

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

**«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»**

Кафедра теоретических основ
компьютерной безопасности и
криптографии

Распознавание лиц с помощью скрытых марковских моделей

АВТОРЕФЕРАТ

дипломной работы

студента 6 курса 631 группы

специальности 10.05.01 Компьютерная безопасность

факультета компьютерных наук и информационных технологий

Андрейченко Александра Дмитриевича

Научный руководитель

к. пед. н., доцент

А.С. Гераськин

23.01.2020 г.

Заведующий кафедрой

д. ф.-м. н., доцент

М. Б. Абросимов

23.01.2020 г.

Саратов 2020

Введение

Эффективные алгоритмы детектирования и распознавания лиц на двумерных изображениях и в видеопотоке весьма актуально с точки зрения обеспечения безопасности режимных объектов.

Происходящие явления можно описывать как сигналы. Сигналы могут быть дискретными, как письменная речь, или непрерывными, как фонограмма или кардиограмма. После некоторого сжатия информации и преобразования к пространству признаков двумерные изображения также могут рассматриваться как сигналы. Сигналы с постоянными статистическими свойствами называются стабильными (стационарными), а с меняющимися — нестабильными (нестационарными). Сигнал может быть чистым, а может и искаженным, с помехами или посторонними сигналами.

Для описания сигналов часто нужны математические модели. В модели сигнала на основе его характеристик может быть предусмотрен определенный механизм обработки, который позволяет получить желаемый выход при анализе сигнала. Например, если надо очистить сигнал, искаженный и зашумлённый при передаче, мы можем смоделировать его и рассмотреть эту модель отвлечённо от искажений и шумов в сигнале. Модели позволяют также генерировать и исследовать сигнал без его источника. В этом случае, имея под рукой хорошую модель, мы можем имитировать сигнал и изучить его по этой имитации.

Модели успешно применяются на практике, позволяя создавать эффективные рабочие системы: системы прогноза, распознавания, идентификации. Грубо все модели можно разделить на детерминистические и статистические. Детерминистические используются, если известны фундаментальные характеристики сигнала. В таком случае достаточно просто описать подобную модель сигнала — для этого нужно всего лишь подобрать (вычислить) параметры этой модели. Второй класс — это вероятностные модели, для разработки которых возможно используются статистические характеристики сигнала. Эти модели описывают гауссовские, пуассоновские, марковские процессы [1], а также подобные им процессы. В общем, вероятностные модели описывают сигнал как

определённый случайный процесс, параметры которого могут быть качественно определены.

При распознавании образов, в частности, при идентификации человека по его фотографии используются специальные вероятностные модели, называемые скрытыми марковскими моделями (СММ) [2, 3]. Основы теории скрытых марковских моделей опубликовал Баум и его коллеги в конце 60-х — начале 70-х годов.

Целью работы является разработка программного модуля, осуществляющего распознавание лиц на основе СММ.

Для достижения поставленной цели требуется решить следующие задачи:

1. Рассмотреть структуру СММ.
2. Изучить основные алгоритмы обучения СММ.
3. Реализовать классификатор на основе СММ.
4. Сравнить качество распознавания с доступным аналогом (OpenCV).

КРАТКОЕ СОДЕРЖАНИЕ

В первом разделе рассмотрена теория скрытых марковских моделей и основные алгоритмы для эффективного решения задачи распознавания лиц на изображениях на основе данного класса математических моделей.

Приведены общие сведения о цепях Маркова и их фундаментальных свойствах.

Скрытая марковская модель [2-4] (СММ, англ. A hidden Markov model, НММ) — модель процесса, в которой процесс считается марковским, причем неизвестно, в каком состоянии s_i находится система (состояния скрыты), но каждое состояние s_i может с некоторой вероятностью $b_{i\omega_j}$ произвести событие ω_j , которое можно наблюдать. Модель представляет из себя марковскую цепь, для которой нам известны начальная вероятность и матрица вероятностей переходов.

