Метод идентификации человека по изображению лица в системах видеонаблюдения на основе научно-методического аппарата иммунокомпьютинга

Method of person identification by the facial image in video surveillance system using scientific methods of immunocomputing

Соломатин / Solomatin A.
Алексей Юрьевич
ООО "Смартег",
генеральный директор.
г. Санкт-Петербург

Зикратов / Zikratov I.
Игорь Алексеевич
доктор технических наук, профессор.
ФГАОУ ВО «Санкт-Петербургский национальный исследовательский университет информационных технологий, механики и оптики»,
заведующий кафедрой БИТ.
г. Санкт-Петербург

Ключевые слова: видеонаблюдение – video surveillance; идентификация человека по изображению лица – facial identification; классификация объектов – object classification; иммунокомпьютинг – immunocomputing.

Предложен подход для решения задачи идентификации человека по изображению лица для систем видеонаблюдения. Идентификация человека реализуется путем представления изображения лица в виде двумерной матрицы с последующим применением математического аппарата иммунокомпьютинга и алгоритмов предобработки исходного изображения для повышения качества идентификации. Процесс идентификации человека по изображению лица построен на принципе вычисления энергии связи между входным изображением и изображениями в обучающей выборке. Данная величина характеризует степень близости изображений друг к другу. Достоинством предлагаемого метода является простота реализации. В отличие от традиционных методов компьютерного зрения метод не требует значительных вычислительных затрат благодаря сведению процедуры анализа к выполнению простых числовых операций.

Approach for the solution of a problem of the person identification by facial image for video surveillance systems is suggested. Person identification is realized by facial image representation as a two-dimensional matrix with the subsequent use of mathematical immunocomputing apparatus and initial image preprocessing algorithms for identification quality improvement. Identification process is based on binding energy calculation between the initial image and images in the training selection. This quantity characterizes images proximity degree. The advantage of the offered method is realization simplicity. Unlike traditional computer vision methods this method doesn't demand considerable computational burden by reducing the procedure of the analysis to simple numerical operations performance.

Введение

В наши дни увеличивается количество установленных систем видеонаблюдения. Одной из наиболее актуальных задач является идентификация человека, попадающего в зону видимости системы видеонаблюдения. Автоматизация решения этой задачи позволяет существенно снижать трудовые и временные затраты [1]. Кроме этого, появляется возможность снижать влияние человеческого фактора в течение всего времени видеонаблюдения за охраняемым периметром.

В рамках этой статьи предлагается идентифицировать человека по изображению лица. Изображения лиц рассматриваются как объекты, которые обладают рядом признаков и показателей, с помощью которых можно отличить одно изображение лица от другого. Информация, описывающая лицо, представляется в виде большого количества регистрационных данных, для обработки и анализа которых предлагается использовать подход на основе иммунокомпьютинга.

Математический базис, применяемый в рамках вычислительных процедур иммунокомпьютинга, основан на свойствах сингулярного разложения матриц. Сингулярное разложение является удобным методом при работе с матрицами, поскольку показывает геометрическую структуру матрицы и позволяет наглядно представить имеющиеся данные.

Одна из наиболее часто решаемых задач с помощью иммунокомпьютинга – это задача распознавания объектов. В данных задачах для определения связи между объектами A и M вводится понятие энергии связи, которое используется в качестве аналога расстояния между объектами. Наличие данного понятия позволяет применить научно-методический аппарат иммунокомпьютинга в задачах идентификации человека по изображению лица, поскольку соответствует общему

подходу к построению систем идентификации, описанному в информационной теории идентификации [2].

Выделение лиц в видеопотоке

Предлагаемый в рамках данной статьи метод идентификации человека по изображению лица использует для детектирования лиц в видеопотоке метод, разработанный Полом Виолой и Майклом Джонсом в 2001 году [3, 4]. Данный метод по-прежнему остается актуальным и в наши дни, являясь основополагающим для поиска лиц на изображениях в режиме реального времени.

В общем виде задача обнаружения лиц в видеопотоке выглядит следующим образом. Изображение, поступаемое с камеры видеонаблюдения, представляется двумерной матрицей пикселей размерностью mxn. Каждый пиксель данной матрицы имеет значение от 0 до 255, поскольку изображение, поступаемое с камеры видеонаблюдения, является черно-белым.

На следующем шаге применяется подход на основе сканирующего окна для поиска лица на изображении. Поиск осуществляется в активной области изображения прямоугольными признаками, с помощью которых описывается найденное лицо:

$$R_i=\{x,y,m,n,a\}$$

где x,y – координаты центра i-го прямоугольника, m,n – ширина и высота прямоугольника, а a – угол наклона прямоугольника к вертикальной оси изображения.

