

Способ оценки четкости фотореалистичных изображений с высоким разрешением

С.В. Сай¹

¹ Тихоокеанский государственный университет, Хабаровск, Россия

Аннотация

В статье предлагается способ оценки четкости фотореалистичных изображений, основанный на сравнении коэффициента детальности оригинального и искаженного изображений. В алгоритме идентификации мелких структур оригинального изображения используются операции сегментации активных пикселей, к которым относятся точечные объекты, тонкие линии и фрагменты текстуры. Количество активных пикселей оценивается значением коэффициента детальности, которое определяется отношением активных пикселей к общему количеству пикселей изображения. Этот же алгоритм используется для вычисления значения коэффициента детальности искаженного изображения, и далее оценивается снижение четкости с помощью сравнения полученных значений. К особенностям способа относится то, что идентификация мелких структур и сегментация активных пикселей выполняется в нормированной системе N-CIELAB. Также в алгоритме учитывается влияние ложных микроструктур на результаты оценки реставрированного изображения. Рассматриваются особенности построения нейронных сетей SRCNN в задачах качественного повышения разрешения изображения с восстановлением мелких структур. Приводятся результаты анализа качества увеличенных изображений по традиционным метрикам PSNR и SSIM, а также по предлагаемому способу.

Ключевые слова: анализ изображений, суперразрешение, мелкие структуры, метрика искажений.

Цитирование: Сай, С.В. Способ оценки четкости фотореалистичных изображений с высоким разрешением / С.В. Сай // Компьютерная оптика. – 2022. – Т. 46, № 1. – С. 121-129. – DOI: 10.18287/2412-6179-CO-899.

Citation: Sai SV. A method for assessing photorealistic image quality with high resolution. Computer Optics 2022; 46(1): 121-129. DOI: 10.18287/2412-6179-CO-899.

Введение

В современных интеллектуальных видеосистемах проблема сохранения качества передаваемых деталей изображения является одной из основных проблем при решении задач поиска, распознавания и идентификации объектов.

Цифровые изображения подвержены воздействию различных искажений, которые появляются в процессе формирования, цифровой обработки и передачи сигналов. Основные искажения качества возникают за счет использования сжатия с потерями, что обусловлено необходимостью передачи сигналов по каналам связи с ограниченной пропускной способностью, а также компактным хранением больших объемов видеоданных.

Четкость относится к основным показателям качества изображения и зависит от разрешающей способности (разрешения), которая традиционно оценивается количеством различимых глазом штриховых линий тестового изображения (таблицы). Увеличение количества пикселей формата видеокadra до стандартов высокой четкости (HD, 4K и 8K) приводит к повышению разрешения изображения, но при этом для обеспечения необходимой скорости передачи требуется применение более эффективных методов сжатия. В

процессе сжатия формат кадра не меняется, но при высоких коэффициентах сжатия «размываются» или «теряются» мелкие детали, что в сильной степени ухудшает визуальное качество изображения. Ухудшение четкости, в свою очередь, приводит к снижению резкости границ крупных деталей изображения.

В целом, задачу восстановления или повышения четкости изображения можно решать следующими способами [1]: повышением качественных характеристик методов и алгоритмов сжатия; применением корректирующих фильтров; использованием технологий искусственного интеллекта (AI) на основе нейронных сетей.

В настоящее время существует достаточно много нейросетевых методов, позволяющих улучшать изображения. Задача улучшения четкости изображения заключается в том, чтобы из изображения с низким качеством с помощью нейронной сети получилось изображение высокого качества с восстановленными мелкими структурами. Методы, позволяющие улучшать изображения, основываются на сверточных нейронных сетях (CNN).

Методы, направленные на качественное увеличение разрешения исходного изображения, называются Super-Resolution (SR), что означает суперразрешение. Изображение может иметь «более низкое

разрешение» из-за меньшего пространственного разрешения (то есть размера) или из-за ухудшения четкости (размытие мелких деталей). Современные методы SR основаны на следующих подходах [2]: предварительное увеличение разрешения; прогрессивное увеличение разрешения; итеративное понижение и повышение разрешения.

Для многих подходов часто применяются попиксельные функции потерь (например, MSE), которые пытаются справиться с неопределённостью при восстановлении текстуры изображения. Минимизация ошибки позволяет находить средние вероятностные решения для каждого пикселя, но они получаются слишком гладкими и, следовательно, имеют низкое качество восприятия. Основными критериями качества [3–4] изображения в задаче SR являются пиковое отношение сигнала к шуму (PSNR) и индекс структурного сходства (SSIM).

В настоящее время имеются программные продукты, с помощью которых можно выполнять улучшение качества и масштабирование фотоизображений на основе методов SR. К наиболее популярным относятся [5]: AI Image Enlarger, Let's Enhance, BigJPG, Topaz Gigapixel AI, Deep Image, ImageUpscaler, Waifu2x, Upscale Pics, PhotoLemur.

К недостаткам использования нейронных сетей относится следующее: методы SR глубокого обучения нельзя применять универсально ко всем типам изображений; реконструкция изображений может занимать длительное время; критерии PSNR и SSIM не всегда соответствуют зрительным оценкам качества фотореалистичных изображений.

В предыдущих авторских работах качество компрессированных изображений оценивалось по разработанной метрике искажений мелких структур (MFSD). В настоящей работе предлагается оригинальный способ оценки качества изображения на основе измерения его четкости в цветовом пространстве N-CIELAB. Отличие предлагаемого способа в том, что в нем измеряется ухудшение четкости, а не искажение любых мелких структур, что является перспективным продолжением исследований метода оценки качества изображений применительно к системам SR.

