

## ОЦЕНКА КАЧЕСТВА ВОССТАНОВЛЕНИЯ ИЗОБРАЖЕНИЙ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ

**Воронин В.В.**

*ФГБОУ ВПО «Донской государственный технический университет», Ростов-на-Дону, Россия (344010, Ростов-на-Дону, пл. Гагарина, 1), e-mail: voronin\_sl@mail.ru*

В последние годы уделяется много внимания реконструкции изображений, соответственно оценка качества является важной задачей для сравнения различных методов восстановления изображений. Во многих случаях методы реконструкции приводят к размытию текстуры и структуры при восстановлении больших областей с искаженными значениями пикселей. Объективная количественная оценка результатов восстановления в настоящее время отсутствует, в связи с чем во многих подходах используется экспертная оценка. В данной статье рассматривается новый подход оценки качества восстановления изображений на основе машинного обучения с использованием модели зрения человека, который заключается в том, что локальные области изображений могут быть представлены дескрипторами в виде некоторых параметрических распределений. Далее метод опорных векторов регрессии позволяет предсказать воспринимаемое качество восстановленных изображений в соответствии с экспертной оценкой. В работе продемонстрировано, что оценка качества, полученная с использованием приведенного подхода коррелирует с субъективной оценкой качества.

Ключевые слова: обработка изображений, реконструкция, визуальное качество, метрика, машинное обучение.

## QUALITY ASSESSMENT FOR IMAGE INPAINTING BASED ON MACHINE LEARNING

**Voronin V.V.**

*Don State Technical University, Rostov-on-Don, Russia (344010, Rostov-on-Don, sq. Gagarina, 1), e-mail: voronin\_sl@mail.ru*

Inpainting has received a lot of attention in recent years and quality assessment is an important task to evaluate different image reconstruction approaches. In many cases inpainting methods introduce a blur in sharp transitions in image and image contours in the recovery of large areas with missing pixels and often fail to recover curvy boundary edges. Quantitative metrics of inpainting results currently do not exist and researchers use human comparisons to evaluate their methodologies and techniques. This paper focuses on a machine learning approach for no-reference visual quality assessment for image inpainting based on the human visual property. Our method is based on observation that when images are properly normalized or transferred to a transform domain, local descriptors can be modeled by some parametric distributions. Next, we use a support vector regression learned on assessed by human images to predict perceived quality of inpainted images. We demonstrate how our predicted quality value repeatably correlate with qualitative opinion in a human observer study. We show that our approach outperforms known and widely used algorithms on a selected image dataset both in terms of correlation coefficient.

Keywords: image processing, reconstruction, visual quality, metric, machine learning.

Объективная метрика качества изображения является важной частью систем обработки изображений. Одним из важных применений объективных метрик оценки качества изображений является оценка эффективности алгоритмов и систем обработки изображений [1]. Несмотря на большое количество публикаций по данной тематике, задача оценки качества реконструированных изображений рассматривается лишь в немногих. В то же время задача восстановления утраченных областей изображения получила значительное внимание в последнее время [2, 3].

Возможны два подхода к оценке качества изображений: количественная оценка с помощью использования математических методов (среднеквадратическая ошибка,  $L_p$ -норма [4], меры, учитывающие особенности восприятия изображения зрительной системой человека) и субъективная оценка на основе экспертных оценок [5].

Оценка качества, полученная с использованием существующих подходов, может значительно отличаться от оценки, полученной с помощью людей-экспертов. Большинство существующих подходов для оценки качества используют эталонное изображение. Но, к сожалению, во многих случаях эталонное изображение недоступно. К таким задачам относится и задача реконструкции утраченных пикселей. Таким образом, задача разработки количественной метрики оценки качества реконструированных изображений является актуальной.

В разработке количественных оценок качества изображения достигнуты значительные успехи. Однако введенные критерии не являются достаточно совершенными. Большинство попыток найти приемлемые оценки качества изображения относится к частным случаям. Предлагается некая оценка, основанная на каких-либо физиологических предпосылках, а чаще просто удобная для анализа и вычислений, а затем оцениваются его свойства. Создание более совершенных оценок качества изображений связано с более глубоким изучением свойств зрительной системы человека.

