

Модификация алгоритма классификации возраста людей по изображению лица на основе метода ранжирования

Андрей Рыбинцев, Татьяна Лукина, Вадим Конушин, Антон Конушин
 Лаборатория компьютерной графики и мультимедиа
 Факультет вычислительной математики и кибернетики
 МГУ имени М.В. Ломоносова, Москва, Россия
 {arybintsev, tlukina, vadim, ktosh}@graphics.cs.msu.ru

Аннотация

В этой работе рассматривается задача определения возраста человека по изображению лица. Предлагается ряд модификаций для алгоритма на основе множества бинарных классификаций. Экспериментальная проверка модифицированного алгоритма на тестовой базе MORPH показала уменьшение средней ошибки определения возраста с 6.07 до 4.52 и 5 лет при времени классификации 3.21 и 0.32 сек. на одно изображение в зависимости от используемых параметров.

Ключевые слова: Классификация возраста человека по изображению лица, Локальные бинарные шаблоны, Метод опорных векторов.

1. ВВЕДЕНИЕ

Интерес к задаче классификации возраста человека по изображению лица в последние годы неуклонно растет. Одной из главных причин является стремительное увеличение числа изображений в Интернете и необходимость их автоматической аннотации для задач поиска. Человек является самым распространённым объектом, запечатлённым на фотографиях, и автоматическая оценка возраста человека может помочь фильтровать выдаваемые поисковиком результаты.

Другим применением автоматической классификации возраста является анализ возрастного состава аудитории, что полезно для анализа эффективности рекламы.

В будущем классификация возраста также найдёт своё место при создании так называемых «аффективных» человеко-машинных интерфейсов, в которых поведение системы будет подстраиваться под конкретного пользователя, ориентируясь, в числе прочего, и на его возраст.

В этой работе предлагается ряд модификаций для одного из лучших на данный момент алгоритмов оценки возраста человека по фотографии, взятого из статьи Chang et. al «Ordinal Hyperplanes Ranker with Cost Sensitivities for Age Estimation» [4]. Для повышения точности алгоритма мы предлагаем заменить использованный в исходной работе дескриптор на локальные бинарные шаблоны. Поскольку это также приводит к существенному увеличению размера классификаторов и уменьшению скорости работы алгоритма, мы предлагаем ряд модификаций для решения этой проблемы.

2. СУЩЕСТВУЮЩИЕ МЕТОДЫ

Большинство алгоритмов классификации возраста человека по изображению лица состоят из трёх шагов: нормализация изображения лица, вычисление дескриптора,



Рис. 1. Примеры изображений из базы MORPH Album 2.

классификация. Различаются алгоритмы в первую очередь реализацией отдельных указанных шагов.

На шаге нормализации стандартным решением является поворот, масштабирование и обрезание изображения таким образом, чтобы глаза (уголки или центры глаз) находились в заданном положении на одной горизонтальной линии. Другим методом нормализации является использование аффинного преобразования и деформации на основе триангуляции Делоне [16]. Для повышения устойчивости к точности выделения центров глаз и других антропометрических точек в тренировочную выборку можно добавлять немного смещенные изображения с неточной нормализацией [12].

В качестве признаков изображения могут использоваться непосредственно интенсивности пикселей, чаще всего с последующим понижением размерности, например, с помощью метода проекций, сохраняющих свойство локальности (Locality Preserving Projections) [6]; геометрические признаки, например, расстояния между антропометрическими точками [13]; параметры «Активной модели внешности» (Active Appearance Models) [7], локальные бинарные шаблоны [11]. В [9] были предложены так называемые «биологически обусловленные признаки» (Biologically Inspired Features, BIF), использование которых позволило достичь одни из наилучших результатов на задаче классификации пола и возраста.

Для собственно классификации возраста обычно используется либо регрессия [9], позволяющая оценивать точный возраст человека, либо многоклассовая классификация [8, 9], предсказывающая возрастную группу человека. В качестве алгоритмов классификации используют нейронные сети [10], случайный лес деревьев [13]. Наиболее широкое применение нашел в этой задаче метод опорных векторов [3, 4, 7, 8, 9].

В статье [8] предложили вначале классифицировать пол и примерный возрастной диапазон человека, после чего применять специализированный классификатор возраста

для данной половозрастной группы. Авторы статьи показали, что такая двухступенчатая процедура заметно повышает точность итоговой классификации.