1. Способ оценки четкости

Четкость традиционно оценивается количеством различимых глазом штриховых ТВ-линий тестового изображения [6–7]. Потенциальная разрешающая способность (разрешение) – это способность камеры оптически воспроизводить самые мелкие детали, она обычно указывается как однозначный показатель, например, форматом кадра или количеством пикселей. Для измерения четкости по рекомендациям стандарта ISO-12233:2017 [8] используется пространственная частотная характеристика (SFR), которая определяется как функция спада контраста мелких деталей от пространственной частоты f_s . В дополне-

ние к SFR используется характеристика спада модуляционной функции (MTF) синусоидального сигнала с увеличением пространственной частоты.

Как правило, контраст уменьшается с увеличением пространственной частоты до уровня, при котором мелкие детали больше не различаются глазом. Это предельное значение частоты и определяет реальное разрешение (четкость) изображения, полученное с камеры или после его сжатия и реставрации. Таким образом, разрешение оценивается самой высокой пространственной частотой, на которой контраст мелких деталей ещё различается глазом.

Пространственная частота, на которой $SFR=0,1$, что означает спад MFT до 10 %, в соответствии с критерием Рэля [9] используется для измерения разрешения в вертикальном, горизонтальном и диагональном направлениях.

Для дискретных изображений значение пространственной частоты f_s можно измерить отношением количества циклов (периодов) на пиксель (cycles/pixel). Значение $f_s=0,5$ соответствует половине частоты дискретизации и принимается как верхний предел SFR, что соответствует потенциальному разрешению цифровой камеры в пикселях.

Для оценки реального разрешения цифровых камер [10] используется отношение измеренной частоты к значению 0,5. Например, если $f_s=0,4$ в вертикальном направлении, то реальную четкость можно оценить величиной $E_v=0,4/0,5=80\%$. Это будет означать, что визуальная четкость изображения по вертикали уменьшилась на 20 %.

Для оценки суммарного разрешения по всем трем направлениям в стандарте ISO-12233:2017 рекомендовано использовать следующую формулу [11]:

$$E_s = \frac{E_d(E_v + E_h)}{2} 100\%, \quad (1)$$

где E_v , E_h и E_d – параметры оценки четкости в вертикальном, горизонтальном и диагональном направлениях. Очевидно, чем ближе значение E_s к 100 %, тем выше четкость визуального изображения. Отметим, что формула (1) обычно используется для оценки разрешения по сигналу яркости Y , так как известно [12], что контрастная чувствительность зрения более высокая к изменениям сигналов яркости и менее высокая к сигналам цветности U и V .

На практике разрешение камеры и спад SFR определяются рядом факторов. К ним относятся: характеристики объектива камеры, количество адресуемых фотоэлементов в оптическом устройстве формирования изображения и электронные схемы в камере, которые могут включать в себя функции сжатия изображения и гамма-коррекции.

Для измерения SFR используется два метрологических метода – на основе выделения границ и на основе синусоидальных сигналов [8]. Для обоих методов используются специальные тестовые таблицы [13],

в изображениях которых присутствуют необходимые объекты для измерения.

В первом методе в изображении таблицы выделяются области интереса (ROI) рядом с наклонными вертикальными и горизонтальными границами и далее используются для вычисления уровней SFR. Использование наклонных границ с низким контрастом позволяет измерять градиент на многих фазах и учитывать влияние артефактов, возникающих за счет нелинейной обработки изображений.

Второй метод метрологии SFR основан на анализе характеристик изображения синусоидальных волн. В качестве тестового образца в стандарте ISO 12233:2017 рекомендовано использовать изображение с синусоидальной модуляцией в полярном формате (например, Siemens Star). С помощью соответствующего программного обеспечения, например, Imatest [14], можно рассчитать SFR на основе синусоидальной волны с точностью до половины частоты дискретизации.

К особенностям стандартных способов относится то, что они в основном используются для калибровки цифровых камер и основаны на анализе искажений изображений тестовых таблиц. При этом контраст элементов тестовой таблицы является фиксированной величиной и обычно задается высоким значением.

Для фотореалистичного изображения такие способы не подходят, так как объекты в виде идеальных границ или штриховых линий, образованных синусоидальными волнами, не являются обязательной частью изображения. Также контраст мелких структур в отличие от элементов тестовых таблиц не является максимальным или постоянным фиксированным значением, а может принимать любое значение.

В настоящей работе предлагается альтернативный способ оценки четкости изображения, основанный на алгоритмах поиска, идентификации и оценки структурного подобия самых мелких деталей, к которым относятся точечные объекты, фрагменты тонких линий или текстур.

В [15] описан способ классификации изображений по уровню детальности (FDL) с применением метода оценки цветовых различий в нормированной колориметрической системе N-CIELAB. Алгоритм идентификации мелких структур оригинального изображения основан на результатах анализа структуры микроблоков с размером 3×3 пикселя. Для такого микроблока имеем шесть цветовых переходов по горизонтали и шесть цветовых переходов по вертикали, т.е. всего 12 переходов.

Мелкие структуры в микроблоке можно разделить на следующие типы: точечные объекты, фрагменты тонких линий, фрагменты контуров, фрагменты текстур. Максимальное количество цветовых переходов (n) между соседними пикселями внутри микроблока будет равно $n = 12$ (фрагмент текстуры – «шахматное поле»). Минимальное количество цветовых переходов равно $n = 2$.