Марковская модель λ задается как $\lambda = \{S, \Omega, \Pi, A, B\}$, где $S = \{s_1, s_2, \dots, s_n\}$ — состояния, $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$ — возможные события, $\Pi = \{\pi_1, \pi_2, \dots, \pi_n\}$ — начальные вероятности, $A = [a_{ij}]$ — матрица переходов, а $B = [b_{i\omega_k}]$ — вероятность наблюдения события ω_k после перехода в состояние s_i . Последовательность наблюдаемых событий суть $O = \{\omega_{j_1}, \omega_{j_2}, \dots, \omega_{j_N}\}$.

Приведено описание алгоритма "прямого-обратного хода" (англ. forward-backward algorithm), позволяющего найти в скрытой марковской модели вероятность попадания в состояние s_i на t -ом шаге при последовательности наблюдений O и (скрытой) последовательности состояний X .

Пусть справедливы предположения:

- Скрытые и наблюдаемые события должны быть последовательностью, которая упорядочена по времени
- Каждое скрытое событие должно соответствовать только одному наблюдаемому.

- Вычисление наиболее вероятной скрытой последовательности до момента t зависит только от наблюдаемого события в этот момент времени и наиболее вероятной последовательности до момента $t - 1$ (динамическое программирование)

Тогда алгоритм Витерби (англ. Viterbi algorithm) позволяет сделать наиболее вероятное предположение о последовательности состояний скрытой Марковской модели на основе последовательности наблюдений.

Путь Витерби (англ. Viterbi path) — наиболее правдоподобная (наиболее вероятная) последовательность скрытых состояний.

Алгоритм распределения по Витерби (англ. Viterbi segmentation) предполагает следующее. Пусть у СММ с относительно небольшим числом состояний наблюдается достаточно длительная выходная последовательность (выходной сигнал) большой длины T .

Алгоритм Витерби позволяет сопоставить ей последовательность наиболее вероятных состояний СММ той же самой длины T . Тем самым алгоритм Витерби позволяет «разметить» достаточно длинный выходной сигнал по состояниям СММ и тем самым на его основе обновить параметры скрыт Алгоритм Баума-Велша (англ. Baum–Welch algorithm) — алгоритм для нахождения неизвестных параметров скрытой Марковской модели, которые максимизируют вероятность известной последовательности наблюдений. Использует алгоритм прямого-обратного хода. скрытойой марковской модели.

Алгоритм k средних (англ. k-means algorithm) используется при обучении СММ в процессе распознавания лиц. Является неиерархическим итерационным методом кластеризации. Основная идея алгоритма k-means заключается в том, что данные произвольно разбиваются на кластеры, после чего итеративно пересчитывается центр масс для каждого кластера, полученного на предыдущем шаге, затем векторы разбиваются на кластеры вновь в соответствии с тем, какой из новых центров оказался ближе по выбранной метрике.

Применительно к распознаванию лиц на изображениях, выходные сигналы по своим свойствам близки к непрерывным. Для СММ с непрерывно рас-

предельными случайными наблюдаемыми величинами пространство наблюдений $\Omega = \mathbf{R}^{N_o}$, а последовательность наблюдений $O = \{\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_T\}$ представляет собой последовательность конечномерных вещественных векторов $\mathbf{o}_t \in \mathbf{R}^{N_o}$, $t = 1, 2, \dots, T$. Вероятность эмиссии рассчитывается на основе линейной комбинации Гауссовых нормальных распределений. При известных значениях параметров распределения рассчитываются вероятности эмиссии наблюдаемых значений, и далее могут быть использованы алгоритмы «прямо-обратного хода» и Витерби. Алгоритм Витерби размечает достаточно длинную последовательность наблюдаемых величин по состояниям скрытой марковской модели. Тем самым, можно выделить последовательности наблюдений, соответствующие эмиссии при переходе в фиксированное состояние, и выполнить переоценку параметров распределения вероятности.

