

**Федеральное государственное автономное образовательное
учреждение высшего образования
«Московский физико-технический институт (национальный
исследовательский университет)»**

ФИЗТЕХ-ШКОЛА ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ

Кафедра когнитивных технологий

Литературный обзор по теме

Исследование инженерных методов оценки качества изображений лиц

Выполнил:
студент М05-012г группы
Попков Андрей

Научный руководитель:
Полевой Дмитрий Валерьевич

Москва, 2020

Исследование по распознаванию лиц мотивировано необходимостью в надежных, эффективных и безопасных методах распознавания для более качественной идентификации.

Однако распознавание лиц по-прежнему является сложной задачей при получении изображений лиц с ухудшенными характеристиками. В последние годы появились недорогие устройства, активным системам распознавания лиц и системам распознавания лиц на базе смартфонов уделяется значительное внимание; такие факты затрудняют обеспечение качества изображения лиц. Качество образцов лиц оказывает значительное влияние на точность биометрического распознавания. Низкое качество образцов является основной причиной ошибок сопоставления в биометрических системах и может быть основным недостатком некоторых приложений. Биометрические подходы к оценке качества изображений используются для измерения качества изображения, и они могут помочь улучшить производительность системы.

В статье «Design and evaluation of photometric image quality measures for effective face recognition»[1] авторы сначала оценивают методы, которые могут измерять такие факторы качества изображения, как контраст, яркость, резкость, фокус и освещенность в контексте распознавания лиц. Во-вторых, они определяют, является ли использование комбинации методов для измерения каждого фактора качества более выгодным с точки зрения качества распознавания лиц, чем использование одного независимого метода. В-третьих, они предлагают новый индекс качества изображения лица (FQI), который объединяет несколько показателей качества и классифицирует изображение лица на основе этого индекса. В исследованиях автора они оценивают преимущества использования FQI в качестве индекса, альтернативного независимым измерениям. Наконец, они проводят Z-тесты статистической значимости, демонстрирующие преимущества предложенного FQI в приложениях для распознавания лиц. В статье, они сначала оценили различные IQM лица, связанные со следующими факторами качества: контраст, резкость, фокус, яркость и освещенность. Они использовали как синтетические, так и реальные данные. Затем они проиллюстрировали, что использование методов обучения с учителем (например, NN) очень важно для понимания взаимосвязи между этими показателями и прогнозированием результатов при использовании в практических сценариях распознавания лиц.

Наиболее часто используются те IQM, которые измеряют следующие показатели качества : яркость, контраст, фокус, резкость и освещенность.

Контрастность изображения - это разница, по интенсивности цвета. Контраст изображения лица можно измерить с помощью следующего уравнения, где μ - среднее значение интенсивности тестового изображения лица $I(x, y)$ размером $N \times M$.

$$C_{RMS} = \sqrt{\frac{\sum_{x=1}^M \sum_{y=1}^N [I(x, y) - \mu]^2}{MN}}$$

Другой метод определения контраста изображения - это мера контраста Майкельсона, где I_{min} и I_{max} - минимальные и максимальные значения интенсивности тестового изображения лица I

$$C_{Mic} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

Яркость - это атрибут визуального ощущения, в соответствии с которым область, в которой представлен зрительный стимул, кажется, излучает больше или меньше света, и диапазон изменения яркости от «яркого» до «тусклого». Измерение яркости изображения лица (обозначим его как B_1) может быть вычисляется как среднее значение компонента яркости после преобразования его в область HSB (оттенок, насыщенность и яркость).