Идентификация микроструктур выполняется по следующему алгоритму. На первом этапе изображение разбивается на микроблоки с размером 3×3 пикселя. В каждом микроблоке t вычисляются двенадцать значений контраста между соседними пикселями:

$$K_{m,p} = \sqrt{\left(\frac{\Delta L_{m,p}^*}{L_{th}^*}\right)^2 + \left(\frac{\Delta a_{m,p}^*}{a_{th}^*}\right)^2 + \left(\frac{\Delta b_{m,p}^*}{b_{th}^*}\right)^2}, \quad (2)$$

где значения $\Delta L_{m,p}^*$, $\Delta a_{m,p}^*$ и $\Delta b_{m,p}^*$ определяют контраст между соседними пикселями по светлоте и по цветности для цветового перехода с номером p ; L_{th}^* , a_{th}^* и b_{th}^* – значения весовых коэффициентов для самых мелких структур размером в один пиксель. На каждом шаге (p) вычислений (2) проверяем условие

$$K_{m,p} > 1, \quad (3)$$

при выполнении которого принимаем решение о том, что изменение контраста между соседними пикселями заметно глазом и увеличиваем значение счетчика (n) обнаруженных цветовых переходов на единицу. После анализа всех цветовых переходов в микроблоке выполняем его идентификацию по следующему критерию. Если количество переходов с цветовым контрастом (3) больше единицы, то в микроблоке присутствует мелкая структура, различимая глазом. Если данное условие не выполняется, то мелкая структура не различается глазом.

После анализа всего изображения получим выделенные участки изображения с различимыми мелкими структурами в виде микроблоков с размерами 3×3 пикселя. Количество таких микроблоков обозначим N_m . Остальные участки классифицируем как микроблоки с равномерными изменениями цветовых координат и обозначим их количество символом N_F . По значению N_m вычисляем уровень детальности (FDL) изображения, который оценивает процент участков изображения с мелкими структурами

$$FDL = \frac{N_m}{(N_m + N_F)} 100\%. \quad (4)$$

В настоящей работе предлагается способ оценки детальности изображения, основанный на алгоритме идентификации активных пикселей.

Пиксель с пространственными координатами (i, j) будет считаться активным, если его цветовой контраст $\Delta K_{i,j}$ относительно двух соседних пикселей превышает нормированный пороговый контраст зрения. В качестве соседних пикселей выбираются следующие комбинации: 1) слева-справа по горизонтали, 2) сверху-снизу по вертикали, 3) снизу-сверху по диагонали и 4) сверху-снизу по диагонали.

Условие идентификации активного пикселя запишем в следующем виде:

$$(\Delta K_{i,j}^{(1)} > 1) \vee (\Delta K_{i,j}^{(2)} > 1) \vee (\Delta K_{i,j}^{(3)} > 1) \vee (\Delta K_{i,j}^{(4)} > 1), \quad (5)$$

где знак (\vee) означает логическую операцию «ИЛИ».

Значения цветового контраста между соседними пикселями определяются уравнением:

$$\Delta K_{j,i} = \sqrt{\left(\frac{\Delta L_{j,i}^*}{L_{th}^*}\right)^2 + \left(\frac{\Delta a_{j,i}^*}{a_{th}^*}\right)^2 + \left(\frac{\Delta b_{j,i}^*}{b_{th}^*}\right)^2}, \quad (6)$$

где $\Delta L_{j,i}^*$, $\Delta a_{j,i}^*$ и $\Delta b_{j,i}^*$ – значения контраста по яркости (светлоте) и по цветности; i и j – координаты анализируемого пикселя по горизонтали и по вертикали.

В отличие от предыдущего алгоритма, в новом алгоритме используется принцип скользящего окна с размером 3×3 пикселя, в центре которого идентифицируется активный пиксель.

Для оценки первого условия $\Delta K_{j,i}^{(1)} > 1$ потребуется выполнение следующих операций.

1. Вычисление двух значений контраста анализируемого пикселя относительно левого $\Delta K_{j,i}^{(1,1)}$ и правого $\Delta K_{j,i}^{(1,2)}$ пикселя. Для вычисления $\Delta K_{j,i}^{(1,1)}$ в (6) подставляются значения: $\Delta L_{j,i}^* = L_{j,i}^* - L_{j,i-1}^*$,

$$\Delta a_{j,i}^* = a_{j,i}^* - a_{j,i-1}^* \text{ и } \Delta b_{j,i}^* = b_{j,i}^* - b_{j,i-1}^*.$$

Для вычисления $\Delta K_{j,i}^{(1,2)}$ подставляются значения:

$$\Delta L_{j,i}^* = L_{j,i}^* - L_{j,i+1}^*, \quad \Delta a_{j,i}^* = a_{j,i}^* - a_{j,i+1}^* \quad \text{и}$$

$$\Delta b_{j,i}^* = b_{j,i}^* - b_{j,i+1}^*.$$

2. Проверится условие превышения полученных значений контраста единичного порога:

$$(\Delta K_{i,j}^{(1,1)} > 1) \wedge (\Delta K_{i,j}^{(1,2)} > 1), \quad (7)$$

где знак (\wedge) означает логическую операцию «И». Также проверяется условие изменения градиента по яркости

$$(L_{j,i-1}^* < L_{j,i}^* > L_{j,i+1}^*) \vee (L_{j,i-1}^* > L_{j,i}^* < L_{j,i+1}^*), \quad (8)$$

выполнение которого обеспечивает положительный или отрицательный контраст анализируемого пикселя.

3. Если условия (7) и (8) выполняются, принимается решение о том, что первое условие в (5) истинно и анализируемый пиксель принимает статус «активный».

Если условия (7) и (8) не выполняются, переходим к оценке второго условия $\Delta K_{j,i}^{(2)} > 1$, где для вычисления $\Delta K_{j,i}^{(2,1)}$ используется соседний пиксель по вертикали с координатами $(j-1, i)$ и для вычисления $\Delta K_{j,i}^{(2,2)}$ соседний пиксель с координатами $(j+1, i)$. Далее проверяются условия (7) и (8).

Аналогично вычисляются значения контраста анализируемого пикселя относительно соседних пикселей по диагонали: $\Delta K_{j,i}^{(3,1)}$, $\Delta K_{j,i}^{(3,2)}$ и $\Delta K_{j,i}^{(4,1)}$, $\Delta K_{j,i}^{(4,2)}$ и проверяются условия $\Delta K_{j,i}^{(3)} > 1$ и $\Delta K_{j,i}^{(4)} > 1$.