Во втором разделе рассмотрены методы и алгоритмы распознавания лиц на изображениях на основе скрытых марковских моделей.

Пикселы двумерных изображений преобразуются к оттенкам серого цвета.

Обычно на фотографиях человека существенные для анализа фрагменты изображения (волосы, лоб, глаза, нос, подбородок) располагаются в естественном порядке сверху вниз, даже при небольших поворотах головы в вертикальной и горизонтальной плоскостях. Соответственно, каждый из этих регионов сопоставляется состояниям «одномерной» СММ.

Вместо всех пикселов фрагмента изображения в множество (вектор), соответствующее отдельному наблюдению, следует включить не все множество пикселов, а лишь некоторые признаки блока. В частности, в [7] предложено применять к прямоугольному блоку двумерное дискретное косинус-преобразование Фурье и использовать в качестве признаков подмножество Фурье-коэффициентов, соответствующих относительно небольшому волновому числам (пространственным аналогам частот в терминологии [7]).

Для обучения СММ дальнейшему распознаванию лиц используется набор

(обучающее множество) фотографий. Изображения в обучающем множестве обычно содержат фотографии «в анфас» в условиях различной освещенности.

Для дальнейшего распознавания, каждому индивидуальному человеку в базе сопоставляется отдельная СММ. Множество различных фотографий одного и того же человека используется для обучения СММ.

После извлечения отдельных блоков из фотографий в обучающем множестве из блоков формируются признаки, составляющие отдельные наблюдения. Далее они используются для обучения СММ.

Сначала выполняется инициализация СММ $\lambda = (A, B, \pi)$. Обучающие данные однородно распределяются между состояниями СММ, и обучающие векторы (множества) в каждом состоянии используются для оценки матрицы вероятностей наблюдения B . Устанавливаются начальные значения для A и π . На следующей итерации равномерное распределение заменяется распределением на основе алгоритма Витерби. Итерации прекращаются, и СММ считается инициализированной, когда вероятность перераспределения на основе алгоритма Витерби становится меньше порогового значения. Далее обучение СММ реализуется на основе рекурсивного применения алгоритма Баума-Вэлша.

Для тестового изображения, которое на насыщенном предметами фоне содержит одно или несколько лиц, детектирование лиц начинается с просмотра каждого прямоугольного окна на тестовом изображении, извлечении для него признаков, составляющих наблюдение, и вычисления вероятности того, что изображение внутри окна соответствует модели лица.

После вычисления признаков для соответствующего изображения, для каждой предварительно обученной СММ на основании алгоритма Витерби вычисляется вероятность соответствия модели лица последовательности наблюдаемых признаков. СММ с наибольшей вероятностью идентифицирует распознаваемое лицо.

В третьем разделе рассмотрены классические алгоритмы детектирования лиц на изображениях и их дальнейшего распознавания, которые использовались для сравнения эффективности с разработанными реализациями.

Рассмотрен классический метод Виолы-Джонса детектирования лиц на изображениях, и каскадный классификатор (детектор лиц на изображениях), реализованный в свободно распространяемой библиотеке компьютерного зрения OpenCV. Приведены алгоритмы вычисления признаков Виолы-Джонса и алгоритма машинного обучения AdaBoost, который одновременно обучает модель и эффективно сокращает количество используемых признаков. Рассмотрен классификатор на основе LBPН-алгоритма, который используется для идентификации лиц на изображениях после их детектирования. Приведены сведения об извлечении признаков в форме гистограмм после разметки изображения однородным LBP-оператором и о метрике для сравнения гистограмм.

