

Методы оценки качества изображений

Л.В. Рябова, ассистент кафедры средств защиты информации института информационно-диагностических систем Национального авиационного университета, lubanau@ukr.net

Знание особенностей человеческого зрения позволяет строить объективные метрики оценки качества изображений, наиболее близко коррелирующие с субъективными экспертными оценками. Объективные техники измерений-это математические модели, которые удачно моделируют результаты субъективной оценки качества, они основаны на критериях и метриках, что могут быть измерены объективно. В основе алгоритмов обработки изображений положены в основном интегральные преобразования: свертка, преобразование Фурье и др. Также используются статистические методы. Методы обработки изображений классифицируют обычно по количеству пикселей участвующих в одном шаге преобразования:

- поточечные методы в процессе выполнения преобразуют значение в точке в значение независимо от соседних точек;
- локальные (окрестностные) методы для вычисления значения используют значения соседних точек в окрестности ;
- глобальные методы определяют значение на основе всех значений исходного изображения.

Объективные методы классифицируются в соответствии с полезностью исходного сигнала, для которого обеспечивается высокое качество. Поэтому они классифицируются по трем категориям: полные эталонные (reference) методы, сокращенные эталонные (reference) методы и не эталонные (no reference) методы. В современной литературе описано достаточно большое количество объективных метрик, которые можно разделить на три класса [2]:

- Эталонные (full-reference, FR) – предполагают наличие исходного изображения, которое рассматривается как опорное или эталонное изображение при сравнении, т.к. оно не зашумлено и имеет идеальное качество.
- Неэталонные (no-reference, NR) – предполагают, что в процессе получения оценки качества изображения опорное или эталонное изображение отсутствует. Такие метрики являются самыми сложными в реализации и зачастую ориентированы на конкретный вид искажения.
- Псевдоэталонные (reduced-reference, RR) – предполагают, что некоторая часть информации об эталонном изображении присутствует вместе с зашумленным изображением, причем количество этой информации значительно меньше объема информации, требуемого для эталонного изображения.

В данной работе рассмотрены две объективных техники измерения PSNR и SSIM [1].

1. Метрика PSNR. Пиковое отношение сигнала к шуму (англ. *peak signal-to-noise ratio*) обозначается аббревиатурой **PSNR** и является инженерным термином, означающим соотношение между максимумом возможного значения сигнала и мощностью шума, искажающего значения сигнала. Поскольку многие сигналы имеют широкий динамический диапазон, PSNR обычно измеряется в логарифмической шкале в децибелах. PSNR наиболее часто используется для измерения уровня искажений при сжатии изображений. Проще всего его определить через среднеквадратичное отклонение (MSE), которое для двух монохромных изображений I и K размера $m \times n$, одно из которых считается зашумленным приближением другого, вычисляется так:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - K(i, j)|^2.$$

PSNR определяется так:

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right),$$

где MAX_I — это максимальное значение, принимаемое пикселем изображения. Когда пиксели имеют разрядность 8 бит, $MAX_I = 255$, т.е. когда значения сигнала представлены линейно с B битами на значение, максимально возможное значение MAX_I будет $2^B - 1$. Для цветных изображений с тремя компонентами **RGB** на пиксель применяется такое же определение PSNR, но MSE считается по всем трем компонентам (и делится на утроенный размер изображения). Данная метрика, по сути, аналогична среднеквадратичному отклонению, однако пользоваться ей несколько удобнее за счет логарифмического масштаба шкалы. Ей присущи те же недостатки, что и среднеквадратичному отклонению. Следует отметить, что «хороший» PSNR не всегда гарантирует хорошее качество изображения, из-за того что зрительная система человека обладает нелинейным поведением.

2. Метрика SSIM. Не так давно было разработано несколько более сложных и точных метрик, одной из которых является - метрика структурного сходства SSIM (англ. *Structural SIMilarity-SSIM*). Они вычисляются по более сложным алгоритмам, но обоснованно считаются более точно учитывающими особенности восприятия человека.

