

УДК 004.932

КАК ОДНИМ ЧИСЛОМ ОБЪЕКТИВНО ОЦЕНИТЬ КАЧЕСТВО ИЗОБРАЖЕНИЯ?

В.В. СТАРОВОЙТОВ

*Объединенный институт проблем информатики НАН Беларуси,
Сурганова, 6, Минск, 220012, Беларусь*

Рассматривается понятие численной оценки качества цифрового изображения. Показано, что обобщенный параметр качества в виде среднего значения множества локальных оценок качества не лучший вариант единой оценки. Представлены результаты исследований описания локальных оценок параметрами распределения Вейбулла.

Ключевые слова: цифровое изображение, локальная оценка, параметры формы, распределение Вейбулла.

Введение

Рассмотрим понятие «оценки качества цифрового изображения». Широко известная статья профессора Бовика [1] всю проблему вместила в следующее название: «Почему оценка качества изображения так сложна?». Рассмотрим рис. 1. и попытаемся дать оценку качества каждому изображению, например, по 10-бальной системе. Это потребует некоторого времени у любого человека, причем оценки будут самые разные.



Рис. 1. Изображения, которые необходимо оценить по 10-бальной системе

А если требуется оценить сходство или отличия двух изображений? На рис. 2. представлены два варианта одного изображения. Корреляция между этими изображениями равна 0,9420. Они отличаются одной строкой.



Рис. 2. Изображения, отличающиеся одной строкой

Популярные оценки

Для сравнения цифровых сигналов используется множество оценок: коэффициент корреляции, среднеквадратичное отклонение (MSE), пиковое отношение сигнал / шум ($PSNR$) и др. Эти оценки часто применяются к цифровым изображениям. Однако они плохо соответствуют субъективным оценкам, что является их главным недостатком (см. рис.3).



Рис. 3. Оригинал (слева) и разные варианты искажений изображения с одинаковым MSE относительно оригинала, равным 225

Для изображений разрабатываются оценочные функции, которые учитывают визуальные особенности (контраст, резкость и другие характеристики). Эти функции условно делятся на три группы:

- сравнение с эталоном (full-reference),
- с использованием частичной информации об эталоне (reduced-reference),
- меры без эталона (no-reference).

Последняя группа – это фактически функции оценки качества. К первой группе относится широко используемая мера SSIM, которая определяет степень сходства с эталоном.

Как правильнее: similarity или dissimilarity?

В научной литературе для схожих понятий используются самые разные термины: метрики качества, меры сходства и различия изображений и т.п. Следует отметить, что большинство подобных функций не являются метриками в математическом смысле.

В 90-е гг. прошлого столетия при сравнении двух изображений чаще применялся термин dissimilarity или различие изображений.

Начиная с 2000-х гг., в основном, используется термин similarity или сходство изображений. Фактически это эквивалентные понятия. Как правило, similarity равно 1, когда анализируемое изображение совпадает с эталоном. При этом их различие или dissimilarity равно единице за вычетом Similarity. Иногда в таких случаях говорят, что качество анализируемого изображения идеально.

В 1997 г. была предложена методика построения функций сравнения изображений типа dissimilarity для оценки отличий в паре изображений [2]. Было показано, что оценка MSE является частным случаем, предложенным в книге. Основным признаком тех лет при сравнении изображений была яркость пикселей. Отличием предложенной методики сравнения был учет не только яркости пикселей с одинаковыми координатами, но в ближайшей окрестности. К сожалению такой, подход до сих пор не используется в мерах сравнения изображений.

Mean Opinion Score – усредненная субъективная оценка

Mean Opinion Score (MOS) [3] или усредненная субъективная оценка качества используется в области телекоммуникационных систем при анализе изображений и видеопоследовательностей. Определяется методикой субъективного оценивания с качественной шкалой: от 3 до 11 оценок (хорошо, плохо и т.п.). Каждый эксперт в заданной шкале ставит бальную оценку качества, после чего оценки, полученные для одного объекта анализа, усредняются. Таким образом, это оценка субъективного типа, выставленная усредненным пользователем.

К недостаткам такого типа оценок относятся следующие:

- при изменении группы экспертов могут измениться выставленные оценки;
- процесс довольно долгий и затратный;
- итоговые оценки субъективны и приближенны.

