

Оценка качества методов масштабирования изображений и результаты сравнений разных методов

Дмитрий Ватолин, Сергей Путилин
Московский Государственный Университет, Лаборатория Компьютерной Графики ВМиК,
Москва, Россия
{dmitriy, sputilin}@graphics.cs.msu.su

Аннотация

В данной статье рассматривается два метода сравнения качества работы алгоритмов масштабирования, основанных на измерении PSNR, и приводится сравнение методов bilinear, precise bilinear, bicubic, precise bicubic, Lanczos3 (в реализации VirtualDub), G&M Lab Impress Normal (bicubic-like), G&M Lab Impress Sharp, G&M Lab Impress Very sharp, G&M Lab Impress Ultra sharp, G&M Lab Impress Best antialiasing, G&M Lab Impress Reduced ringing, G&M Lab Impress No Overshoot & No Ringing, Photoshop 7.0 bicubic rescaling.

В статье приводятся качественные и количественные результаты применения алгоритмов масштабирования, как на специальных тестовых изображениях, так и на изображениях, предоставленных для сравнения. Работы были выполнены по заказу компании Samsung.

Ключевые слова: масштабирование, rescaling measurement, PSNR, bilinear, bicubic, ringing, antialiasing, sharpness.

1. ВВЕДЕНИЕ

В последнее время для показа изображений и видео все чаще начинают применяться устройства с фиксированным разрешением, например, плазменные панели или LCD-дисплеи. Одновременно все больше распространяются HDTV телевизоры, также ориентированные на фиксированное разрешение, большее, чем разрешение фильмов, записанных на DVD или передаваемых по каналам вещания. Качество изображения на таких устройствах начинает существенно зависеть от качества используемых алгоритмов масштабирования.

Современный электронно-лучевые трубки позволяют изменять шаг и частоту вертикальной и горизонтальной развертки с достаточно мелким шагом. Как следствие, появляется возможность показывать экран дисплея в разных разрешениях без внесения существенных искажений в изображение.

Однако на дисплеях, основанных на жидких кристаллах (LCD), где базовая решетка пикселей фиксирована, приходится использовать аппаратное масштабирование. Как правило, оно производится алгоритмом билинейной интерполяции, из-за чего становится практически невозможно работать, например, на ноутбуках в разрешении, отличном от базового. Такие же проблемы возникают в проекторах и при показе изображения в PAL/NTSC на телевизорах высокой четкости (HDTV), например, на плазменных панелях.

Основные артефакты, которые возникают при масштабировании таковы:

- 1) Ringing – возникновение волны около резкой границы на изображении.
- 2) Overshooting - возникновение 2 и 3 волн.
- 3) Aliasing – “лестничный эффект” – неравномерности изображения на резких диагональных границах изображения.
- 4) Unsharpening – размывание – недостаточная четкость изображения после масштабирования. Очевидно, что повышение четкости, как правило, приводит к увеличению остальных артефактов и наоборот – подавление артефактов подавляет также и четкость.
- 5) Sub-pixel shift – субпиксельный сдвиг изображения, связанный, как правило, с особенностями реализации алгоритма. Практически не влияет на визуальное качество, однако существенно влияет на формальную метрику.

Классический подход к оценке качества изображений основан на спектрах [1,2] и в виде поправки к PSNR метрике был реализован в рамках этой работы. Однако эксперименты показали, что при малых изменениях изображений поправка на спектральную составляющую [6] становится несущественна.

Также были рассмотрены работы формализации визуального анализа алгоритмов масштабирования [3] с использованием метрик, дающих разумные результаты для достаточно низкокачественных методов. К сожалению, в нашем случае (тонкая доводка алгоритмов) соответствующие методы были слабо применимы. В результате был предложен достаточно простой метод анализа, основанный на наборе тестовых изображений.

Данная работа является частью проекта по созданию алгоритма масштабирования, дающего четкое изображение и, по возможности, избавленного от артефактов других методов. Разработанная методика измерений использовалась для подбора оптимальных параметров алгоритмов. Построение самого семейства оптимальных алгоритмов масштабирования будет являться темой отдельной статьи.