После проверки всех четырех условий в соответствии с (5) идентифицируется активный пиксель по критерию выполнения одного или более условий. При этом в изображении маркируются пиксели в окне (3×3) с координатами центрального пикселя (i, j) .

Если выражение (5) будет ложным, анализируемому пикселю присваивается статус «пассивный». Это будет означать, что такой пиксель не будет различаться глазом на фоне соседних пикселей.

После анализа всего изображения выполняется подсчет маркированных пикселей N_a и вычисляется коэффициент детальности:

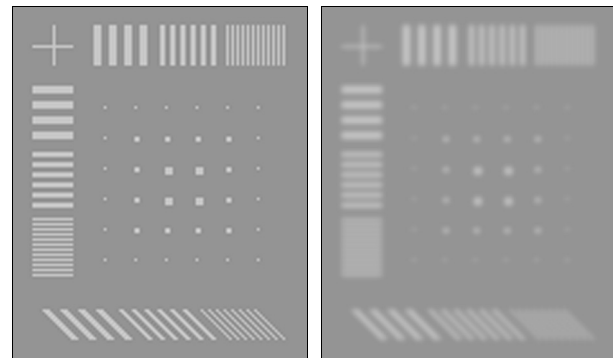
$$FDL_a = \frac{N_a}{W \cdot H} \cdot 100\%, \quad (9)$$

который оценивает относительное количество мелких деталей изображения в процентах.

Отметим, что в отличие от уровня детальности (4) коэффициент детальности (9) учитывает самые мелкие структуры изображения, к которым относятся: точечные объекты, тонкие линии, фрагменты текстуры. При этом не учитываются фрагменты границ более крупных объектов. Например, если все четыре условия в (5) истинны, то очевидно, что данный пиксель будет точечным объектом. Если выполняются три условия, то это будет фрагмент тонкой линии.

Пассивные пиксели относятся к мало изменяющимся по яркости и по цветности областям изображения, а также к фрагментам их границ.

Очевидно, что четкость изображения зависит от значения коэффициента детальности (9). Покажем это на примере тестового изображения. На рис. 1 показан фрагмент (116×139) изображения тестовой таблицы по яркости, где слева показано оригинальное изображение, справа – после Гауссовой фильтрации. Коэффициент детальности FDL_a искаженного изображения вычислен по формуле (9).



а) $FDL_a = 18,97$

б) $FDL_a = 0,11$

Рис. 1. Фрагмент изображения тестовой таблицы после Гауссовой фильтрации

Из рис. 1 видно, что одиночные пиксели и штриховые линии толщиной в один пиксель после Гауссовой фильтрации полностью «размыты». При этом коэффициент детальности уменьшился в 172 раза.

Следовательно, коэффициент детальности FDL_a позволяет оценивать ухудшение четкости оригинального изображения, возникающее в процессе его цифровой обработки.

На рис. 2 показан фрагмент (98×115) фотореалистичного изображения «Женщина в шляпе» после Гауссовой фильтрации, где коэффициент детальности рассчитан для всего изображения (480×720).

Сравнивая два значения детальности, получим, что детальность «размытого» изображения уменьшилась в 42 раза и при этом мелкие структуры практически не различаются глазом.

Таким образом, разработанный алгоритм оценки качества изображения позволяет объективно оценивать ухудшение четкости путем сравнения коэффициента детальности двух изображений: до и после цифровой фильтрации. Также можно сравнивать изображения после выполнения разных алгоритмов фильтрации, сжатия, реставрации и выбирать наиболее эффективные алгоритмы. Однако для сложных нелинейных алгоритмов обработки сигналов изображения необходимо учитывать следующие особенности.



а) $FDL_a = 19,26$ б) $FDL_a^- = 0,46$

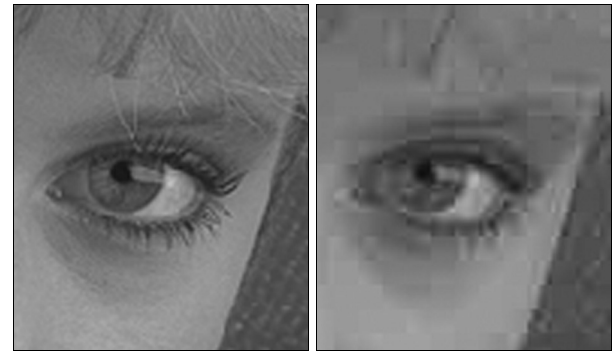
Рис. 2. Фрагмент изображения после Гауссовой фильтрации

В процессе цифровой обработки и восстановления изображения могут возникать ложные компоненты сигналов в виде дополнительных мелких структур, которые отсутствовали в оригинальном изображении. Например, после выполнения сжатия по стандарту JPEG [16] в изображении могут появляться такие артефакты, как блокинг-эффект или выбросы на резких границах, которые будут заметны глазом. Такие ложные элементы будут идентифицироваться в виде активных пикселей и, следовательно, влиять на коэффициент детальности. На рис. 3 показан предыдущий фрагмент изображения после сжатия JPEG в 30 раз в приложении Adobe Photoshop CS6 с низким качеством.

Из сравнения коэффициентов детальности следует, что детальность сжатого изображения уменьшилась всего в 2,4 раза, но при этом визуальное качество значительно ухудшилось.

Для оценки коэффициента детальности с учетом влияния ложных компонент предложен алгоритм, основанный на оценке структурного подобия микроблоков оригинального и искаженного изображений.

На первом этапе выполняем идентификацию активных пикселей оригинального и искаженного изображений по описанному выше алгоритму.