Целью данной работы является разработка метрики оценки качества изображений при обработке методами реконструкции на основе машинного обучения.

### Математическая модель

В статье используются обозначения, аналогичные обозначениям, принятым в работе [2]. Все изображение  $I$  состоит из двух непересекающихся областей: реконструированной области  $\Omega$ , и известной области  $\Phi$  ( $\Phi = I - \Omega$ ). На рисунке 1 показан пример расположения этих областей.

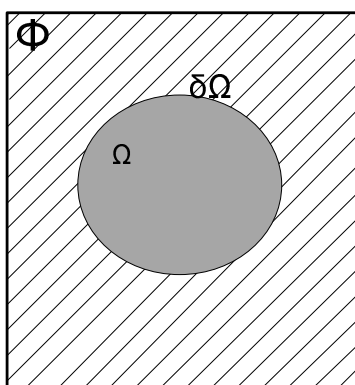


Рис 1. Модель изображения

Известно изображение  $I$  и область  $\Omega$  внутри его. Задача реконструкции состоит в модификации значений пикселей изображения внутри области  $\Omega$ , таким образом, чтобы область не выделялась на фоне окружающего изображения [1]. Цель реконструкции может состоять в

восстановлении поврежденных частей изображения (например, царапин и трещин на старых фотографиях) или удаления нежелательных объектов на изображении. Показанная на рисунке 1 область  $\Omega$  всегда определяется пользователем, т.е. определение области  $\Omega$  не является частью задачи реконструкции.

### Алгоритм оценки качества восстановления изображений

В общем случае для успешного построения метрики качества изображения на основе машинного обучения требуется решение трех следующих задач:

1. Определение пространства признаков  $F$ , служащих описанием входных сигналов.
2. Выбор функции отображения из пространства признаков в пространство оценок качества  $\gamma$ .
3. Обучение системы и проверка ее устойчивости (проверка на переобучение и т.д.).

Структурная схема выбранного подхода представлена на рисунке 2 и содержит следующие этапы:

1. Выбор области интереса (с использованием карты внимания);
2. Вычисление низкоуровневых признаков изображения;
3. Построение дескриптора восстановленной области на базе низкоуровневых особенностей;
4. Решение задачи регрессии с целью получения численной оценки качества на основе полученного вектора-дескриптора.

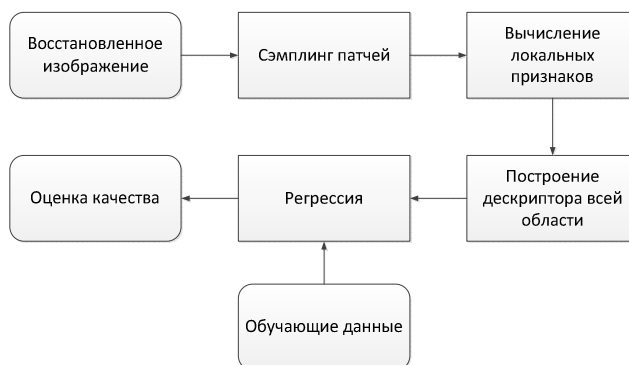


Рис. 2. Блок-схема алгоритма

В работе [5] показано, что визуальное внимание играет важную роль в зрительном восприятии человека. В каждый момент времени человеческий глаз ясно видит лишь небольшую часть сцены, в то же время гораздо более значительная область сцены воспринимается как «размытая». Этой «размытой информации» достаточно для оценки важности различных областей сцены и привлечения внимания к важным областям зрительного поля. Большинство методов позволяют получить карту внимания - двумерное изображение, в котором значения каждого пикселя связано с важностью соответствующей области.