2. МЕТОДИКА ОЦЕНКИ АЛГОРИТМОВ МАСШТАБИРОВАНИЯ

Специально для сравнения алгоритмов масштабирования было построено 7 искусственных векторных изображений, обычно называемых «мирами».

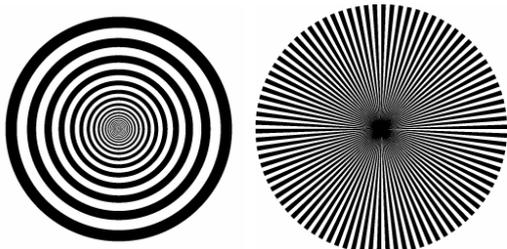


Рисунок 1. Тестовые изображения 1 и 3 для оценки качества приближения контрастных границ, ginging и aliasing на них

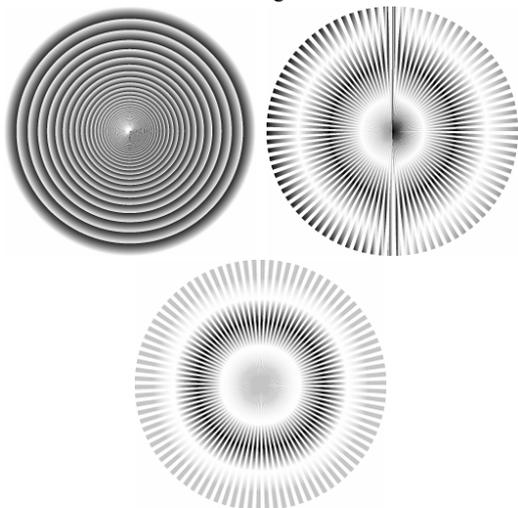


Рисунок 2. Тестовые изображения 2, 4 и 5 для оценки систематического субпиксельного сдвига в разных направлениях в разных частях изображений и оценки приближения неконтрастных деталей

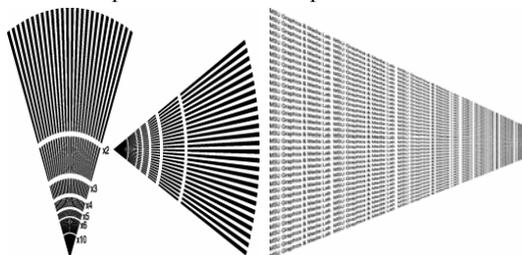


Рисунок 3. Тестовые изображения 6 и 7 для оценки качества передачи небольших элементов. Седьмое изображение представляет собой надпись «MSU Graphics&Media Lab», повторенную много раз и показанную с перспективой.

Соответствующие векторные изображения были получены в следующих разрешениях: 256x256, 512x512, 800x800, 1024x1024.

Помимо 7 векторных изображений были также использованы 9 фотографий высокого качества (1024x1024). Ниже приведен пример двух фотографий:



Рисунок 4. Тестовые изображения 10 и 16. Использовались для оценки сохранения мелких деталей и диагональных границ в реальных изображениях

Соответствующие векторные изображения масштабировались вверх и вниз в кратное и некратное количество раз между полученными ранее разрешениями. Для оценки качества масштабирования использовалась метрика Y-PSNR. Т.е. классическая PSNR, применяемая к яркостной составляющей изображения:

$$d(x,y) = 10 \cdot \log_{10} \frac{255^2 \cdot n^2}{\sum_{i=1, j=1}^{n,n} (x_{ij} - y_{ij})^2}$$

Авторы понимают все минусы L_2 мер [4, 5], однако здесь измерялись именно минимальные отличия в изображениях, на которых L_2 адекватна. Для визуального сравнения также использовалась PSNR, но уже в виде LUV-PSNR, которая учитывает также и цветовые искажения, вносимые при масштабировании в изображение.