Предобработка изображений

На качество видеоизображения влияет большое количество самых разнообразных факторов. Из этого следует, что все данные, используемые в процессе идентификации человека по изображению лица, должны быть приведены к общему виду. Стоит отметить, что приведение изображений к общему виду может выполняться с применением разных методов предварительной обработки изображений. В рамках данной статьи предварительная обработка сводилась к преобразованию видеоизображений в монохроматичный формат и корректировке контрастности.

Для перевода цветного изображения в монохроматическое были взяты значения красного, зеленого и синего компонентов пикселя, а затем по формуле:

$$Y'=0.2126R+0.7152G+0.0722B$$

произведены вычисления новых значений для каждого пикселя [5].

На следующем шаге была произведена корректировка контрастности изображений. Под контрастностью изображения понимается отношение яркостей самой светлой и самой темной частей изображения. У более светлого изображения значения пикселей сконцентрированы ближе к концу диапазона градаций серого, у более темного – ближе к началу. Для улучшения качества идентификации физического лица необходимо входное изображение, содержащее максимально широкий спектр значений из диапазона градаций серого. Для получения необходимого результата был применен метод корректировки контрастности изображения, известный как «эквализация гистограммы» [6].

Построение обучающей выборки

Изображение каждого лица, принадлежащее одному классу, представляется в виде вектора:

$$V_i=[p_1\ p_2\ \dots\ p_n]^T,$$

где $i=\overline{1,N}$, где N – количество изображений данного класса. Каждый элемент в таком векторе является значением пикселя исходного изображения – p_j, где $j=\overline{1,n}$, $n=l$xh – размерность вектора, l – ширина рисунка в пикселях, h – высота рисунка в пикселях.

Полученные вектора записываются в виде одного большого вектора:

$$V=[V_1^T V_2^T \dots V_N^T]^T.$$

На следующем шаге вектора для каждого класса изображений в обучающей выборке представляются в виде общей матрицы, где вектор каждого класса записывается в виде колонки в общей матрице:

$$M=[V_1 V_2 \dots V_n]^T.$$

Далее для каждой строки в общей матрице рассчитывается среднее значение по формуле:

$$a_i=\frac{1}{N}\sum_{j=1}^{N} m_{ij},$$

где m_{ij} – элемент i-й строки и j-го столбца матрицы M.

В результате выполнения этой операции полученные значения записываются в виде вектора средних значений:

$$A=[a_1 a_2 \dots a_n]^T.$$

Вектор средних значений A представляется в виде матрицы, на выходе получается усредненное изображение.

$$M_A=[C_1 C_2 \dots C_l],$$

где $C_i=[a_{(i-1)\cdot h+1} a_{(i-1)\cdot h+2} \dots a_{i\cdot h}]$, $i=\overline{1,l}$.

Затем вектор средних значений A вычитается из каждого изображения в обучающей выборке, представленного в виде вектора:

$$S_i=|A-V_i|, i=\overline{1,N}.$$

Далее по вектору разницы считается S_i матрица изображения.

На следующем шаге работы алгоритма по построению обучающей выборки полученное изображение лица представляется в виде матрицы A размерностью MxN, элементы которой – вещественные числа. Из формулы:

$$A=USV^T$$

следует, что любая матрица такого типа может быть представлена с помощью сингулярного разложения.

В результате разложения получаются три матрицы. Столбцы матрицы U и V^T содержат левые и правые сингулярные векторы, которые записываются в качестве обучающей выборки, четко характеризующей изображение лица.

На финальной стадии для каждого класса вычисляется пороговое значение, которое позволяет определить объекты, не относящиеся ни к одному из имеющихся в обучающей выборке классов изображений.

Пусть есть матрицы пикселей для множества изображений данного класса M_z размерности N, где $z=\overline{1,K}$, K– количество классов. Тогда для них считается усредненная матрица M_A, в которой каждый элемент считается по следующей формуле:

$$m_{ij}=\frac{1}{N}\sum_{k=1}^{N} m_{ijk},$$

где m_{ijk}– элемент i-й строки j-го столбца матрицы k-го изображения.

Далее для каждого изображения в выбранном классе рассчитывается энергия связи с матрицей M_A по формуле:

$$\omega=-U^T M_A V$$

где U,V – левый и правый собственные вектора для наибольшего собственного числа SVD разложения матрицы M_A. Из полученных значений выбирается максимальное значение энергии связи, которое дополняется константой:

$$c=\max(\omega)+B$$

Результирующее значение записывается в качестве порогового значения, используемого в процессе идентификации человека по изображению лица.