На рис. 4 показано, что субъективные оценки качества *MOS*, выставленные людьми, не имеют линейной зависимости от количественных оценок типа *PSNR*.

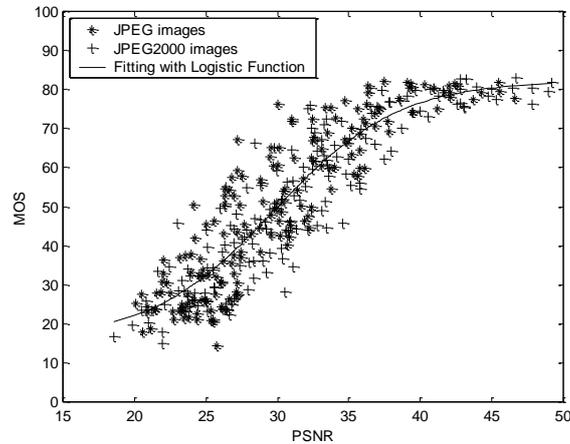


Рис. 4. Сложная зависимость *MOS* от *PSNR*

Исследуемая гипотеза об оценке качества нового типа

В настоящее время автор изучает следующую гипотезу. Оценку качества следует вычислять в 2 этапа – сначала оценить ряд локальных признаков, затем вместо их среднего значения использовать параметр формы их пространственного распределения.

Рассмотрим пример. Средние значения двух множеств локальных признаков, представленных на рис. 5, будут одинаковыми, но форма их расположения различна. Левое множество примерно образует круг, правое – эллипс. Из уравнения эллипса $kx^2 + y^2 - r^2 = 0$ следует, что форма расположения этих данных определяется параметром k . При $k = 1$ это круг, при k , близком к единице – почти круг, при большем или меньшем k – это эллипс.

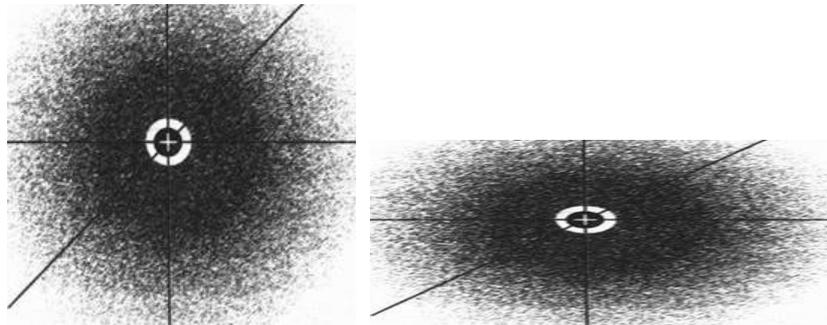


Рис. 5. Два множества локальных характеристик, имеющих одинаковый центр, но разную форму

Аналогичный вывод можно сделать для гистограмм локальных признаков, вычисленных для разных изображений. Например, на рис. 6 показаны три варианта кривых Гаусса, имеющих одинаковое среднее значение, но разные формы из-за различий в дисперсии данных. Проведенными авторами исследования показали, что при оценке сходства двух реально похожих изображений локальные признаки не имеют нормального распределения. Это означает, что глобальная оценка в виде среднего значения не совсем верно отражает локальные отличия.

В качестве альтернативы нормальному распределению локальных признаков рассматривается распределение Вейбулла. Оно имеет достаточно разнообразную форму при изменении параметра формы k (см. рис. 7).