Оценка качества масштабирования фотографий проводилось по другой методике. Они масштабировались вниз, а потом восстанавливались вверх и далее измерялась Y-PSNR между исходным и получившимся изображением. Поскольку методы масштабирования, разработанные в лаборатории компьютерной графики ВМиК МГУ дают разные результаты при масштабировании вверх и при масштабировании вниз (что хорошо видно на графика масштабирования «мир»), также было разработано несколько методов, когда масштабирование вниз осуществлялось одним методом, а масштабирование вверх – другим. Использовались следующие варианты преобразований:

- 1024x1024 to 256x256 to 1024x1024
- 1024x1024 to 300x300 to 1024x1024
- 1024x1024 to 512x512 to 1024x1024

Были протестированы следующие методы:

1. VirtualDub bilinear
2. VirtualDub precise bilinear
3. VirtualDub bicubic
4. VirtualDub precise bicubic (A=-1.00)
5. VirtualDub Lanczos3
6. G&M Lab Impress Normal (bicubic-like)
7. G&M Lab Impress Sharp
8. G&M Lab Impress Very sharp
9. G&M Lab Impress Ultra sharp

10. G&M Lab Impress Best antialiasing
11. G&M Lab Impress Reduced ringing
12. G&M Lab Impress No Overshoot & No Ringing
13. Photoshop 7.0 bicubic rescaling

G&M Lab Impress – семейство методов, разработанных в нашей лаборатории.

В идущей ниже таблице представлено число замеров для различных типов тестов. Было использовано 7 изображений «мир» и 9 фотографий разного типа разрешением 1024x1024.

Type of test	Number of images	Num. of measure.
ScaleUp 256→1024	7 miras × 13 methods	91
ScaleUp 512→800	7 miras × 13 methods	91
ScaleUp 512→1024	7 miras × 13 methods	91
ScaleUp 800→1024	7 miras × 13 methods	91
ScaleDown 512→256	7 miras × 13 methods	91
ScaleDown 800→512	7 miras × 13 methods	91
ScaleDown 1024→256	7 miras × 13 methods	91
ScaleDown 1024→800	7 miras × 13 methods	91
ScaleDownUp 1024→256→1024	16 images × 18 methods	288
ScaleDownUp 1024→300→1024	16 images × 18 methods	288
ScaleDownUp 1024→512→1024	16 images × 18 methods	288
Total measurements		1592

Всего было проведено более 2000 измерений. Результатирующими были 1592 замера, чьи изображения метрик и сами изображения (3184 картинок) заняли порядка 9Гб. Ниже приведены части фрагменты изображений, показывающие разницу между алгоритмами масштабирования и соответствующие им метрики.

3. РЕЗУЛЬТАТЫ

При разработке достаточно сложного алгоритма масштабирования, зависящего от многих параметров, основной проблемой является определение оптимальных параметров алгоритма (оптимизация в n-мерном пространстве). При этом фактически невозможно подбирать параметры вручную. Начиная с какого-то момента изменения в изображениях становятся практически не видны на глаз, и, кроме того, приходится отсматривать большое количество изображений, в силу того, что хорошие результаты на одном изображении зачастую не гарантируют высокого качества на другом (не все зависимости в сложном алгоритме можно предсказать).

Поэтому основной задачей данной работы было предоставить удобный механизм для визуального и формального анализа наборов параметров (пресетов) алгоритма.

3.1 Масштабирование вверх

На приведенных ниже графиках значение PSNR метрики приведено относительно соответствующего значения для Photoshop bicubic — весьма аккуратной реализации масштабирования. Это позволяет учесть поправку на разный характер изображений, проявляющийся в существенно отличающихся абсолютных значениях метрики.

Из-за ограничений на размер приведены только два графика из всех вариантов масштабирования вверх.