Чтобы преобразовать цвета RGB (красный, зеленый и синий) в диапазон HSB, каждый компонент сначала нормализуется до диапазона $[0, 1]$.

$$B_1 = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [\max(r, g, b)]$$

Bezryadinetal. предложил другое измерение яркости изображения, где X , Y и Z - трехцветные значения. Чтобы преобразовать цвета RGB в XYZ, каждый компонент сначала нормализуется до диапазона $[0-1]$.

$$\begin{bmatrix} D \\ E \\ F \end{bmatrix} = \begin{bmatrix} 0.2053 & 0.7125 & 0.4670 \\ 1.8537 & -1.2797 & -0.4429 \\ -0.3655 & 1.0120 & -0.6104 \end{bmatrix} \times \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$B_2 = \sqrt{D^2 + E^2 + F^2}$$

Фокус изображения лица - это степень размытия изображения лица. Для тонких линз, когда объект (лицо) находится на расстоянии O_d , и изображение

формируется на расстоянии I_d ; фокусное расстояние линзы определяется как:

$$\frac{1}{f} = \frac{1}{O_d} + \frac{1}{I_d}.$$

Если лицо смещено на O_d , свет от лица через линзу камеры распределяется по круглому участку на плоскости изображения, таким образом, образуется размытое изображение лица. Яп и Равендран представили несколько измерений фокуса изображения, таких как L1-норма градиента изображения и энергия Лапласиана. L1-норма изображения определяется как:

$$F_{L_1} = \sum_{x=1}^M \sum_{y=1}^N |G_{xx}(x, y)| + |G_{yy}(x, y)|$$

а энергия Лапласиана изображения как:

$$F_{EL} = \sum_{x=1}^M \sum_{y=1}^N [G_{xx}(x, y) + G_{yy}(x, y)]^2$$

где G_{xx} и G_{yy} - вторые производные в горизонтальном и вертикальном направлениях, соответственно.

Резкость изображения описывает четкость деталей изображения лица и показывает как четки грубые, так и мелкие детали изображения. В литературе было предложено несколько способов измерения резкости изображения. Крыщук и Дригайло определили меру резкости изображения как:

$$S_1 = \frac{1}{2} \left[\frac{1}{(N-1)M} \sum_{x=1}^M \sum_{y=1}^{N-1} |I_{x,y} - I_{x,y+1}| + \frac{1}{(M-1)M} \sum_{x=1}^{M-1} \sum_{y=1}^N |I_{x,y} - I_{x+1,y}| \right]$$

Гао определил меру резкости изображения, как:

$$S_2 = \sum_{x=1}^{M-2} \sum_{y=1}^{N-2} G(x, y) \quad \text{где } G(x, y) - \text{значение градиента в точке } (x, y).$$

Мера резкости по Тененграду определяется как:

$$S_3 = \sum_{x=1}^M \sum_{y=1}^N (L_x \cdot I_x^2 + L_y \cdot I_y^2)$$

$$L_x(x, y) = [I(x+1, y) - I(x-1, y)]^p$$

$$L_y(x, y) = [I(x, y+1) - I(x, y-1)]^p$$

где L_x, L_y - веса в горизонтальном и вертикальном направлениях, а I_x, I_y - горизонтальный и вертикальный градиенты, полученные путем применения Sobelfilter.

Адаптивная мера резкости по Тененграду определяется как:

$$S_4 = \sum_{x=1}^M \sum_{y=1}^N L(x, y) [I_x^2 + I_y^2]$$

$L(x, y) = [I(x-1, y) + I(x+1, y) - I(x, y-1) - I(x, y+1)]^P$ где $L(x, y)$ - вес, а P - индекс мощности, который может определять степень подавления шума.

Освещенность - это один из атрибутов изображения, связанного с освещением. Термин «освещенность» используется для описания количества света, который проходит через определенную область изображения или излучается из нее. UQI - это комбинация трех основных факторов: потеря корреляции, искажение освещенности и искажение контрастности. Искажение освещенности определяется как:

$$I_1 = \frac{2 \sigma_{rt} \bar{r} \bar{t}}{[r^2 + t^2]}$$

где r и t - дисперсии эталонного (r) и тестового изображения (t) соответственно, а σ_{rt} - ковариация (r) и (t).