В [3] задача классификации возраста человека была сведена к набору бинарных классификаций. Для каждого возраста строится классификатор, который определяет, старше ли человек, изображенный на тестовом изображении, чем заданный возраст, или младше. Возраст человека определяется как число классификаторов, выдавших положительный результат. Предложенная схема показала более высокую точность классификации, чем существующие аналоги.

В [4] данный метод получил развитие. В частности, в [3] использовался метод опорных векторов, и бинарные классификаторы отличались между собой только сдвигом разделяющей поверхности, в то время как в [4] бинарные классификаторы обучались независимо друг от друга, и в каждом классификаторе могло использоваться своё ядро. Также было предложено вместо минимизации ошибки классификации минимизировать цену ошибки. Для этого для каждой из двух используемых метрик качества алгоритма были предложены специальные функции расчёта цены ошибки.

3. БАЗОВЫЙ АЛГОРИТМ

По результатам проведенного обзора существующих методов мы выбрали для последующей модификации алгоритм Chang et. all «Ordinal Hyperplanes Ranker with Cost Sensitivities for Age Estimation» [4] как показавший один из лучших результатов. Поэтому в данной секции опишем его подробнее.

Обозначим i -ый тренировочный пример из обучающей выборки как x_i , а через y_i - соответствующий ему возраст, где $y_i \in \{1, \dots, K\}$, K - максимальный возраст. Конкретное значение K зависит от используемой базы, в [4] взят равным 80.

Для каждого возраста k построим 2 выборки X_k^+ и X_k^- следующим образом:

$$\begin{aligned} X_k^+ &= \{(x_i, +1) \mid y_i > k\} \\ X_k^- &= \{(x_i, -1) \mid y_i \leq k\} \end{aligned} \quad (1)$$

По этим выборкам X_k^+ и X_k^- обучим бинарный классификатор, который будет отвечать на вопрос, больше заданному человеку, чем k лет, или нет.

Т.к. на практике важность, т.е. цена, разных ошибок отличается, то вводится функция стоимости $\text{cost}_k(l)$, где $k, l = 1 \dots K$. Эта функция показывает цену неправильной классификации примера возраста l для k -го классификатора и зависит от используемой меры качества итогового алгоритма.

Наиболее часто применимой мерой качества в задаче определения возраста является MAE (mean absolute error), которая определяется следующим образом:

$$MAE = \sum_{j=1}^N |\bar{y}_j - y_j| / M,$$

где \bar{y}_j - предсказанный возраст, y_j - истинный возраст, M - количество тестовых изображений.

Для этой меры в качестве функции стоимости предлагается использовать модуль разности между возрастaми:

$$\text{cost}_k(l) = |l - k|, \quad (2)$$

т.к. чем больше разница между истинным и предсказанным возрастом, тем больше вклад в итоговую ошибку.

Второй широко используемой мерой качества является суммарный счёт («cumulative score»), который показывает число примеров с ошибкой предсказания ниже заданного уровня:

$$CS(L) = (M_{e < L} / M) * 100\%,$$

где $M_{e < L}$ - количество тестовых примеров с ошибкой предсказания e меньше L лет.

Т.к. для этой меры разница до L лет не вносит вклад в общую ошибку, то можно в функции стоимости сделать цену такой ошибки нулевой, а цену всех остальных ошибок одинаковой (например, единичной):

$$\text{cost}_k(l) = \begin{cases} 0, & \text{if } (l - L) \leq k \leq (l + L) \\ 1, & \text{otherwise} \end{cases} \quad (3)$$

Для обучения классификатора используется метод опорных векторов с RBF ядром. Вес каждого примера из обучающей выборки задаётся с помощью функции стоимости ошибки.

Для определения возраста человека нужно применить все $K - 1$ классификаторов, после чего итоговая оценка считается равной числу положительных откликов классификаторов плюс один:

$$r(x) = 1 + \sum_{k=1}^{K-1} \|f_k(x) > 0\|, \quad (4)$$

где $f_k(x)$ - выход k -го классификатора, $\|\cdot\|$ равна 1, если внутреннее условие верно, и 0 иначе.

В качестве признаков используются параметры Активной модели внешности [5].

Для тестирования алгоритма использовались база FG-NET [1] и часть большей по размеру базы MORPH [15]. Примеры изображений из базы MORPH приведены на рисунке 1.