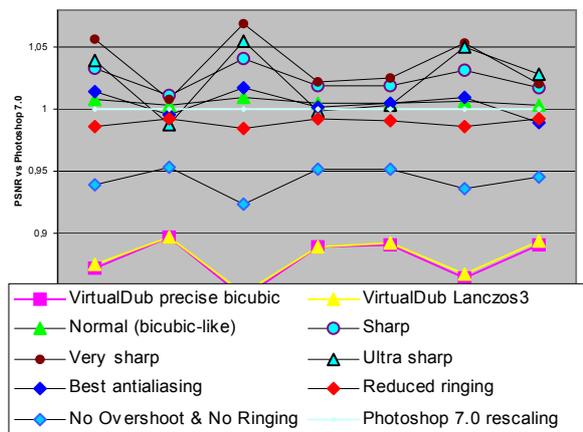


Рисунок 5. Измерения слева направо на графиках соответствуют номерам изображений, приведенным выше. Масштабирование 0512 to 1024

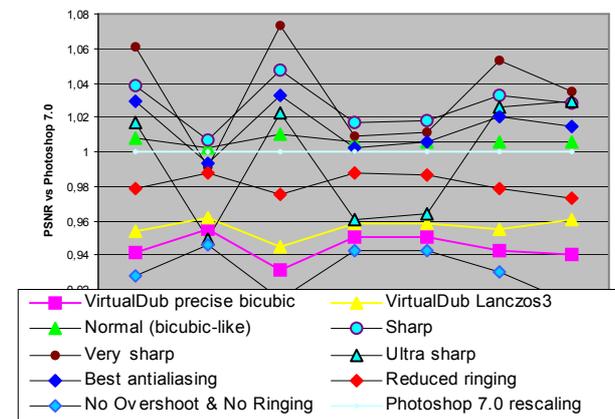


Рисунок 6. Масштабирование 0800 to 1024

В целом можно сделать следующие выводы:

Вариант алгоритма “Very sharp” дает стабильно лучшие результаты, по сравнению с Photoshop’s bicubic на всех разрешениях и почти для всех изображений. “Ultra sharp” показывает хорошие результаты на контрастных изображениях.

Замеры методов масштабирования из программы VirtualDub’s показывают сравнительно низкие результаты из-за вносимого при реализации алгоритма сабпиксельного сдвига. На визуальных сравнениях ниже будет видно достаточно высокое качество масштабирования для Lanczos3.

“Normal” показывает стабильно хорошие результаты, немного лучшие, чем у Photoshop в силу лучшего подавления лестничного эффекта.

3.2 Масштабирование вниз

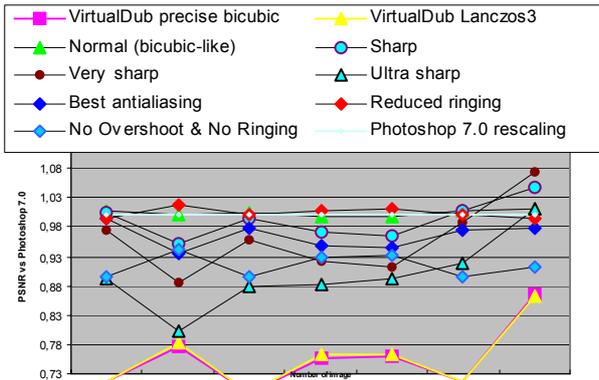


Рисунок 7. Масштабирование 1024 to 0256

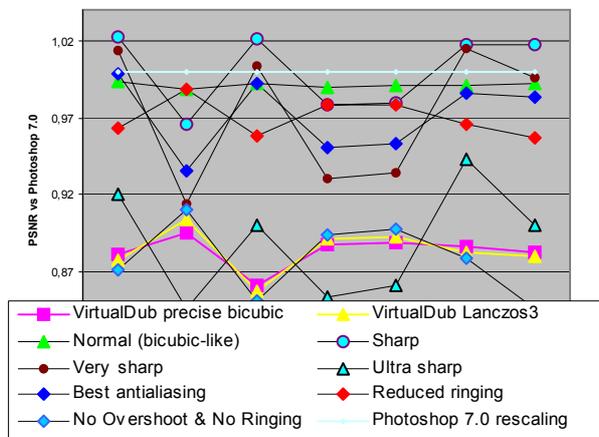


Рисунок 8. Масштабирование 1024 to 800

Здесь также приведены не все измерения, а только два наиболее отличающихся варианта.