Другой показатель освещенности изображения рассчитывается как взвешенная сумма средней интенсивности освещенности изображения, разделенного на блоки (4×4)

$$I_2 = \sum_{i=1}^4 \sum_{j=1}^4 \omega_{ij} \cdot \bar{I}_{ij}$$

$$\bar{I}_{ij} = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N I(x, y)$$

где ω_{ij} - весовой коэффициент каждого блока. Абдель-Мотгалебанд Махур определил гауссову маску для добавления веса различным блокам лица. Это приводит к присвоению больших весов блокам в середине изображения и малых весов границам изображения.

Каждый из вышеупомянутых IQM может дать оценку только одного фактора качества изображения. Однако для некоторых биометрических приложений потребуется использовать общий индекс качества для обозначения общего качества входных данных.

Грозер и Табасси объединили метрики качества с использованием методов :

- $q = \min(q_1, q_2)$ - minimum
- $q = \max(q_1, q_2)$ - maximum
- $q = \frac{q_1 + q_2}{2}$ - mean

- $q = \sqrt{q_1 \cdot q_2}$ - geometric-mean

Сравнение нескольких показателей качества схем объединения метрик , чтобы классифицировать входное изображение как "хорошее" или "плохое"

Fusion Rule	Linear, % Normalisation, %	Gaussian, % Models, %
minimum	50.00	70.37
maximum	48.61	51.39
mean	48.15	72.69
geometric-mean	50.00	75.46
dempster-shafer	48.61	65.28
logistic-regression	71.30%	
support-vector-regression	76.39%	
neural-network	81.02%	

Среднее геометрическое оказалось лучшим правилом объединения для интеграции вышеупомянутых мер качества . Таким образом, индекс качества лица определяется следующим образом:

$$FQI = \sqrt[5]{Gc(C) * Gb(B) * Gf(F) * Gs(S) * Gi(I)}$$

где FQI - индекс качества лица полученный с помощью геометрического среднего(geometric-mean), факторы качества подверглись нелинейной нормализации моделью Гаусса(Gaussian models)

Вывод:

1. Хотя NN дает лучшую производительность классификации среди других методов, в нашей работе мы рассматриваем только инженерные методы оценки качества.
2. Наиболее хорошо показал себя метод geometric-mean с нелинейной нормализацией факторов качества.
3. Также удовлетворительный результат показали методы minimum и mean, также с нелинейной нормализацией факторов качества в качестве параметров.

В статье «Automatic Face Image Quality Prediction»[3], рассказывается о методах оценки качества изображения, на основе оценок людей. Описываются 2 метода:

1. Human Quality Values (HQV)

Этот метод использует относительные попарные сравнения качества изображения лиц, задав следующий вопрос: «Какое лицо (левое или правое) имеет лучшее качество?» После сбора случайных наборов парных сравнений лиц от 194 людей, они используем Matrix Completion, предложенный Yi et al, чтобы вывести полный набор оценок качества для каждого оценщика во всей базе данных LFW (всего 13 233 изображения лиц). Цель состоит в том, чтобы вывести $F \in \mathbb{R}^{n \times m}$, матрицу оценок лиц, где n - количество оценщиков, а m - количество изображений лиц.

2. Matcher Quality Values (MQV)

Они определили метку качества для запроса j субъекта i как

$z_{ij} = (s_{ij}^G - \mu_{ij}^I) / \sigma_{ij}^I$ где s_{ij}^G - истинная оценка лица, а μ_{ij}^I и σ_{ij}^I - среднее значение и стандартное отклонение оценок «не истинных». Вручную было выбрано изображение наилучшего качества для 1680 объектов в базе данных LFW минимум с двумя изображениями лиц. Лучшее изображение, помещается в галерею (1680 изображений, по одному на каждый объект), а оставшиеся 7484 изображения этих объектов используются в качестве пробного набора. Дополнительные 4069 изображений в базе данных LFW используются для увеличения размера галереи.