В четвертом разделе приведены сведения об использованных средствах разработки, о разработанном программном обеспечении и его эффективности при распознавании лиц. При разработке программного обеспечения в ходе выполнения работы использован язык программирования высокого уровня Java (Java SDK 10.2). Для создания графического интерфейса пользователя использовалась библиотека Swing. Для детектирования границ лица по фотографии использованы стандартный детектор из свободно распространяемой библиотеки OpenCV 4.1.2 для платформы Windows x64. Для сравнения эффективности алгоритмов классификации использована версия LBPН-алгоритма, которая не входит в стандартный дистрибутив OpenCV. В связи с этим исходный код OpenCV был дополнен рядом модулей, свободно распространяемых сообществом пользователей OpenCV. Далее библиотека OpenCV была сконфигурирована и собрана на основе кросс-платформенного конфигулятора СМАКЕ 3.16.0 и компилятора Microsoft Visual C++ 2017.

Были разработаны приложения для распознавания лиц как на основе СММ, так и на основе LBPН-классификатора. Оба приложения поддерживают как консольный режим работы для тестирования и автоматического сбора информации о результативности распознавания, так и режим графического интерфейса пользователя.

Для обучения были использованы фотоизображения 40 людей из следу-

ющих баз изображений: Face Recognition Data университета Эссекса, FaceScrub [18], университета штата Джорджия Georgia Tech face Database. Для обучения классификатора, моделирующего лицо отдельного человека, использовалось 10 черно-белых изображений различных размеров. Для тестирования классификаторов было использовано 80 черно-белых фотографий, содержащих изображения лиц 40 людей (по 2 на человека).

Из представленных в работе данных можно сделать вывод о том, что обучение классификаторов на изображениях большего размера ведет к улучшению качества распознавания. В целом LВРН-классификатор [16, 19] более эффективен, чем классификатор на основе СММ. Это связано с тем, что скрытые марковские модели развиваются в одномерном дискретном времени, и в большей степени оптимизированы для анализа «одномерных» сигналов. Для улучшения качества распознавания с помощью СММ можно использовать другие методы извлечения признаков [20]. В частности, для этого можно использовать алгоритм LВРН.

Показано, что эффективность распознавания лица может значительно снизиться из-за очень большой контрастности освещения, наклона и поворота головы.

Можно предположить, что использование большего количества изображений при обучении модели приведет к лучшим результатам при распознавании, поскольку при извлечении векторов признаков улучшится точность кластеризации признаков.

ЗАКЛЮЧЕНИЕ

Рассмотрена структура скрытых марковских моделей (СММ).

Изучены и реализованы основные алгоритмы обучения и распознавания на основе скрытых марковских моделей.

Реализован классификатор для распознавания лиц на изображениях на основе СММ.

Детектирование лиц на изображениях с заполненным фоном эффективно реализуется на основе каскадного классификатора Хаара из свободно распространяемой библиотеки компьютерного зрения OpenCV.

Для сравнения с разработанным классификатором на основе СММ был сконфигурирован и собран классификатор на основе LBPН-алгоритма, не входящий в стандартный дистрибутив OpenCV.