“Reduced ringing” показал стабильно высокий результат. “Very sharp”, “Ultra sharp” and “Normal” пресеты показали худшую производительность, чем Photoshop, однако они дают визуально заметно лучшие изображения. Причина в том, что метод растеризации изображений «мир» был близок к растеризатору Photoshop.

3.3 Проблемы масштабирования в одну сторону

Некоторые алгоритмы масштабирования дают сдвиг картинки по диагонали. Этот эффект влияет как на PSNR, так и на визуальную метрику – для симметричных картинок она становится несимметричной (см. рисунок 12). Сдвиг кратный размеру пиксела при измерениях можно достаточно легко компенсировать, однако в большинстве случаев встречается некратный сдвиг. Компенсировать этот эффект можно применяя попеременно масштабирование вниз и вверх.

Кроме того, если все замеры проводить на векторных изображениях, то есть опасность подобрать параметры для алгоритма масштабирования, которые будут оптимальными именно для контрастных векторных изображений. Поэтому проводились также замеры на реальных фотографиях.

3.4 Измерение масштабирования фотографий

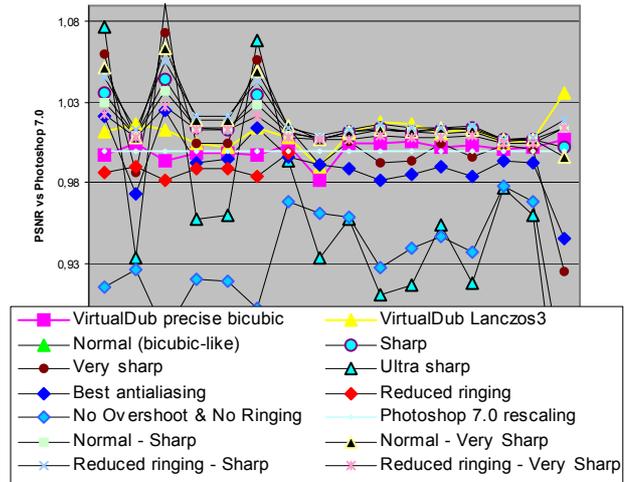


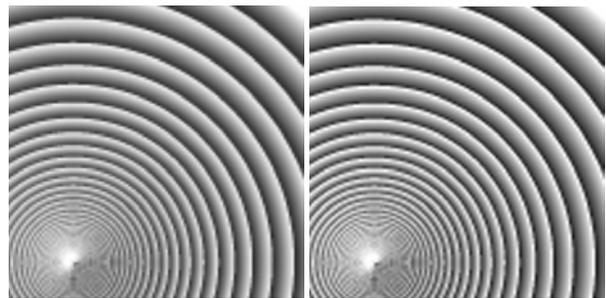
Рисунок 9. Масштабирование 1024-256-1024

Комбинация “Reduced ringing” для уменьшения и “Sharp” для увеличения дает стабильно хороший результат. Алгоритм Lanczos3 дает хороший результат на фотографиях, однако хуже работает на векторной графике. Комбинация “Normal” и “Sharp” показывает хорошие результаты на контрастных изображениях, но работает хуже чем “Reduced ringing” и “Sharp” на фотографиях.

3.5 Визуальное сравнение методов на мирах

Визуальная метрика строилась по попиксельным значениям PSNR и приводится в условных цветах. Черный означает совпадение изображений, синий – незначительные отличия, зеленый и желтый – большие отличия, а красный – хорошо видимые на глаз отличия. Благодаря логарифмической шкале PSNR условные цвета достаточно хорошо коррелируют с визуальной разницей между изображениями.

Для того чтобы было лучше видно разницу, приводятся только фрагменты изображений.