В статье использовалось только 7484 изображения из LFW как для методов MQV, так и для методов HQV. Результат исследования показывает, что удаление низкокачественных изображений в ранжированных по HQV уменьшает FNMR для всех трех систем распознавания, указывая на то, что оценки качества человека коррелируют с производительностью распознавания. Однако MQV гораздо эффективнее снижает FNMR. Полезность целевых значений качества с точки зрения снижения FMR на рис. 1b не так очевиден; на самом деле удаление низкокачественных изображений на основе HQV явно увеличивает FMR для COTS - B, хотя величина этого увеличения невелика (удаление 25% зондовых изображений увеличивает FMR на 0,14%).

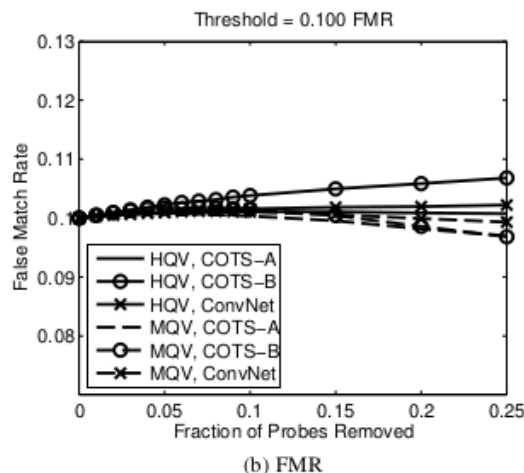
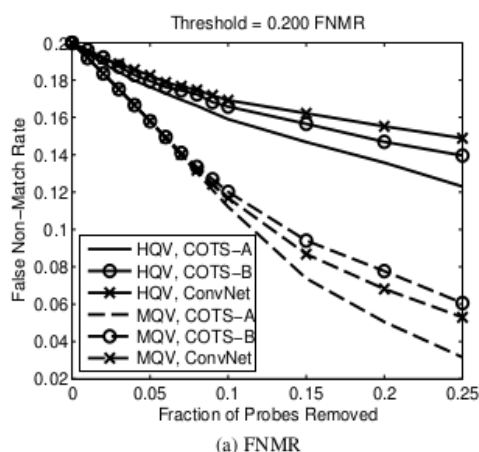


Рис. 1. Кривые зависимости ошибки от количества отбракованных изображений для (a) FNMR и (b) FMR на LFW (размер галереи из 5749 изображений лиц и 7484 пробных изображений лиц из LFW). Изображения отклонялись в порядке ранжирования целевых значений по (HQP) или значений качества на основе баллов (MQV). Пороговые значения фиксированы на (a) 0,20 FNMR и (b) 0,10 FMR для сравнения трех систем распознавания лиц (COTS-A, COTS-B и ConvNet).

Вывод:

1. Для оценки качества изображений человеком, целесообразнее будет выбрать MQV подход, т.к его оценки позволяют определить изображения не подходящие для системы распознавания лиц более точно.

В статье «Face Quality Estimation and Its Correlation to Demographic and Non-Demographic Bias in Face Recognition»[2] авторы представляют углубленный анализ корреляции между смещением при распознавании лиц и оценкой качества лиц. Эксперименты проводились с двумя общедоступными наборами данных, собранными в контролируемых и неконтролируемых обстоятельствах с двумя популярными эмбедингами лица. Авторы оценили четыре современных решения для оценки качества лица на предмет смещения позы, этнической принадлежности и возраста. В статье говорится, что решения по оценке качества лица присваивают значительно более низкие значения качества подгруппам, затронутыми разным смещением, демонстрируя, что эти подходы также являются смещенными. Это поднимает этические вопросы в отношении справедливости и дискриминации, которые должны быть адресованы в будущих работах.

В статье «Performance evaluation of no-reference image quality metrics for face biometric images»[4] авторы оценивают эффективность 13 выбранных IQM без справочных данных по биометрии лица. Результаты экспериментов показывают, что некоторые из них могут оценить качество выборки лиц в соответствии с производительностью системы. Они также анализируют сильные и слабые стороны различных IQM, а также почему некоторые из них не смогли оценить качество выборки лиц.

