшение касается более тонкой настройки уровня белого. Величина, выбранная на первом этапе, обеспечивает хорошие результаты, когда выбранный уровень белого превышает в небольшом количестве пикселей и перенасыщенность в этих точках не оказывает существенного влияния на восприятие изображений,

поскольку большая, основная часть изображения имеет яркость существенно ниже выбранного уровня белого. Но этот подход работает плохо, когда значительная часть пикселей изображения имеет яркость близкую к выбранному уровню белого. В этом случае вся эта часть изображения снова становится перенасыщенной. Для преодоления этой проблемы уровень белого на втором этапе выбирается так, чтобы уровень белого, выбранный первоначально, соответствовал бы заданному уровню яркости монитора. Эмпирическая величина 0.98 (для канонического диапазона [0, 1]) дает хорошие результаты для большинства тестируемых сцен. Обозначим эту величину как  $L_{dwt}$  – уровень белого для монитора. Новый финальный уровень белого  $L_{fwhite}$  вычисляется путем совместного решения уравнений полученных из (1) (6), используя уже вычисленные  $L_{wa}$ ,  $L_{dwt}$  и первоначальное значение  $L_{white}$  в пространстве физических величин яркостей:

$$\begin{cases} L_d = m(L_{wa}) \cdot L_{da} \cdot \left( \frac{L_{white}}{L_{wa}} \right)^{\left( \frac{\gamma_w}{\gamma_d} \right)} \\ L_{dwt} = \frac{L_d \cdot \left( 1 + \frac{L_d}{L_{fwhite}^2} \right)}{1 + L_d} \end{cases}$$

### 4. РЕЗУЛЬТАТЫ<sup>2</sup>

Разработанный метод был реализован в нескольких программных системах компании INTEGRA Inc. [21] и показал хорошие результаты для широкого набора изображений с большим динамическим диапазоном яркостей, созданных в этих системах. Мы также экспериментировали с различными изображениями реальных сцен, имеющих большой динамический диапазон.

<sup>2</sup> Результаты сравнения с полноцветными изображениями можно посмотреть на <http://www.keldysh.ru/pages/cgraph/articles/index.htm>

Во всех случаях наш метод дал удовлетворительные результаты без дополнительной настройки параметров. Результаты некоторых из этих экспериментов приведены ниже.

На рис. 1 показаны четыре различных изображения “Stanford Memorial church” [7, 10]. Динамический диапазон для этого изображения превышает 250000:1. Левое верхнее изображение было получено Фаталлом (Fattall) [10], правое верхнее Тумблином и Турк (Tumblin and Turk’s LCIS method) [16], левое нижнее Вард Ларсоном и др. (Ward Larson et al) [22] и правое нижнее нашим методом.

На рисунках 2 and 3 показано аналогичное сравнение для изображений “streetlight on a foggy night” [17] и “Belgium House” [10]. Динамический диапазон для первого изображения превышает 100000:1, а для второго - 500000:1.

На рис. 4 показано сравнение для изображения “Nave” [7]. Левое изображение получено Рейнхардом (Reinhard) [12] и правое нашим методом.

Из сравнения видно, что наш метод дает хорошую проработку деталей, как для ярких, так и для темных частей изображения. Недостатком является некоторое снижение контраста, особенно заметное в очень ярких частях изображения. Наш метод избавлен от артефактов, характерных для LCIS метода. Так на изображении рис. 1, полученного LCIS методом, можно заметить повышенную шероховатость пола, а на изображении рис. 2 некоторую избыточную зернистость.

Кроме того, наш метод является также наиболее эффективным из рассмотренных методов. Обработка изображения размером 3600 на 2060 пиксела методом Рейнхарда [12] потребовала около 5 минут. Тогда как нашему методу понадобилось только 20 секунд для того же изображения и на том же компьютере. Т.е. наш метод быстрее метода Рейнхарда в 15 раз.

Наименьшая потеря деталей изображения при одновременном сохранении контраста и отсутствии артефактов достигнута, безусловно, в результате применения метода Domain Gradient из работы Фаталл и др. [10]. Одновременно этот метод также считается и наиболее быстрым из известных локальных методов. Тем не менее, время обработки весьма значительно. Оно составило от 1.1 секунды для изображения размером 512 на 384 пиксела и до 5 секунд для изображения размером 1130 на 751 пиксела на компьютере Pentium 4 1800 MHz [10].

По нашему методу обработка таких изображений составила 0.26 секунды для изображения размером 768 на 512 пиксела и 0.56 секунды для изображения размером 1025 на 769 пиксела на более слабом компьютере Pentium 3 866 MHz.