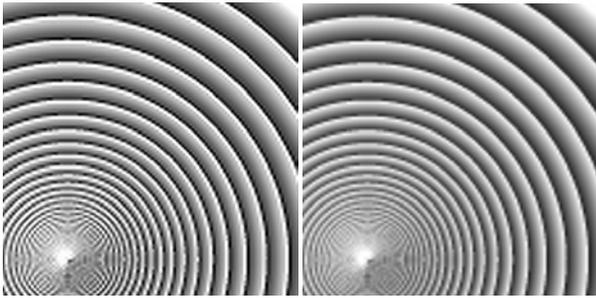


Рисунок 10. Тестовое изображение 2, масштабирование 256->1024, VirtualDub bilinear, Photoshop bicubic, Impress Ultra Sharp, Impress Reduced Ringing.

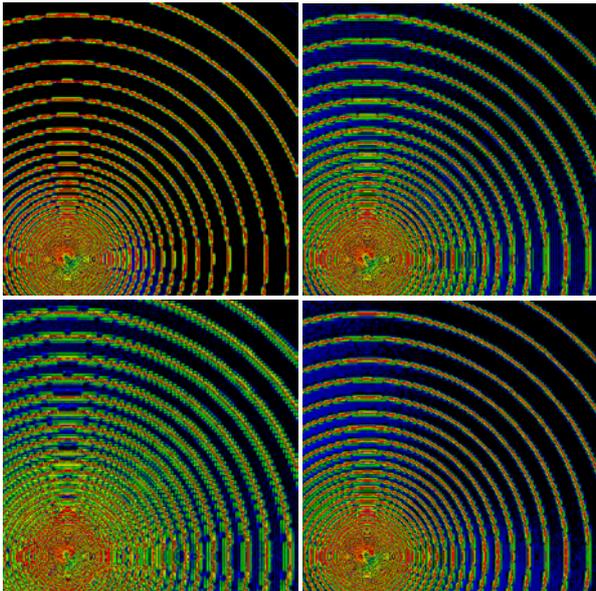


Рисунок 11. Метрики для тестового изображения 2, масштабирование 256->1024, VirtualDub bilinear, Photoshop bicubic, Impress Ultra Sharp, Impress Reduced Ringing.

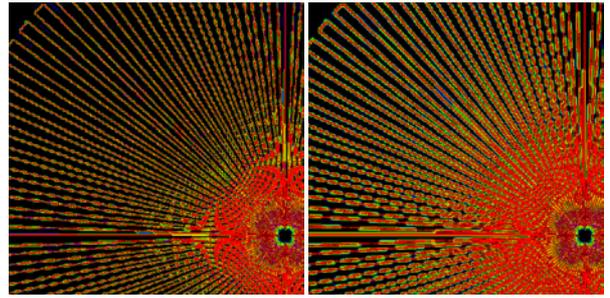
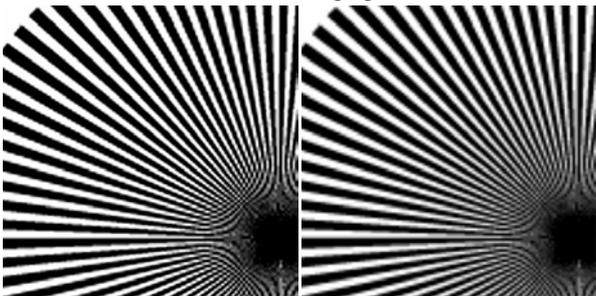


Рисунок 12. Изображения и метрики для изображения 3, масштабирование 256->1024, Impress Very Sharp, VirtualDub bilinear. Поскольку есть риск, что после печати не будет видна разница даже с ним.

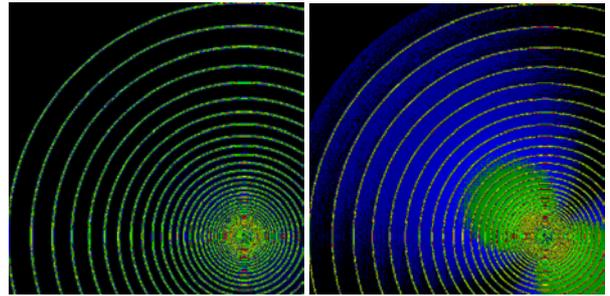


Рисунок 13. Изображения и метрики для изображения 2, масштабирование 512->1024, Impress Normal, VirtualDub bilinear.

