

ПАРНОЕ СРАВНЕНИЕ ИЗОБРАЖЕНИЙ ДЛЯ ПОЛУЧЕНИЯ ОЦЕНКИ КАЧЕСТВА

Евстеева Виктория Юрьевна^{1,2}

1: студент, факультет ВМК МГУ имени М. В. Ломоносова, Москва, Россия

2: стажер, научно-исследовательский центр компании Schlumberger, Москва,
Россия

E-mail: evsteeva.vika@gmail.com

В ряде задач анализа изображений важно дать безреференсную оценку качества изображений. Разрабатываемые алгоритмы оценки качества необходимо обучать и/или верифицировать. Одним из способов решения проблемы является составление выборки, в которой каждому изображению сопоставлена оценка качества, выставленная комитетом экспертов. В работе [1] анализируются достоинства и недостатки различных подходов выставления оценки качества. Метод парных сравнений обеспечивает более достоверные результаты по сравнению с выставлением оценки для каждого изображения отдельно, так как различия между изображениями имеют большую статистическую значимость. Описанные в [1] подходы предлагают наличие референсных изображений, хотя на практике они часто отсутствуют.

В данной работе рассматривается способ парных сравнений изображений для получения оценки качества при отсутствии референсного изображения. Опрос проводился по подвыборке из датасета LIVE[2], для них уже известна экспертная оценка качества изображений. Предполагается, что дана выборка из n изображений. Процедура опроса состоит в том, что эксперту показывают два изображения из выборки, и ему предлагается оценить какое изображение лучше и насколько (Рис. 1.). Конечный результат - непрерывная оценка изображений от самого худшего до самого лучшего по качеству.



Какое изображение лучше?

☐ Правое

☐ Левое

☐ Одинаковые

На сколько лучше, если
они неодинаковые?

☐ Немного лучше

☐ Лучше

☐ Сильно лучше

Рис. 1. Пример интерфейса для проведения опроса.

При проведении процедуры опроса возникают следующие проблемы: небольшое число экспертов (характерно для узкоспециализированных изображений), субъективность оценки, транзитивность данных опроса, временные затраты для проведения опроса (с n изображениями необходимо оценить $n \cdot (n - 1)/2$ пар изображений).

Для обработки результатов и получения непрерывной шкалы качества используется *матрица парных сравнений* $A_k = (a_{ij}^k)$, где $i, j = 1, 2, \dots, n$, a_{ij}^k - оценка k -го эксперта, показывающая во сколько раз изображение i превосходит по качеству изображение j (данные преобразовываются к виду: "одинаковые" приравнивается к оценке 1, "немного лучше" - 2, "лучше" - 3, "сильно лучше" - 4). Эта матрица обладает следующими свойствами.

$$a_{ii}^k = 1, i = 1, 2, \dots, n \quad (1)$$

$$a_{ij}^k = \frac{1}{a_{ji}^k}, \forall i, j = 1, 2, \dots, n \quad (2)$$

Если матрица парных сравнений $A_k = (a_{ij}^k)$ обладает следующим свойством, то она называется *согласованной*.

$$a_{ij}^k = a_{im}^k \cdot a_{mj}^k, \forall m = 1, 2, \dots, n \quad (3)$$

Для получения непрерывной шкалы качества над матрицей A_k были сделаны следующие преобразования:

$$\tilde{a}_{ij}^k = \frac{a_{ij}^k - \bar{a}_k}{\sigma_k}, i, j = 1, 2, \dots, n \quad (4)$$

где $\bar{a}_k$ - среднее значение всех оценок k -го эксперта, σ_k - среднеквадратичное отклонение оценок k -го эксперта.

Данное преобразование помогает получить центрированные и нормированные оценки изображений для каждого эксперта. Для получения непрерывной шкалы качества берется поэлементное среднее оценок всех изображений всех экспертов, а затем среднее по всем столбцам. Полученный вектор из n значений, нормированный в диапазон от 0 до 1, будет представлять собой непрерывную оценку качества изображений.

Предложенный в работе [3] метод помогает решить проблему транзитивности данных, преобразуя матрицу A_k к согласованной. Время опроса можно сократить, пользуясь свойством согласованно-

сти матрицы парных сравнений. Таким образом, эксперту достаточно провести сравнение $n - 1$ различных пар изображений, которые образуют связный граф. Для частичного решения проблемы небольшого числа экспертов было предложено воспользоваться идеей бутстрепа, когда каждый эксперт оценивает различные пары изображений таким образом, что получается несколько различных связных графов. Пользуясь свойством согласованности матрицы парных сравнений, удастся виртуально увеличить количество экспертов, при этом физически их останется столько же. Достоинство такого подхода в том, что в результате получается более точная оценка качества изображений, но для этого эксперту требуется делать больше сравнений пар изображений.

Для опроса использовалась подвыборка из n изображений из датасета LIVE [2]. Для этих изображений известно экспертное мнение и, соответственно, непрерывная шкала качества - Difference Mean Opinion Score (DMOS). Эта шкала была построена по результатам опроса группы экспертов на основе референсных и искаженных изображений. Полученные нами результаты опроса хорошо коррелируют с DMOS, коэффициент корреляции Пирсона равен 0.91. Из этого можно сделать вывод, что метод дает адекватную непрерывную оценку качества изображений. И стоит отметить, что этот метод в отличие от методов, описанных в [1], работает и в случае отсутствия референсных изображений высокого качества.

Литература

1. Safonov I. V., Mavrin G. N., Kryzhanovsky K. A. Segmentation of convex cells with partially undefined boundaries //Pattern recognition and image analysis. – 2006. – Vol. 16(1). – С. 46-49.
2. Žunić J., Hirota K., Martinez-Ortiz C. Compactness measure for 3d shapes //Informatics, Electronics & Vision (ICIEV), 2012 International Conference on. – IEEE, 2012. – P. 1180-1184.
3. Хемди А. Таха. Введение в исследование операций: пер. с англ./Хемди А. Таха.–7-е изд //М.: Издательский дом «Вильямс», 2005.–912 с.