## 5. СПИСОК ЛИТЕРАТУРЫ

- [1] Dicarlo, J., and Wandell, B. 2000. Rendering high dynamic range images. *Proceedings of the SPIE: Image Sensors 3965*, 392–401.
- [2] Aggarwal, M., and Ahuja, N. 2001. High dynamic range panoramic imaging. In *Proc. IEEE ICCV*, vol. I, 2–9.
- [3] Schechner, Y. Y., and Nayar, S. K. 2001. Generalized mosaicing. In *Proc. IEEE ICCV*, vol. I, 17–24.
- [4] Aggarwal, M., and Ahuja, N. 2001. Split aperture imaging for high dynamic range. In *Proc. IEEE ICCV*, vol. II, 10–17.
- [5] Nayar, S. K., and Mitsunaga, T. 2000. High dynamic range imaging: Spatially varying pixel exposures. In *Proc. IEEE CVPR*.
- [6] SpheronVR. <http://www.spheron.com/>.
- [7] Debevec, P. E., and Malik, J. 1997. Recovering high dynamic range radiance maps from photographs. In *SIGGRAPH 97 Conference Proceedings*, AddisonWesley, T. Whitted, Ed., Annual Conference Series, ACM SIGGRAPH, 369–378.
- [8] Durand, F. and Dorsey, J. 2000. Interactive tone mapping. In *Eurographics Workshop on Rendering*, 219–230.
- [9] Durand, F. and Dorsey, J. 2002. Fast Bilateral Filtering for the Display of High-Dynamic-Range Images. *ACM Transactions on Graphics*, 21(3), July 2002 (Proceedings of SIGGRAPH 2002).
- [10] Fattall, R. Lischinski, D. and Werman, M. 2002 Gradient Domain High Dynamic Range Compression. *ACM Transactions on Graphics*, 21(3), July 2002 (Proceedings of SIGGRAPH 2002).
- [11] Ferwerda, J. 1998. Fundamentals of spatial vision. In *Applications of visual perception in computer graphics*. Siggraph '98 Course Notes.
- [12] Reinhard, E., Stark, M., Shirley, P., and Ferwerda, J., 2002. 'Photographic Tone Reproduction for Digital Images', *ACM Transactions on Graphics*, 21(3), July 2002 (Proceedings of SIGGRAPH 2002).
- [13] Reinhard, E., 2002, Parameter estimation for photographic tone reproduction. *Journal of graphics tools*, 7(1):45-52, 2002
- [14] Barladian B. Robust Parameter Estimation for Tone Mapping Operator, *Proceedings of Graphicon'2003*, Moscow, Russia, 2003.
- [15] Kang, S.B., Uyttendaele, M., Winder, S., Szeliski, R. High Dynamic Range Video. *Proceedings of SIGGRAPH 2003*.
- [16] Tumblin, J., and Turk, G. 1999. LCIS: A boundary hierarchy for detail-preserving contrast reduction. In *Siggraph 1999, Computer Graphics Proceedings*, Addison Wesley Longman, Los Angeles, A. Rockwood, Ed., Annual Conference Series, 83–90.
- [17] Tumblin, J., Hodgins, J. K., and Guenter, B. K. 1999. Two methods for display of high contrast images. *ACM Transactions on Graphics* 18, 1 (Jan.), 56–94.
- [18] Tumblin, J. 1999 *Three Methods For Detail-Preserving Contrast Reduction For Displayed Images*. PhD thesis, Georgia Institute of Technology, 1999.
- [19] Tumblin, J., and Rushmeier, H. 1993. Tone reproduction for computer generated images. *IEEE Computer Graphics and Applications* 13, 6 (November), 42–48.
- [20] Stevens, S.S and Stevens, J.C.. Brightness function: parametric effects of adaptation and contrast. *Journal of the Optical Society of America*, 50(11):1139A, November 1960.
- [21] <http://www.integra.jp>
- [22] Ward Larson, G., Rushmeier, H., and Piatko, C. 1997. A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Transactions on Visualization and Computer Graphics* 3, 4, 291–306.





Рис 1: Stanford Memorial Church. Левое верхнее изображение получено Gradient Domain методом [10], правое верхнее LCIS методом, левое нижнее Вард Ларсоном и др. и правое нижнее нашим методом.





Рис. 2: Streetlight on a Foggy Night. Левое верхнее изображение получено Gradient Domain методом [10], правое верхнее LCIS методом, левое нижнее Вард Ларсоном и др. и правое нижнее нашим методом.







Рис. 3: Belgium House. Левое верхнее изображение получено Gradient Domain методом [10], правое верхнее LCIS методом, левое нижнее Вард Ларсоном и др. и правое нижнее нашим методом.

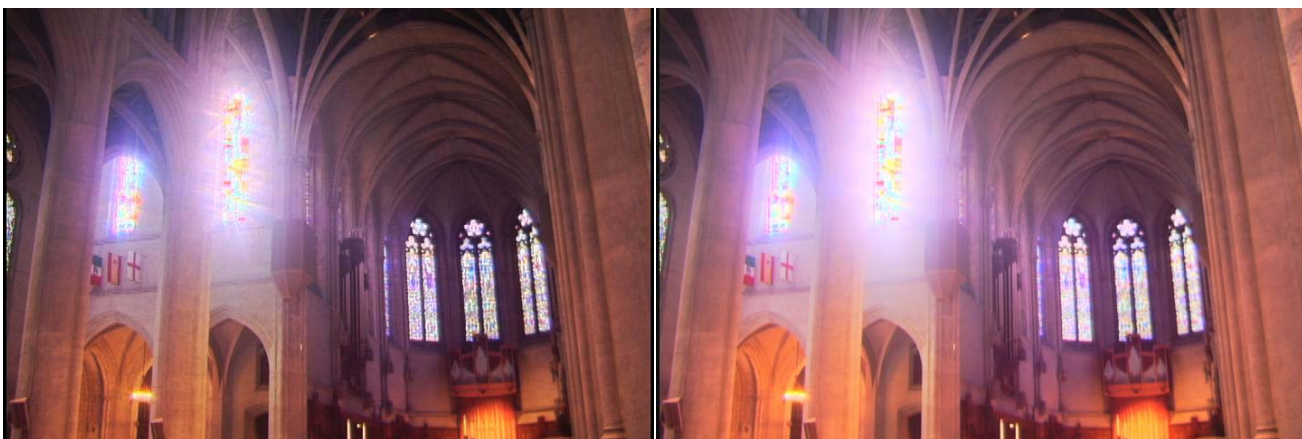


Рис. 4: Nave. Левое изображение получено Рейнхардом [12] и правое - нашим методом.