Для сравнения приведем хотя бы один пример масштабирования вниз. Хорошо видно, в каких областях изображения меньше метрика у алгоритма «No overshoot & no ringing» по сравнению с «Lanzcos3»:

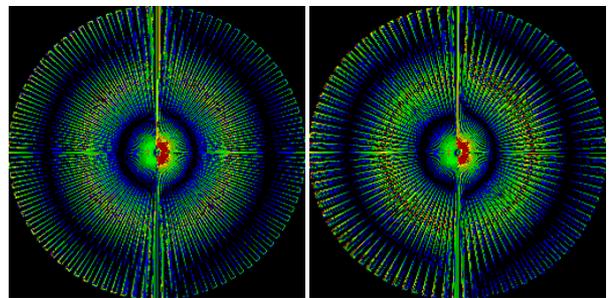


Рисунок 14. Изображения и метрики для изображения 4, масштабирование 1024->256, Impress No overshoot & no ringing, VirtualDub Lanczos3.

3.6 Визуальное сравнение методов на фотографиях

К сожалению, из 1600 фотографий и метрик здесь приведены только три фрагмента, показывающие наиболее заметные на глаз отличия между методами. Суть сравниваемых методов масштабирования будет изложена в другой статье, поэтому необходимо сказать, что «группа Sharp» методов неэквивалентна предварительному применению фильтра Sharp к изображению именно потому, что они дают меньшие ringing-эффекты.



Рисунок 15. Изображение 10, масштабирование 1024->256->1024, VirtualDub precise bilinear.



Рисунок 16. Изображение 10, масштабирование 1024->256->1024, Photoshop bicubic.



Рисунок 17. Изображение 10, масштабирование 1024->256->1024, Impress Very Sharp.

4. ЗАКЛЮЧЕНИЕ

Предлагаемые в статье методы показали себя как удобный инструмент, позволяющий достаточно просто массово тестировать алгоритмы масштабирования и улучшать их параметры.

На данный момент в процессе разработки находится следующий вариант алгоритма масштабирования. У него несколько выше вычислительная сложность, однако, его

характеристики на резких границах (т.е. на всех мирах) будут стабильно выше при низкой ступенчатости границ.

5. БЛАГОДАРНОСТИ

Авторы выражают свою признательность своим коллегам Андрею Крылову, Алексею Лукину и Анастасии Куликовой за работу по другим метрикам, реализацию алгоритмов масштабирования и множество ценных идей, а также компании Samsung за любезное разрешение опубликовать эту часть работы.

6. ЛИТЕРАТУРА

- [1] S. Haikin. *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, NJ, second edition, 1991
- [2] Terrance E. Boulton and George Wolberg, *Local Image Reconstruction and Subpixel Restoration Algorithms*, Computer Vision, Graphics, and Image Processing. Graphical Models and Image Processing, 55, N1, pp. 63-77, 1993,
- [3] Richard Harvey, Stephen King, Richard Aldridge and J. Andrew Bangham, *Comparing Image Resamplers Via a Model of the Human Vision System*, British Machine Vision Conference http://scooby.sys.uea.ac.uk/papers/bmvc97_2.ps.gz
- [4] Bernd Girod. *What's wrong with mean-squared error?* In Andrew B. Watson, editor, *Digital Images and Human Vision*, pages 207–220. MIT Press, 1993.
- [5] H.Marmolin. Subjective mse measures. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-16 (3): 486–489, May/June 1986.
- [6] Christian J. van den Branden Lambrecht and Joyce E. Farrel. *Perceptual quality metric for digitally coded color images*. In *Proceedings European conference on signal processing*, 1996.

Об авторах

Дмитрий Ватолин — к.ф.-м.н., старший научный сотрудник лаборатории компьютерной графики ВМиК МГУ. Основатель сервера «Все о сжатии» <http://compression.graphicon.ru/>

E-mail: dmitriy@graphics.cs.msu.su

Сергей Путилин — студент, сотрудник лаборатории компьютерной графики ВМиК МГУ.

E-mail: sputilin@graphics.cs.msu.su