а) $FDL_a = 19,26$ б) $FDL_a^- = 8,04$

Рис. 3. Фрагмент изображения после сжатия JPEG

Для оценки структурного подобия в процессе вычислений последовательно проверяем следующие четыре логические выражения:

$$\begin{aligned} & \text{if}((\Delta K_{i,j}^{(n)} > 1) \wedge (\Delta \tilde{K}_{i,j}^{(n)} > 1)) P_1 = 1 \\ & \text{else } P_1 = 0 \end{aligned} \quad (10)$$

где $n=1, 2, 3, 4$ – номера условий в соответствии с (5); $\Delta \tilde{K}_{j,i}^{(n)}$ – контраст центрального пикселя относительно соседних в искаженном изображении; P_1 – параметр проверки.

Если после проверки всех условий (10) значение параметра равно $P_1 = 1$, принимается решение о том, что структура микроблока искаженного изображения подобна структуре микроблока оригинального изображения. При этом в искаженном изображении маркируются пиксели в окне (3×3) с координатами центрального пикселя (i, j).

После анализа изображения выполняется подсчет маркированных пикселей и вычисляется коэффициент детальности:

$$FDL_a^\Delta = \frac{N_a^\Delta}{W \cdot H} 100\%, \quad (11)$$

где значение N_a^Δ определяет количество маркированных пикселей искаженного изображения.

Относительное количество ложных микроструктур получим с помощью простого выражения:

$$FDL_{false} = FDL_a^- - FDL_a^\Delta. \quad (12)$$

Разработанный алгоритм позволяет устранить влияние ложных компонент на результаты оценок детальности искаженного изображения, и, следовательно, такие результаты будут более объективными.

Снижение четкости удобнее оценивать относительным коэффициентом детальности:

$$R_d = FDL_a^\Delta / FDL_a, \quad (13)$$

т.е. отношением коэффициента детальности (11) искаженного изображения к коэффициенту детальности (9) оригинального изображения. Очевидно, что чем ближе значение R_d к единице, тем ближе качество восстановленного изображения к оригинальному.

Для примера, показанного на рис. 3, получим, что искаженное изображение после сжатия JPEG имеет коэффициент (11) детальности $FDL_a^{\Delta} = 3,82$ и относительный коэффициент (13) $R_d = 0,2$, т.е. по сравнению с оригинальным изображением количество активных пикселей с подобными микроструктурами уменьшилось в 5 раз. При этом процент ложных микроструктур (12) равен $FDL_{false} = 4,2$.

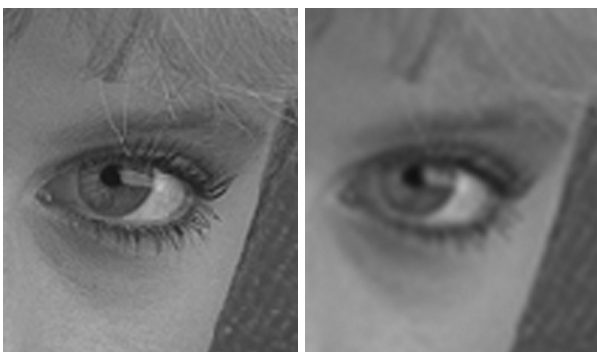
Для сравнения JPEG – изображение «Женщина в шляпе» с установленным максимальным параметром качества ($Q = 12$), сжатое всего в 3,5 раза, имеет следующие параметры: $R_d = 0,97$ и $FDL_{false} = 0,75$.

2. Особенности увеличения разрешения изображения с использованием нейронных сетей

Одним из простых способов увеличения разрешения изображения является его масштабирование с помощью бикубической интерполяции. Однако такой способ масштабирования приводит к существенной потере четкости изображения.

Покажем это на примере. Выберем тестовое изображение «Женщина в шляпе» с разрешением 480×720 пикселей и уменьшим его размер пропорционально в два раза. В итоге получим уменьшенную копию с разрешением 240×360 пикселей и с объемом файла в четыре раза меньше оригинального. Такое изображение становится более компактным, так как требует в четыре раза меньшую пропускную способность канала передачи данных или меньший объем памяти для хранения. При получении изображения или извлечения из памяти восстанавливаем исходный размер с помощью бикубической интерполяции.

Оценим качество восстановленного изображения. С этой целью вычислим интегральные параметры оценки качества PSNR и SSIM, а также параметр оценки четкости R_d . В итоге получим: PSNR = 32,4 дБ, SSIM = 0,914 и $R_d = 0,05$. На рис. 4 показан фрагмент восстановленного изображения после интерполяции.



а) $FDL_a = 19,26$ б) $R_d = 0,05$
Рис. 4. Фрагмент изображения после бикубической интерполяции

Параметры PSNR и SSIM дают общую интегральную оценку качества изображения. При этом считается [17], что хорошее качество обеспечивается при значениях PSNR > 40 дБ и SSIM > 0,98.

Параметр R_d оценивает ухудшение четкости относительным значением различных мелких структур искаженного изображения к количеству мелких структур оригинального изображения.

Полученные результаты показывают, что простое увеличение разрешения уменьшенной копии изображения с помощью бикубической интерполяции приводит к значительному снижению его четкости.

Такой результат объясняется тем, что уменьшение масштаба изображения с помощью передискретизации является необратимым процессом, так как при этом теряются высокочастотные компоненты исходного изображения.

В настоящее время популярными методами, направленными на качественное увеличение разрешения изображения, являются методы суперразрешения (SR), основанные на нейронных сетях. Обзор современных методов SR представлен в [2]. Для решения задач SR применяются различные методы глубокого обучения на основе сверточных нейронных сетей (CNN) (например, SRCNN [18]) или с использованием генеративных состязательных сетей (GAN) (например, SRGAN [19]).

Рассмотрим основные принципы повышения качества увеличенных изображений на примере нейронной сети SRCNN [18].

На входе сети имеем изображение с низким разрешением (LR). На этапе предварительной обработки выполняем масштабирование изображения до заданного размера с помощью бикубической интерполяции. В итоге получим изображение Y .