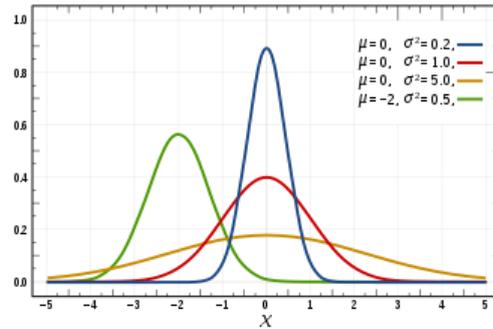


Рис. 6. Нормально распределенные локальные признаки

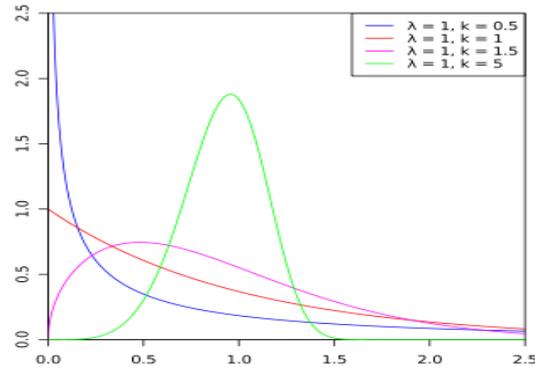


Рис. 7. Варианты распределения Вейбулла

Экспериментальные исследования

Один из независимых вариантов проверки вышеописанной гипотезы, использованный авторами, заключался в исследовании широко известной меры оценки качества двух изображений (т.е. первого типа) *SSIM*. Использовалась ее авторская реализация в системе Matlab. Для пары изображений, представленных на рис. 2, *SSIM* равен 0,9177, коэффициент корреляции – 0,9420, параметр масштаба кривой Вейбулла – 0,9553. Визуальный анализ показывает, что изображения практически одинаковы, т.е. оценка их сходства должна быть максимальна (в нормированном варианте – 1).

На рис. 8 представлены локальные отличия в паре этих изображений, полученные при вычислении меры *SSIM*. Значение меры *SSIM* равно 0,9177 и соответствует максимуму кривой Гаусса.

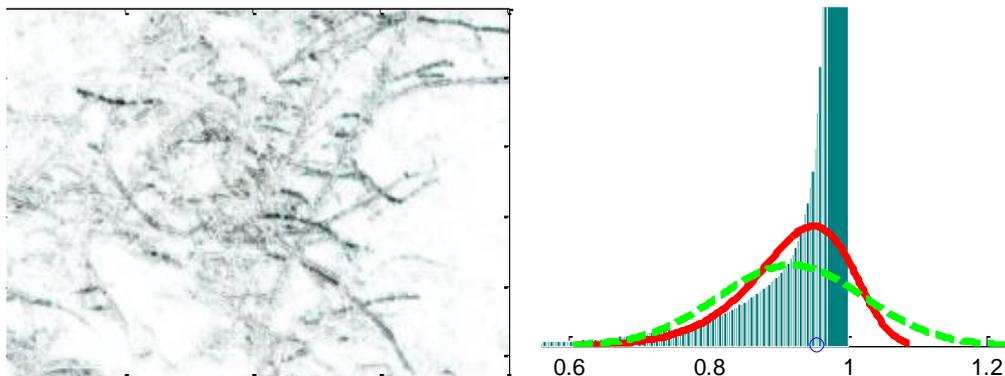


Рис. 8. Слева – карта локальных признаков меры *SSIM*, справа – их гистограмма, аппроксимированная кривыми Вейбулла (сплошная линия) и Гаусса (пунктирная линия)

Заключение

В работе рассмотрено понятие численной оценки качества цифрового изображения. Такие оценки должны удовлетворять противоречивым требованиям: быть точными и соответствовать визуальным оценкам, сделанным человеком. Показано, что среднее арифметическое локальных оценок не лучший вариант глобальной оценки качества. Исследована гипотеза его замены параметром формы распределения локальных данных. На основе первых результатов выполненного исследования можно сделать вывод о возможности получения принципиально новой количественной оценки качества цифровых изображений.

HOW EVALUATE OBJECTIVELY THE IMAGE QUALITY WITH USE OF ONE NUMBER?

V.V. STAROVOITOV

Abstract

The concept of numerical estimation of digital image quality is considered. It is shown that the generalized quality parameter in the form of the mean value of set of local quality estimates is not the best variant of a single assessment. Study results of the description of local estimates by the parameters of the Weibull distribution are described.

Keywords: digital image, local estimates, shape parameters, Weibull distribution.

Список литературы

1. Wang Z., Bovik A.C., Lu L. Why is image quality assessment so difficult? // IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002. Vol. 4. P. 3313–3316.
2. Старовойтов В.В. Локальные геометрические методы цифровой обработки и анализа изображений. Минск: ИТК НАН Беларуси. 1997.
3. Recommendation ITU-R BT. 1788 Methodology for the subjective assessment of video quality in multimedia applications, 2007.