Используемая часть базы MORPH содержит 5492 изображения людей в возрасте от 16 до 77 лет. Во время тестирования используемая выборка случайным образом 30 раз разбивалась на две части - 80% данных использовались для обучения и 20% для тестирования. Для подбора параметров на обучающей выборке использовался метод кросс-валидации с разбиением на 5 сегментов. Результаты усреднялись по всем 30 разбиениям. Итоговый алгоритм показал ошибку MAE в 6.07 лет.

4. ПРЕДЛОЖЕННЫЕ МОДИФИКАЦИИ

4.1 Локальные бинарные шаблоны

Вместо используемых в [4] параметров Активной модели внешности мы предлагаем использовать в качестве признаков локальные бинарные шаблоны [14], хорошо себя

зарекондовавшие в задачах распознавания человека по изображению лица.

Локальные бинарные шаблоны вычисляются путем применения определенного оператора к каждому пикселю изображения. Этот оператор работает следующим образом. Вначале значение интенсивности в пикселе сравнивается со значениями во всех пикселях из некоторой окрестности, например, размером 3×3 пикселя. Результат сравнения записывается как 0, если значение рассматриваемого пикселя меньше центрального, и как 1 в противном случае. Для рассматриваемой окрестности 3 на 3 получается 8 цифр, из которых составляется двоичный вектор, который интерпретируется как двоичная запись целого числа. Это число и является результатом применения оператора к пикселю. Итоговые признаки получаются после разбиения всего изображения решеткой на прямоугольные области, подсчета гистограмм частот появления чисел в каждой области и конкатенации гистограмм по всем областям в один вектор. Процесс вычисления локальных бинарных шаблонов показан на рисунке 2.

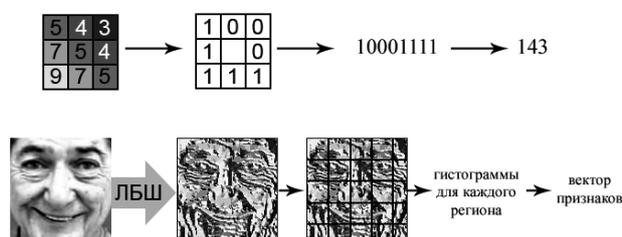


Рис. 2. Вычисление локальных бинарных шаблонов: вверху – для одного пикселя, внизу – для целого изображения.

Эти признаки устойчивы к небольшим изменениям освещенности и небольшим сдвигам в положении лица. Эта устойчивость достигается за счет того, что подсчет ведется не индивидуально для каждого пикселя, а используются области значительного размера.

При замене в базовом алгоритме используемых признаков на локальные бинарные шаблоны, а также использовании линейного метода опорных векторов, ошибка MAE алгоритма на выборке из базы MORPH понизилась с 6.07 до 4.61. При этом использовалась точно такая же процедура тестирования, как и в [4].

В последующих экспериментах мы использовали упрощенную процедуру тестирования. Параметры модели, такие как параметры метода опорных векторов, подбирались с помощью кросс-валидации с 5 разбиениями однократно на всей выборке. После этого вся выборка 30 раз случайным образом делилась на 80% и 20%, на большей части обучались классификаторы, а тестировались на 20%. Данная процедура существенно быстрее исходной, поэтому она лучше подходит для тестирования различных модификаций алгоритма.

Основными недостатками исходного алгоритма являются высокие требования по памяти и длительное время классификации из-за большого количества независимых классификаторов. В работе [4] использовалось 80 классификаторов. При замене параметров Активной модели внешности на локальные бинарные шаблоны данная проблема значительно обостряется, т.к. размер

вектор-признака увеличивается в несколько раз, с нескольких десятков или сотен параметров¹ до 2065.

4.2 Повышение точности классификации

Далее нами были предложены две модификации, дающие дополнительный прирост в точности классификации.

Первая модификация заключается в использовании вероятностного выхода классификатора. В базовом алгоритме подсчитывается число классификаторов, которые выдают результат больше 0, т.е. всех тех, которые считают человека старше, чем сравниваемый возраст. Данная схема была бы правильной, если бы все первые классификаторы выдавали бы положительный результат, а после определенного возраста – отрицательный. Однако на практике в окрестности правильного возраста человека классификаторы могут давать разный результат. Например, среди соседних отрицательных результатов может встречаться и положительный.