Задача нейронной сети заключается в восстановлении из Y изображения $F(Y)$, которое будет максимально похоже на исходное изображение X с высоким разрешением (HR). Функциональное преобразование F для рассматриваемой сети SRCNN состоит из трех операций, которые образуют слои сверточной нейронной сети.

1. Выделение и описание фрагментов (патчей) с размером $f_1 \times f_1$. Операция выделяет фрагменты изображения Y и описывает каждый фрагмент в виде многомерного вектора. Вектор представляет собой набор из n_1 характеристик, полученных с помощью набора фильтров. Первую операцию F_1 опишем с помощью выражения:

$$F_1(Y) = \max(0, W_1 * Y + B_1), \quad (14)$$

где W_1 – двумерный цифровой фильтр с размерностью $f_1 \times f_1 \times n_1$ на каждый цветовой канал изображения; B_1 – смещение; n_1 – количество фильтров; * – операция свертки. Отметим, что выражение (14) соответствует математической модели нейрона, где в качестве функции активации ReLU использовано [18] максимальное значение отклика фильтра: $\max(0, x)$.

2. Нелинейное отображение. Операция нелинейно отображает каждый многомерный вектор n_1 на другой многомерный вектор n_2 . Каждый сопоставленный

вектор описывает фрагмент с высоким разрешением. Эти вектора составляют еще один набор карт характеристик. Во втором слое вычисляется функция:

$$F_2(Y) = \max(0, W_2 * F_1(Y) + B_2), \quad (15)$$

где W_2 – цифровой нелинейный фильтр с размерностью $f_2 \times f_2 \times n_1$; B_2 – вектор смещения; n_1 – количество фильтров. Каждый из выходных n_1 -мерных векторов описывает фрагмент с высоким разрешением, который будет использоваться для реконструкции.

3. Реконструкция. Операция объединяет фрагменты с высоким разрешением для создания окончательного изображения с высоким разрешением. В традиционных методах прогнозируемые перекрывающиеся фрагменты с высоким разрешением усредняются для получения окончательного полного изображения. Третий сверточный слой представим в следующем виде:

$$F_3(Y) = W_3 * F_2(Y) + B_3, \quad (16)$$

где W_3 – линейный усредняющий фильтр с размерностью $f_3 \times f_3 \times n_2$; B_3 – вектор смещения.

В итоге выполнения трех операций реставрированное изображение будет похоже на оригинальное изображение X . На рис. 5 показана структура сети SRCNN, реализующая функциональное преобразование F .

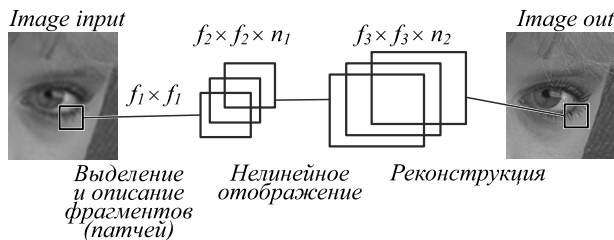


Рис. 5. Структура SRCNN

Для обучения функции F потребуется оценка параметров $\theta = \{W_1, W_2, W_3, B_1, B_2, B_3\}$ для каждого фрагмента из оригинальных тестовых фотореалистичных изображений. Это достигается за счет минимизации функции потерь $L(\theta)$ между восстановленным изображением $F(Y; \theta)$ и фотореалистичным изображением X с высоким разрешением. В процессе обучения используется набор из фрагментов $\{X_i\}$ оригинальных изображений. Для вычисления $L(\theta)$ обычно используется среднеквадратичная ошибка (MSE):

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \theta) - X_i\|^2, \quad (17)$$

где n – количество обучающих выборок. Потери минимизируются с помощью стохастического градиентного спуска со стандартным обратным распространением. Минимальное значение $L(\theta)$ позволяет оптимизировать параметры сети SRCNN и получить карты характеристик фрагментов изображений.

Поскольку SRCNN решает задачу преобразования изображения LR в изображение HR, в которой вход-

ное изображение сильно коррелирует с выходным, часто используется глобальное остаточное обучение. В этом случае не используется сложное преобразование из полного изображения в другое, а требуется только обучение по остаточной карте для восстановления недостающих высокочастотных деталей. При этом сложность модели и вычислительные затраты на обучение значительно снижаются.

Таким образом, для восстановления качества увеличенного изображения с использованием нейронных сетей используется принцип дорисовки [20] элементов изображения между пикселями в узлах решетки интерполированного входного изображения. При этом с помощью сети выбираются те тестовые фрагменты, которые будут наиболее подобными фрагментам входного изображения. Этот основной принцип заложен во всех используемых алгоритмах SR.

В целом, семейства алгоритмов SR отличаются друг от друга по следующим характеристикам [2]: разные типы сетевых архитектур, разные типы функций потерь, различные типы принципов и стратегий обучения и т. д.

3. Результаты оценок качества реставрированных HR-изображений

Из популярных программных приложений [5], реализующих методы SR, выберем следующие: Deep Image, Let's Enhance, Topaz Gigapixel и Waifu2x. Далее используем эти приложения для проверки качества реконструированных изображений на основе традиционных метрик PSNR и SSIM, а также с помощью разработанного способа.

Для сравнительного анализа использовался набор фотореалистичных изображений из базы LIVE [21]. На первом этапе экспериментов размер тестового изображения уменьшался пропорционально в два раза с помощью фоторедактора Adobe Photoshop CS6. На втором этапе разрешение изображения увеличивалось также в два раза на основе метода SR с использованием выбранного приложения.

Для проведения экспериментов реализован программный модуль в среде C++, выполняющий следующие функции.