SSIM это альтернатива PSNR, которая может лучше коррелировать с ощущаемым качеством сжатого изображения. Вычисляется она следующим образом. Рассматривается i -ый кадр размером $K \times L$ и значениями яркостных компонент $Y_{k,l}^{(i)}$, $k=1 \div K$ и $l=1 \div L$. Выбирается окно W и веса ω_j для каждой j -ой точки окна, $j=1 \div J$, J —количество пикселей в окне. Веса ω_j нормируются на единицу: $\sum_{j=1}^J \omega_j = 1$. На практике, обычно выбирают квадратное окно с весами, распределенными симметрично относительно некоторого центра по Гауссу. Для каждой точки данного кадра центр окна размещается в этой точке. Вычисляются средневзвешенные значения

яркостной компоненты исходного и закодированного окна ($Y-AVG_{k,l}^{(i)}$ и $Y-AVG_{k,l}^{(i)}$), средневзвешенные дисперсии яркостной компоненты исходного и закодированного окна ($Y-VAR_{k,l}^{(i)}$ и $Y-VAR_{k,l}^{(i)}$), средневзвешенная ковариация между яркостными компонентами исходного и закодированного окон ($Y-COV_{k,l}^{(i)}$) по следующим формулам:

$$Y-AVG_{k,l}^{(i)} = \sum_{j=1}^j \omega_j * Y_{k,l(j)}^{(i)},$$

$$Y-VAR_{k,l}^{(i)2} = \sum_{j=1}^j \omega_j * \left(Y_{k,l(j)}^{(i)} - Y-AVG_{k,l}^{(i)} \right)^2,$$

$$Y-COV_{k,l}^{(i)} = \sum_{j=1}^j \omega_j * \left(Y_{k,l(j)}^{(i)} - Y-AVG_{k,l}^{(i)} \right) * \left(Y_{k,l(j)}^{(i)} - Y-AVG_{k,l}^{(i)} \right),$$

где $Y_{k,l(j)}^{(i)}$ – значение яркостной компоненты j -ой точки окна, центр которого расположен в точке с координатами (k,l) i -ого кадра.

Далее рассчитывается значение метрики $SSIM$ в точке:

$$Y-SSIM_{k,l}^{(i)} = \frac{\left(2 * Y - AVG_{k,l}^{(i)} * Y - AVG_{k,l}^{(i)} + c_1 \right) * \left(2 * Y - COV_{k,l}^{(i)} + c_2 \right)}{\left(Y - AVG_{k,l}^{(i)2} + Y - AVG_{k,l}^{(i)2} + c_1 \right) * \left(Y - VAR_{k,l}^{(i)2} + Y - VAR_{k,l}^{(i)2} + c_2 \right)},$$

где $c_1 = (k_1 * MAX_Y)^2$ и $c_2 = (k_2 * MAX_Y)^2$. MAX_Y – максимально возможное значение Y -компоненты. Например, если для задания одного значения Y -компоненты используется 8 бит, то максимальное значение есть $2^8 - 1 = 255$. По умолчанию коэффициенты $k_1 = 0,01$ и $k_2 = 0,03$.

$Y-SSIM_{k,l}^{(i)}$ усредняется по i -ому кадру в отдельности:

$$Y-SSIM_i = \frac{1}{K * L} \sum_{k=1, l=1}^{K, L} Y-SSIM_{k,l}^{(i)},$$

$Y-SSIM_i$ усредняется по всем N кадрам:

$$Y-SSIM = \frac{1}{N} \sum_{i=1}^N Y-SSIM_i.$$

Математические методы оценки качества изображений позволяют значительно упростить процесс оценки и значительно сократить затраты. В результате анализа изображений объективными методами оценки качества изображений были сформированы основные требования как к системе в целом, так и к отдельным ее компонентам (наполнение базы данных изображений и функциональность модуля обработки и кодирования изображений).

Литература

1. Winkler S. Digital Video Quality. Vision models and metrics. Wiley, 2005. 192 p.
2. Wang Z., Bovik A., Sheikh H., Simoncelli E. Image Quality Assessment: From Error Visibility to Structural Similarity // IEEE Trans. On Image Proc., Vol.13, № 4, 2004.