Поэтому мы предлагаем суммировать непосредственно вероятностный выход всех классификаторов:

$$r'(x) = 1 + \sum_{k=1}^{K-1} f_k(x)$$

Предварительно к вероятностным выходам классификаторов применяется медианная фильтрация.

Также мы предлагаем изменить функцию стоимости. Исходная функция стоимости (2) линейно зависит от ошибки в оценке возраста, вне зависимости от значения возраста. При этом одна и та же разница в возрасте для детей обычно гораздо заметнее, чем для пожилых людей. Поэтому мы предложили модифицированный вариант этой функции:

$$\text{cost}_k(l) = \alpha \cdot (K - k) \cdot |l - k|,$$

где K - максимальный возможный возраст, α - настраиваемый параметр.

Суммарно, указанные модификации уменьшают ошибку MAE до 4.52 года.

4.3 Ускорение алгоритма

Как было сказано выше, основным недостатком используемого алгоритма является большой размер классификатора и медленная классификация.

Наиболее прямолинейным решением является непосредственное уменьшение числа классификаторов. Например, можно оставить лишь часть классификаторов, различающихся на n лет друг от друга. Это автоматически влечёт за собой уменьшение размера общего классификатора и ускорение времени обработки в n раз. В таблице 1 указана точность алгоритма, использующего все модификации из пункта 4.2, если оставлять лишь каждый n -ый классификатор.

Разумно было бы предположить, что оптимальный набор возрастов, классификаторы которых стоит оставить при заданном количестве классификаторов, отличается от простого выбора каждого n -ого. Для выбора оптимального набора возрастов мы исследовали применимость жадного алгоритма, который по одному «выкидывает» отдельные классификаторы так, чтобы ошибка классификации оставалась минимальной. Однако это приводит к

¹ Точный размер вектор-признака в [4] не указан.

n = 1 (все классификаторы)	n = 2	n = 3	n = 4	n = 5
4,52	4,61	4,65	4,72	4,96

Таблица 1. Средняя ошибка классификации возраста при использовании каждого n-го классификатора

переобучению, т.к. алгоритм естественным образом выкидывает те возрасты, для которых в выборке меньше всего фотографий. Таким образом, жадный алгоритм можно применять только при наличии полностью равномерной выборки.

Также для ускорения классификации возраста можно применять не все классификаторы, а лишь до того момента, когда n последних классификаторов не дадут отрицательный результат. Суммировать результаты нужно также только по примененным классификаторам. При этом нельзя использоваться вероятностные выходы, и приходится округлять их до 0 и 1. В таблице 2 приведены результаты работы алгоритма при различных n.

n=1	n = 2	n = 3	n = 4	n = 5
4,75	4,7	4,65	4,63	4,62

Таблица 2. Средняя ошибка классификации возраста при остановке классификации после n отрицательных результатов

Таким образом, при остановке после 3 отрицательных результатов средняя ошибка увеличилась до 4.65, при этом среднее время классификации уменьшилось более чем 2 раза. Однако уменьшение времени происходит по-разному для различных возрастов. Для старших возрастов уменьшение времени меньше, т.к. приходится применять большее количество классификаторов.

Если же скомбинировать описанные оптимизации вместе, например, используя лишь каждый 3-й классификатор и остановку классификации после 2 отрицательных результатов, можно добиться 8-кратного ускорения и 3-кратного уменьшения размера классификатора при увеличении ошибки определения возраста до 4.72.

Наконец, еще одним способом ускорения является использование бинарного поиска. При этом вначале проверяется средний возраст. Если классификатор даёт положительный результат, то проверяется старшая половина возрастов, если отрицательный - то младшая половина. При этом также можно использовать не все классификаторы, а только каждый n-ый. Проведенные эксперименты показали, что при этом можно в 3 раза уменьшить классификатор и в 10 раз снизить время классификации при увеличении средней ошибки до 5 лет.

4.4 Выводы

Наибольшей точности определения возраста удалось достичь при использовании локальных бинарных шаблонов, суммирования вероятностных выходов и модифицированной функции стоимости.

Компромисс между увеличением точности и временем классификации можно обеспечить за счёт использования каждого третьего классификатора и остановки после 2 отрицательных результатов (обозначим этот алгоритм как «ускоренный вариант 1») или за счёт замены остановки после 2 отрицательных результатов на бинарный поиск («ускоренный вариант 2»).