1. Открытие оригинального и искаженного изображений и преобразование цифровых сигналов из RGB в $L^*a^*b^*$ формат.
2. Идентификация мелких структур оригинального изображения и вычисление коэффициента детальности FDL_a (8).
3. Идентификация мелких структур искаженного изображения, вычисление уровня детальности FDL_a^i (10) и относительного коэффициента R_d (12).
4. Вычисление параметров PSNR и SSIM.

На рис. 6 показаны фрагменты тестовых изображений разной детальностью: «Женщина в шляпе» и «Попугай» – после реставрации. В табл. 1 приведены расчетные параметры.

Анализ полученных результатов показывает, что по традиционным метрикам PSNR и SSIM наиболее высокое качество реставрированных изображений обеспечивается при использовании приложения Waifu2x, а наихудшее качество будет при использовании приложения Let'sEnhance.

Однако по параметру R_d лучшую четкость обеспечивает приложение Deep Image, а худшие результаты дает приложение Waifu2x. По количеству ложных компонент FDL_{false} лучшие результаты обеспечивает Waifu2x, а худшие результаты Let'sEnhance, что соответствует метрикам PSNR и SSIM.

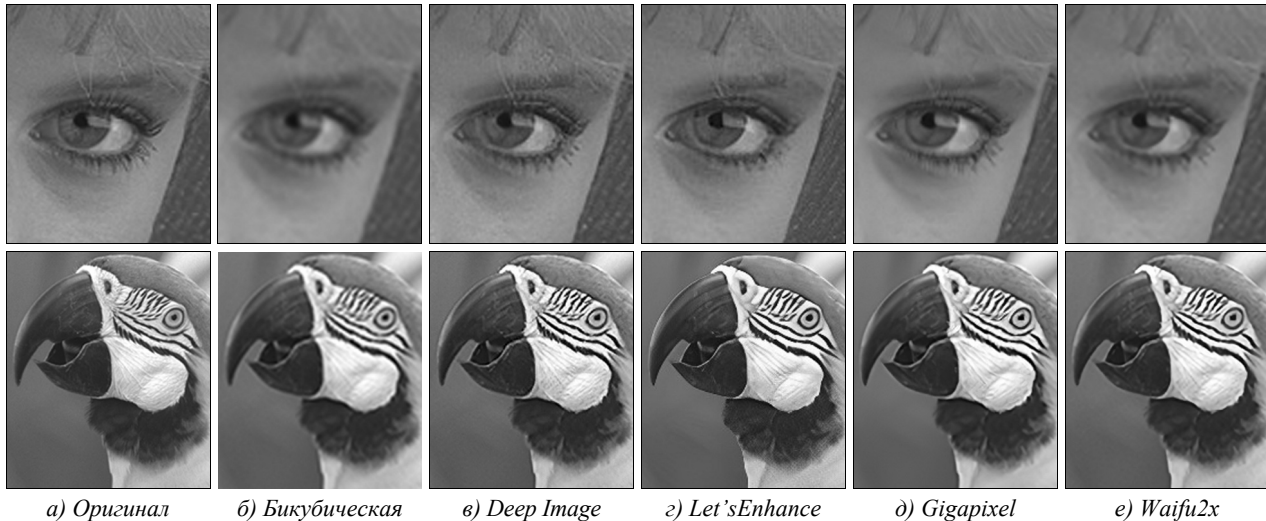


Рис. 6. Фрагменты реставрированных тестовых изображений

Табл. 1. Оценка искажений

	Бикубическая	Deep Image	Let'sEnhance	Gigapixel	Waifu2x	
PSNR	32,41	32,56	30,45	31,90	34,47	Женщина в шляпе ($FDL_a = 19,26$)
SSIM	0,914	0,927	0,892	0,918	0,943	
FDL_a	1,23	19,45	20,07	11,47	6,19	
R_d	0,05	0,51	0,40	0,33	0,25	
FDL_{false}	0,33	9,61	12,33	5,11	1,47	
PSNR	34,33	33,57	29,29	31,54	34,51	Попугай ($FDL_a = 6,74$)
SSIM	0,964	0,958	0,901	0,941	0,969	
FDL_a	2,81	7,92	16,87	7,98	5,30	
R_d	0,30	0,69	0,61	0,64	0,57	
FDL_{false}	0,80	3,26	12,75	3,66	1,46	

Визуальное сравнение качества реставрированных изображений также дает лучшие субъективные оценки для приложения Deep Image.

Таким образом, можно сделать вывод о том, что разработанный способ оценки четкости является более объективным по сравнению с традиционными метриками и позволяет обоснованно выполнять сравнительный анализ качества реставрированных изображений на основе SR-методов.

Заключение

В результате исследований в данной работе получены следующие выводы. Развитие систем искусственного интеллекта с целью улучшения четкости изображения на основе методов SR несомненно является перспективным направлением. Однако традиционные метрики качества реставрированных изображений не позволяют однозначно сопоставить результаты численных оценок с визуальными оценками.

В настоящей работе предложен и описан подход к решению этой проблемы, основанный на оригинальном способе измерения четкости изображения. Оригинальность заключается в том, что по сравнению со стандартным способом измерения разрешения на основе SFR в предлагаемом способе используется алгоритм идентификации активных пикселей эталонного и искаженного изображений с учетом свойств зрительного восприятия контраста мелких деталей и наличия ложных микроструктур. При этом в процессе анализа не требуются тестовые таблицы и используются характеристики фотореалистичных изображений.

Получены численные критерии – относительный коэффициент детальности R_d и коэффициент ложных компонент FDL_{false} , по значениям которых можно делать выводы о качестве реставрированного изображения. Чем ближе значение R_d к единице и значение FDL_{false} к нулю, тем ближе качество восстановленного изображения к оригинальному.

В частности, для фотореалистичных изображений, сжатых по стандарту JPEG или JPEG2000 с высокими параметрами качества, эти критерии в среднем соответствуют значениям $R_d > 0,95$ и $FDL_{false} < 2\%$.