По результатам тестирования показано, что классификатор на основе LBPН-алгоритма более эффективен. Однако обучение на большем количестве изображений и применение других алгоритмов извлечения признаков (например, LBPН) может значительно улучшить качество работы классификатора на основе СММ.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Гнеденко, Б.В. Курс теории вероятностей/ Б.В. Гнеденко. М.: Наука, 1988. 448 с.
2. Моттль, В.В. Скрытые марковские модели в структурном анализе сигналов/ В.В. Моттль, И.Б. Мучник. М.: Физматлит, 1999. 352 с.
3. Форсайт, Д.А. Компьютерное зрение. Современный подход/ Д.А. Форсайт, Ж. Понс. М.: Вильямс, 2004. 928 с.
4. Vasheghi, S.V. Advanced Digital Signal Processing and Noise Reduction/ S.V. Vashegi. Wiley, 2006. 453 p.
5. Mouhcine, R. Recognition of cursive Arabic handwritten text using embedded training based on HMMs/ R. Mouhcine, A. Mustapha, M. Zouhir// Journal of Electrical Systems and Information Technology. 2018. Vol. 5. Pp. 241-251.
6. Yang, W. Continuous sign language recognition using level building based on fast hidden Markov model/Wenwen Yang , Jinxu Tao, Zhongfu Ye// Pattern Recognition Letters. 2016, vol. 78. Pp. 28-35.
7. Nefian, A.V. Hidden Markov Models for Face Recognition/ A. V. Nefian, M. H. Hayes// Proc. International Conf. on Acoustics, Speech and Signal Processing, Seattle, 1998, Vol. 5, pp. 2721-2724.
8. Nefian, A.V. An Embedded HMMBased Approach for Face Detection and Recognition/A. V. Nefian, M. H. Hayes // ICASSP'99.
9. Малла, С. Вэйвлеты в обработке сигналов/ С. Малла. М: Мир, 2005. 672 с.
10. Фаворская, М.Н. Прогнозирование в системах распознавания образов на основе скрытых марковских моделей/ М.Н. Фаворская, Н.Д. Торгашин, А.Г. Зотин// Вестник Сибирского государственного аэрокосмического университета имени академика М.Ф. Решетнева. 2006. Вып. 1 (8). С. 59-63.
11. Nefian, A.V. An embedded HMM-based approach for face detection and recognition/A.V. Nefian, M.H. Hayes // 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. IEEE. 1999. Vol. 6. Pp. 3553-3556.
12. Othman, H Hybrid hidden markov model for face recognition/ H. Othman, T.

- Aboulnasr// 4th IEEE Southwest Symposium on Image Analysis and Interpretation. Los Alamos: IEEE, 2000. Pp. 36-40.
13. Dinesh Kumar, D.S. Analysis and Design of Principal Component Analysis and Hidden Markov Model for face recognition/ D.S. Dinesh Kumar, P.V.Rao // Procedia Materials Science. 2015. Vol. 10. Pp 616-625.
 14. Sharif, M. Sub-Holistic Hidden Markov Model for Face Recognition/ M. Sharif, J.H. Shah, S. Moshin, M. Rasa// Research Journal of Recent Sciences. 2013. Vol. 2(5). Pp. 10-14.
 15. Miari-Naimi, H. A New Fast and Efficient HMM-Based Face Recognition System Using a 7-State HMM Along With SVD Coefficients/ H. Miari-Naimi, P. Davari// Iranian Journal of Electrical & Electronic Engineering. 2008. Vol. 4, Nos. 46 1 & 2. Pp. 46-57.
 16. Chihaoui, M. Face Recognition Using HMM-LBP/ Chihaoui M., Bellil W., Elkefi A., Amar C.B. //Abraham A., Han S., Al-Sharhan S., Liu H. (eds) Hybrid Intelligent Systems. HIS 2016. Advances in Intelligent Systems and Computing, Springer, Cham, 2016. Vol 420. Pp. 249-258.
 17. Viola, P. Robust Real-Time Face Detection/ P. Viola, M.H. Jones// International Journal of Computer Vision. 2004. Vol. 57(2). Pp. 137–154.
 18. Ng, H.-W. A data-driven approach to cleaning large face datasets. / H.-W. Ng, S. Winkler//Proc. IEEE International Conference on Image Processing (ICIP). Paris, France, Oct. 27-30, 2014.
 19. Aftab, A. LBPH Based Improved Face Recognition At Low Resolution/ A. Aftab, G. Jiandong, A. Fayaz, D. Fahra, A. Awais// 2018 International Conference on Artificial Intelligence and Big Data. IEEE, 2018. Pp. 144-147.
 20. Chen, L. Face recognition based on multi-class mapping of Fisher scores/ L. Chen, H. Man, A.V. Nefian// Pattern Recognition. 2005. Vol. 38. Pp. 799-811.