В таблице 3 показано сравнение указанных вариантов алгоритмов друг с другом и с исходным алгоритмом. Поскольку в исходной работе [4] не указан размер классификаторов и времена работы, эти параметры в таблице не приводятся. Время указано для реализации алгоритма на Матлабе, непосредственно классификация реализована на С++ в виде тех функций. Для классификации использовалась библиотека lib-svm [2].

Алгоритм	Размер классификатора, Гб	Время классификации, с.	MAE, лет
Исходный алгоритм	-	-	6.07
Лучший по точности	1,17	3,21	4.52
Ускоренный вариант 1	0.29	0,40	4.72
Ускоренный вариант 2	0,29	0,32	5

Таблица 3. Сравнение различных модификаций алгоритма на основе бинарных классификаторов.

5. ЗАКЛЮЧЕНИЕ

В данной статье предлагается ряд модификаций алгоритма [4] определения возраста человека по фотографии лица. Благодаря замене параметров Активной модели внешности на локальные бинарные шаблоны в качестве признаков достигнуто существенное повышение точности классификации возраста, однако при этом алгоритм становится медленнее и существенно более требовательным к объёму необходимой памяти. Для обеспечения компромисса между повышением точности, временем работы и требованиями по памяти предложено несколько дополнительных модификаций.

В дальнейшем планируется исследовать возможность повысить точность классификации за счёт использования дополнительных признаков, поиска оптимального разбиения на возрастные группы, а также провести тестирование алгоритма на базе FG-NET [1].

6. БЛАГОДАРНОСТИ

Работа была выполнена при частичной поддержке гранта РФФИ № 11-01-00957-а.

7. ССЫЛКИ

- [1] The FG-NET aging Database, available at <http://www.fgnet.rsunit.com/>.
- [2] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. <http://csiewiki.org/~cjlin/libsvm>, 2001.
- [3] K.-Y. Chang, C.-S. Chen, and Y.-P. Hung. A ranking approach for human age estimation based on face images. In *ICPR*, 2010.
- [4] K.-Y. Chang, C.-S. Chen, and Y.-P. Hung. Ordinal Hyperplanes Ranker with Cost Sensitivities for Age Estimation. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 585--592, 2011.
- [5] T. Cootes, G. Edwards, and C. Taylor. Active appearance models. *IEEE TPAMI*, 2001.

- [6] Y. Fu, Y. Xu, and T. S. Huang. Estimating human ages by manifold analysis of face pictures and regression on aging features. In *Proc. IEEE Conf. Multimedia Expo.*, pp. 1383–1386, 2007.
- [7] G. Guo, Y. Fu, T. S. Huang, and C. Dyer. A probabilistic fusion approach to human age prediction. In *IEEE CVPR-SLAM workshop*, 2008.
- [8] G. Guo, G. Mu, Y. Fu, C. Dyer, and T.S. Huang. A Study on Automatic Age Estimation using a Large Database. *Proc. of ICCV*, pp. 1986-1991, 2009
- [9] G. Guo, G. Mu, Y. Fu, and T. Huang. Human age estimation using bio-inspired features. In *CVPR*, 2009.
- [10] A. Lanitis, C. Draganova, and C. Christodoulou. Comparing different classifiers for automatic age estimation. *IEEE TSMC-Part B*, 2004.
- [11] H.-C. Lian and B.-L. Lu. Multi-view gender classification using local binary patterns and support vector machines. In *Proc. 3rd Internat. Sympos. on Neural Networks (ISNN'06)*, Chengdu, China, vol. 2, pp. 202–209, 2006.
- [12] M. Mayo and E. Zhang. Improving Face Gender Classification By Adding Deliberately Misaligned Faces To The Training Data. *Image and Vision Computing New Zealand*, pp. 1-5, 2008.
- [13] A. Montillo and H. Ling. Age Regression from Faces Using Random Forests. *IEEE International Conference on Image Processing*, pp. 2465-2468, 2009
- [14] T. Ojala, M. Pietikainen and D. Harwood. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognition* 29, pp. 51–59, 1996.
- [15] K. Ricanek Jr. and T. Tesafaye. Morph: A longitudinal image database of normal adult age-progression. In *AFGR*, 2006.
- [16] Z. Yang, M. Li, and H. Ai. An experimental study on automatic face gender classification. *Proc. of the 18th International Conference on Pattern Recognition*, 2006.