Результаты оценок качества реставрированных HR-изображений (табл. 1) показывают, что существующие методы SR ещё далеки от совершенства. Однако по сравнению с простой бикубической интерполяцией имеют очевидное преимущество. Сохранение более 50 % четкости мелких структур увеличенного изображения с помощью приложения Deep Image – это очень хороший результат.

References

- [1] Sai SV, Kamensky AV, Kuryachy MI. Modern methods of analyzing and improving the digital images quality. Khabarovsk: Publishing house of the Pacific State University; 2020.
- [2] Wang Z, Chen J, Hoi Steven CH. Deep learning for image super-resolution: A survey. IEEE Trans Pattern Anal Mach Intell 2021; 43: 3365-3387.
- [3] Lin W, Jay Kuo C-C. Perceptual visual quality metrics: A survey. J Vis Commun Image Represent 2011; 22(4): 297-312.
- [4] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: From error visibility to structural similarity. IEEE Trans Image Process 2004; 13(4): 600-612.
- [5] Top 15 Best Image Enlarger Review 2021. Source: <https://topten.ai/image-enlargers-review/>.
- [6] Barten PGJ. Contrast sensitivity of the human eye and its effects on image quality. Knegsel: HV Press; 1999.
- [7] Dvorkovich VP, Dvorkovich AV. Measurements in video information systems (theory and practice). Moscow: "Technosphere" Publisher; 2015.
- [8] ISO 12233:2017. Photography – Electronic still picture imaging – Resolution and spatial frequency responses. Source: <https://www.iso.org/standard/71696.html>.
- [9] Born M, Wolf E. Principles of optics: Electromagnetic theory of propagation, interference and diffraction of light. 7th ed. Cambridge: Cambridge University Press; 2019.
- [10] Burns PD, Williams D. Sampling efficiency in digital camera performance standards. Proc SPIE 2008; 6808: 680805.
- [11] Williams D, et al. A pilot study of digital camera resolution metrology protocols proposed under ISO 12233, edition 2. Proc SPIE 2008; 6808: 680804.
- [12] Gonzalez RC, Woods RE. Digital image processing. 3rd ed. Prentice Hall; 2008.
- [13] High resolution test patterns. Source: <http://www.bealecorner.org/red/test-patterns/>.
- [14] Imatest. Source: <https://www.imatest.com/>.
- [15] Sai SV. Metric of fine structures distortions of compressed images. Computer Optics 2018; 42(5): 829-837. DOI: 10.18287/2412-6179-2018-42-5-829-837.
- [16] Pennebaker WB, Mitchell JL. JPEG still image data compression standard. New York, USA: Springer; 1992.
- [17] Bovik A, Mittal A. No-reference image quality assessment in the spatial domain. IEEE Trans Image Process 2012; 21(12): 4695-4708.
- [18] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. IEEE Trans Pattern Anal Mach Intell 2016; 38(2): 295-307.
- [19] Ledig C, Theis L, Huszar F, et al. Photo-realistic single image super-resolution using a generative adversarial network. arXiv Preprint 2017. Source: <https://arxiv.org/abs/1609.04802>.
- [20] Yang C, Lu X, Lin Z, et al. High-resolution image-inpainting using multi-scale neural patch. Proc IEEE Conf on Computer Vision and Pattern Recognition 2017: 6721-6729.
- [21] Sheikh HR, Wang Z, Cormack L and Bovik AC. LIVE image quality assessment database. Source: <http://live.ece.utexas.edu/research/quality>.

Сведения об авторе

Сай Сергей Владимирович, 1960 года рождения, в 1983 году окончил Томский институт автоматизированных систем управления и радиоэлектроники (ТИАСУР) по специальности «Радиоэлектронные устройства». Доктор технических наук, доцент, заведующий кафедрой вычислительной техники Тихоокеанского государственного университета (ТОГУ). Область научных интересов: анализ изображений и распознавание образов. E-mail: sai1111@rambler.ru.

ГРНТИ: 28.23.15

Поступила в редакцию 31 марта 2021 г. Окончательный вариант – 5 июля 2021 г.

A method for assessing photorealistic image quality with high resolution

S.V. Sai¹

¹ Pacific National University, Khabarovsk, Russia

Abstract

The article proposes a method for assessing photorealistic image quality based on a comparison of the detail coefficients in the original and distorted images. An algorithm for identifying fine structures of the original image uses operations of active pixels segmentation, which include point objects, thin lines and texture fragments. The number of active pixels is estimated by the value of a fine detail factor (FDF), which is determined by the ratio of active pixels to the total number of image pixels. The same algorithm is used to calculate the FDF of the distorted image and, further, the image quality deterioration is estimated by comparing the obtained values. Special features of the method include the fact that the identification of small structures and the segmentation of active pixels are performed in the normalized system N-CIELAB. The algorithm also takes into account the influence of false microstructures on the results of the restored image estimating. Features of the construction of neural networks SRCNN in the tasks of a qualitative increase in the image resolution with the restoration of fine structures are considered. Results of the analysis of the quality of enlarged images by the traditional metrics PSNR and SSIM, as well as by the proposed method are also presented.

Keywords: image analysis, super resolution, fine structures, distortion metric.

Citation: Sai SV. A method for assessing photorealistic image quality with high resolution. *Computer Optics* 2022; 46(1): 121-129. DOI: 10.18287/2412-6179-CO-899.

Author's information

Sergey Vladimirovich Sai, born in 1960, in 1983 graduated from the Tomsk Institute of Automated Control Systems and Radioelectronics (TUSUR), specializing in Radioelectronics Devices. Doctor of Technical Sciences, Professor, Head of the Computing Technology department, Pacific National University (TOGU). Area of scientific interests: image analysis and pattern recognition. E-mail: sai1111@rambler.ru.

Received March 31, 2021. The final version – July 5, 2021.