Для получения карт внимания используется Saliency Toolbox, описанный в работе [6]. Данный набор инструментов использует модель зрительной системы человека. Важно отметить, что нет смысла сравнивать восстановленную область на исходном и восстановленном изображении, так как общее содержание может значительно измениться. Для выбора областей интереса предлагается использовать следующее выражение:

$$GD_{out} = \frac{1}{\|\Phi\|} \cdot \sum_{p \in \Phi} S'(p) \cdot$$

Здесь  $S'$  - карта внимания для реконструированного изображения, а  $S'(p)$  значение карты внимания, соответствующее пикселю  $p$ . В выражении, приведенном выше, плотность взгляда вычисляется внутри и снаружи восстановленной области изображений. Значение  $GD_{out}$  используется в качестве порогового значения при принятии решения о том, какие части изображения будут использоваться при оценке, а какие нет. Принимаются во внимания только те области, для которых  $S'(p) > GD_{out}$ .

В качестве низкоуровневых признаков локальных областей используются спектральные представления. Далее предлагается анализ следующих базисов Фурье, Уолша, Хаара с использованием вектора эффективности. Для корректного вычисления составляющих системного критерия эффективности при наличии помех и искажений требует применения статистического усреднения.

При синтезе алгоритмов обработки сигналов и систем чаще всего используется критерий минимума среднего риска, позволяющий учесть статистику помех и сигналов. При реализации частотных преобразований и оценке вычислительных затрат существенное значение имеет выбор базиса спектрального разложения. Для оптимизации выбора базиса разложения сигналов целесообразно использовать критерий минимума среднего риска. Для этого необходимо, чтобы был задан класс используемых сигналов и процессов и были известны их вероятностные характеристики.

Для заданного класса двумерных процессов  $\{Y_\ell(i, j)\}_{\ell=1}^L$ ,  $Y_\ell(i, j) = \{y_\ell^{(k)}(i, j)\}_{k=1}^\infty$  предполагается известной вероятность каждого из подклассов  $P_\ell$ , где индекс  $\ell$  – номер подкласса с некоторыми общими свойствами, а  $k$  – номер реализации процесса  $\ell$ -го подкласса. Будем сравнивать некоторую совокупность базисных систем  $\{\phi_{n,m}^{(h)}(i, j)\}_{n,m=1}^\infty$ ,  $h = 1, 2, \dots, H$ . Разложение  $y_\ell^{(k)}(i, j)$  в обобщенный ряд Фурье по  $n, m$ -й базисной системе в общем виде имеет вид:

$$y_{\ell h}^{(k)}(i, j) = \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} a_{n,m} \phi_{n,m}^h(i, j) \cdot$$

При конечном числе членов ряда Фурье можно охарактеризовать погрешностью:  $\varepsilon_{\ell h}^{(k)} = \rho[y_{\ell}^{(x)}(i, j); \tilde{y}_{\ell}^{(k)}(i, j; K)]$ , где  $\rho[\cdot]$  – расстояние в некоторой метрике,  $\tilde{y}_{\ell}^{(k)}(i, j; K)$  – частичная сумма  $K$  членов ряда Фурье.

Аппаратурное определение коэффициентов ряда Фурье или их вычисление связано с определенными вычислительными затратами  $S_h(\ell, K)$ . Введем функцию потерь, учитывающую как потери, связанные с погрешностью усечения ряда Фурье, так и затраты аппаратурных и вычислительных ресурсов:

$$W_{\ell h}^{(k)} = w \left[ \varepsilon_{\ell h}^{(x)}(\ell, K); S_h^{(k)}(\ell, K) \right].$$

Значение условного риска зависит как от подкласса сигналов, так и от базиса и вычисляется путем усреднения функции потерь по реализациям:

$$r_{\ell h} = \langle W_{\ell h}^{(k)} \rangle_k = \int w \left[ \varepsilon_{\ell h}^{(x)}(\ell, N); S_h^{(k)}(\ell, N) \right] f(x) dx,$$

где  $f(x)$  – плотность вероятности анализируемых сигналов и помех; а угловые скобки означают операцию статистического усреднения.

Средний риск определяется путем усреднения условного риска по подклассам сигналов:

$$R_h = \langle r_{\ell h} \rangle = \sum_{\ell=1}^L P_{\ell} \cdot r_{\ell h},$$

где  $P_{\ell}$  – вероятность  $\ell$ -го подкласса сигналов.