В статье «Image-Quality-Based Adaptive Face Recognition»[5] представлены адаптивные подходы к распознаванию лиц, позволяющие преодолеть неблагоприятные эффекты различных условий освещения. Качество изображения, которое измеряется с точки зрения искажения яркости по

сравнению с известным эталонным изображением, будет использоваться в качестве основы для адаптации применения процедур нормализации глобального и регионального освещения. Качество изображения также используется для адаптивного выбора параметров объединения для многопоточного распознавания лиц на основе вейвлетов.

Для оценки производительности исследуемых методов будет использован часть данных из набора MIDV-500 [6]. В этих данных 11 видов документов, имеющих фотографию владельца. Каждый документ распечатали, ламинировали для имитации отражения, и сняли с помощью iPhone 5 и смартфона Samsung Galaxy S3.

Каждый из кадров сопровождался файлом JSON с координатами углы документа в следующем формате:

```
{
  "quad":
    [
      [0, 0],
      [111, 0],
      [111, 222],
      [0, 222]
    ]
}
```

В случае, если углы документа выходили за пределы рамки, координаты были экстраполированы так, чтобы отрезки, соединяющие вершины четырехугольника, соответствовали к видимым границам документа. Для каждого из 50 документов в наборе данных был файл JSON, содержащий информацию о расположении полей (на выровненном изображении документа) и значениях его текстовые поля в следующем формате:

```
{
  "field01": {
    "value": "Erika",
    "quad": [ [983, 450], [1328, 450],[1328, 533], [983, 533] ]
  },
  // ...
  "photo": {
    "quad": [ [78, 512], [833, 512],[833, 1448], [78, 1448] ]
  }
}
```

Так как в нашей работе рассматривается качество изображений лиц на фотографиях документов, то нам надо предобработать датасет MIDV-500 и вырезать только область фотографии.

Затем нам предстоит посторить оценки качества на основе оценок людей, для этого используем MQV подход [3].

Из рассмотренных оценок качества изображений и способов их объединения внимания заслуживают geometric-mean, minimum и mean из [1] статьи. На следующем этапе исследования нужно будет вести на нашем наборе данных и построить эти метрики качества.

Затем нам нужно посмотреть как коррелируют метрики качества изображения, построенные с помощью IQM, с метриками качества построенных на основе оценок людей.

Оценка фотографии владельца на документе отличается от оценки изображения лиц на фотографии.

Во первых на документе область изображения лица довольно мала и поэтому качество изображения лица может быть значитель хуже, чем на фотографии.

Во вторых многие документы ламинируют, это может давать блики во время съемки, что также нужно учитывать при оценки качества изображения.

Список литературы

1. Ayman Abaza и др. "Design and evaluation of photometric image quality measures for effective face recognition". в: IET Biometrics 3.4 (2014), с. 314324.
2. Philipp Terhörst и др. "Face Quality Estimation and Its Correlation to Demographic and Non-Demographic Bias in Face Recognition". в: arXiv preprint arXiv:2004.01019 (2020).
3. Automatic Face Image Quality Prediction. Lacey Best-Rowden, Student Member, IEEE, and Anil K. Jain, Life Fellow, IEEE. arXiv:1706.09887v1 [cs.CV] 29 Jun 201
4. Performance evaluation of no-reference image quality metrics for face biometric images. Xinwei Liu, Marius Pedersen, Christophe Charrier, Patrick Bours J. Electron. Imaging 27(2), 023001 (2018), doi: 10.1117/1.JEI.27.2.023001.
5. Image-Quality-Based Adaptive Face Recognition Harin Sellaheewa and Sabah A. Jassim. IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, VOL. 59, NO. 4, APRIL 2010
6. Datasets of ID documents: MIDV-500, <https://smartengines.com/wp-content/uploads/2020/03/Datasets-of-ID-documents-MIDV-500.pdf>