В соответствии с критерием минимума среднего риска из  $H$  базисов выбирается тот, для которого средний риск минимален.

Для оценки эффективности системного критерия качества обработки изображений рассматриваются тестовые изображения в виде текстур, полученных на основе моделирования гауссовских полей с заданными корреляционными функциями. Генерирование однородных нормальных случайных полей, как и стационарных нормальных случайных процессов, наиболее просто производится методом формирующего фильтра [7].

В качестве примера в статье рассматривается представление случайных реализаций с различными корреляционными функциями в базисах тригонометрических функций (Фурье), Уолша и Хаара. Проведем анализ в выбранных базисах для созданных моделей изображений размером 256 на 256 пикселей. Зададимся также тремя типами распределения вероятностей подклассов: 1) равномерное:  $P_1 = P_2 = P_3 = 0,33$ ; 2) убывающее:  $P_1 = 0,5, P_2 = 0,3, P_3 = 0,2$ ;

3) возрастающее:  $P_1 = 0,2, P_2 = 0,3, P_3 = 0,5$ . Выберем функцию стоимости в виде:

$$W_{qm}^{(k)} = \varepsilon_{mq}^2 H_m.$$

Средний риск определяется путем усреднения условного риска по подклассам сигналов с использованием принятых априорных вероятностей подклассов сигналов, рассчитанные значения представлены в таблице 1.

Таблица 1

Величины среднего риска

Базис	Виды распределения вероятностей		
	1	2	3
Фурье	0,71	0,97	0,53
Уолша	0,26	0,3	0,24
Хаара	0,043	0,06	0,03

Результаты расчетов, представленные в таблице показывают, что для принятых моделей двумерных сигналов и распределений их вероятностей базис Хаара имеет наименьший средний риск, а базис Фурье наибольший.

На основе проведенного анализа выберем базис Хаара для представления локальных областей изображений. Следует отметить, что размер восстановленной области  $\Omega$  различен для различных изображений. В связи с этим на основе низкоуровневых признаков следует формировать высокоуровневое представление фиксированного размера. В качестве высокоуровневого представления используется подход «мешок слов». Процедура построения дескриптора (сигнатуры) восстановленной области состоит из двух шагов. На первом шаге строится словарь. Для этого используются низкоуровневые особенности, извлеченные из всех изображений обучающего набора изображений. Для построения словаря извлеченные особенности делятся на 100 классов с помощью алгоритма кластеризации k-средних. Каждый элемент словаря представляет собой центроид для одного из классов, найденных процедурой кластеризации. Каждое слово в словаре представляет преобразование Хаара в блоке изображения размером 8x8. Полученный словарь используется на втором этапе при построении гистограмм частот для слов из словаря в качестве вектора признаков – дескриптора восстановленной области (рис. 3). Полученный набор дескрипторов используется для обучения машины регрессии (Support Vector Regression). Для получения гистограммы частот слов из словаря извлекаются все визуально заметные области (заметность определяется с использованием карт внимания) конкретного изображения. Затем к каждому из извлеченных блоков применяется преобразование Хаара и выполняется классификация согласно полученному словарю на основе евклидова расстояния [8].

Каждый бин результирующей гистограммы содержит число низкоуровневых особенностей конкретного класса в данной восстановленной области. После нормализации гисто-

граммы получается «сигнатура» изображения – высокоуровневое представление восстановленной области [8].

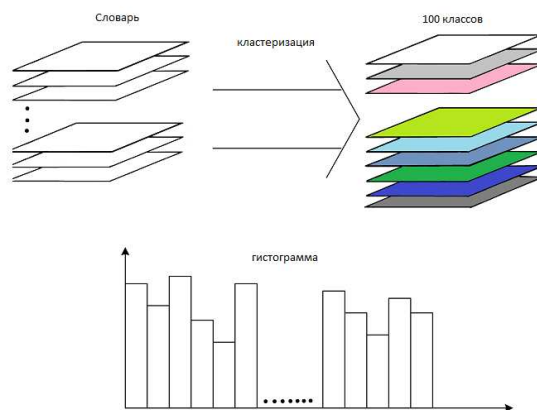


Рис.3. Построение гистограммы

### Оценка эффективности алгоритма оценки качества восстановления изображений

С целью оценки эффективности разработанной метрики использовался набор тестовых изображений. Набор состоит из 300 изображений. В качестве методов восстановления выбраны следующие подходы: метод, основанный на поиске самоподобных областей, метод, основанный на спектральных преобразованиях, метод, основанный на вычислении частных производных. Для каждого изображения была получена экспертная оценка, с привлечением 30 человек. Результаты были разделены на два непересекающихся набора. Первый использовался для обучения, а второй для проверки результата.

Эксперты оценивали качество по шкале, в которой 5 соответствует «Отлично», а 1 – соответствует «Очень плохо». Для оценки эффективности полученной метрики используется коэффициент корреляции между векторами полученных с помощью объективной метрики и экспертным методом оценок качества. Анализ полученных результатов в таблице 2 показывает, что предложенный подход превосходит известные метрики качества на выбранном наборе тестовых данных.

Таблица 2

Коэффициент корреляции для различных методов вычисления объективной метрики качества изображений

DN	SSIM	ASVS	Предложенный подход
0,854	0,768	0,834	0,892

### Заключение

В статье представлена объективная метрика оценки качества изображений на основе машинного обучения. Количественные меры качества изображения необходимы для проектирования и оценки систем воспроизведения изображений. Эти меры во многом помогут избавиться от трудоемкой и неточной современной методики оценки изображений посред-

ством субъективной экспертизы. Кроме того, на основе количественных мер можно развивать методы оптимизации систем обработки изображений. Продемонстрировано, что оценка качества, полученная с использованием приведенного подхода коррелирует с субъективной оценкой качества.

*Работа поддержана Минобрнауки России в рамках ФЦП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2014–2020 годы» (соглашение №14.586.21.0013).*

### Список литературы

1. Gastaldo P. Machine learning solutions for objective visual quality assessment / 6th international workshop on video processing and quality metrics for consumer electronics, VPQM. — Vol. 12. — 2012.
2. Bertalmio M., Bertozzi A., Sapiro G. Navier-Stokes, fluid dynamics, and image and video inpainting/ Hawaii: Proc. IEEE Computer Vision and Pattern Recognition (CVPR) . — 2001.– PP. 213–226.
3. Criminisi A., Perez P., Toyama K. Region filling and object removal by exemplar-based image inpainting / IEEE Trans. Image Process. — 13(9) . — 2004. — PP. 28–34.
4. Vijay M., Cheung, S.S. Eye tracking based perceptual image inpainting quality analysis/ Image Processing (ICIP), 17th IEEE International Conference on IEEE. — 2010. — PP. 1109 - 1112.
5. Ardis P.A., Singhal A. Visual salience metrics for image inpainting /SPIE Electronic Imaging. International Society for Optics and Photonics. — 2009.
6. Cheung S.S., Zhao J., Venkatesh V. Efficient object-based video inpainting / Image Processing, 2006 IEEE International Conference on. — 2006. — PP. 705-708.
7. Перетягин Г.И. Представление изображений гауссовыми случайными полями/ Автоматрия. – № 6. – 1984. – С. 42 – 48.
8. Frantc V.A., Vroni V.V., Marchuk V.I., Sherstobitov A.I., Agaian S., Egiazarian K. Machine learning approach for objective inpainting quality assessment/ Proc. SPIE 9120, Mobile Multimedia/Image Processing, Security, and Applications. – Vol. 91200S. – 2014.
9. Paul A., Singhal A., and. Brown C. Inpainting quality assessment / Journal of Electronic Imaging. – Vol. 19. – 2010. – PP. 011002-011002.

### Рецензенты:

Федосов В.П., д.т.н., профессор, заведующий кафедрой ТОР инженерно-технологической академии Южного Федерального Университета, г.Ростов-на-Дону;



Марчук В.И., д.т.н., профессор, заведующий кафедрой «Радиоэлектронные и электротехнические системы и комплексы» ИСОиП (филиал ДГТУ), г.Шахты.