Идентификация человека по изображению лица

В процессе идентификации с камеры видеонаблюдения поступает изображение лица неизвестного человека, которое представляется в виде матрицы M и проходит все процедуры предобработки. Далее вычисляются энергии связи между входным изображением лица и элементами в обучающей выборке по формуле [7]:

$$\omega_j=-U_j^T M V_j,$$

где

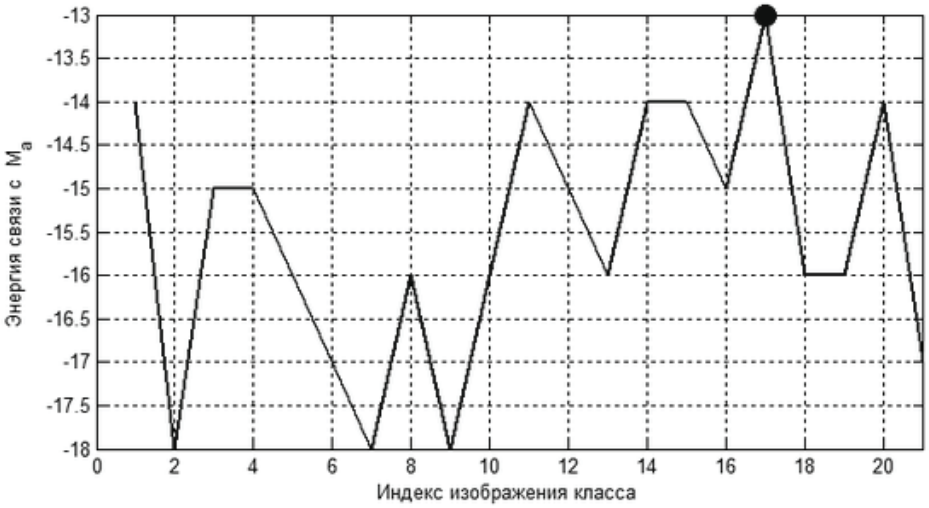

Рис. 1. Значения энергий связи для выбранного класса. Выделено максимальное

$$j=\overline{1,\sum_{l=1}^{k} N_l}.$$

Процесс идентификации физического лица определяется минимальным значением энергии связи по формуле:

$$c = \min(\omega),$$

где ω – вектор всех значений энергий связи с изображениями.

На основе полученного значения энергии связи делается предположение о принадлежности входного изображения лица человека к одному из ранее известных лиц, записанных в обучающую выборку.

Затем энергия связи сравнивается с пороговым значением класса, к которому предположительно относится входное изображение. Если энергия связи превышает пороговое значение, то это свидетельствует о том, что входное изображение не относится ни к одному из представленных лиц в обучающей выборке.

$$\mathrm{Ind}(c,\omega) = z,\ c > Th_z \Rightarrow M \notin M_z$$

где $\mathrm{Ind}(c,\omega)$ – функция поиска индекса класса изображения с минимальной энергией, Th_z – пороговое значение для класса z.

Если энергия связи меньше или равна пороговому значению, то это свидетельствует о том, что процесс идентификации выполнен успешно.

$$\mathrm{Ind}(c,\omega) = z,\ c \le Th_z \Rightarrow M \in M_z.$$

Сравнение метода с другими известными методами

Предлагаемый в данной статье метод сравнивался с методом на основе фильтров Габора и методом EigenFace, которые обычно используются в задачах распознавания человека по изображению лица.

Для 11 классов изображений рассматривалась выборка 11 случайных изображений, принадлежащих классам. Для случайной выборки рассчитывалось коли-

Таблица 1

	Среднее время распознавания картинки, с учетом подгрузки из ФС, мс.	**Средняя вероятность ошибки распознавания изображения, принадлежащего классу, %**
EigenFaces	4.75 ±0.75	5.4 ± 4.9
Gabor Filter	3.75 ± 0.6	20 ± 8.8
Предлагаемый метод	4.5 ±0.045	4.8 ± 5.3

Таблица 2

	Количество ошибок распознавания для первой выборки	**Количество ошибок распознавания для второй выборки**
EigenFaces	0	4
Gabor Filter	79	41
Предлагаемый метод	9	4

чество ошибок распознавания в 40 экспериментах для каждого из трех методов. Результаты представленые в табл. 1 в виде среднего значения с индивидуальным для каждого метода стандартным отклонением.

Также был проведен второй вычислительный эксперимент. Для 11 классов изображений распознавались две выборки, содержащие по 100 изображений одного класса. Для трех методов рассматривалось количество ошибочных распознаваний изображений. Результаты эксперимента представлены в табл. 2.

Заключение

В статье рассмотрена задача идентификации человека по изображению лица, которая представляет научный и практический интерес благодаря автоматизации процесса обработки видеоизображений, которые поступают с камеры видеонаблюдения. Новизна заключается в применении нового метода идентификации человека по изображению лица на основе математического аппарата иммунокомпьютинга.

Изображения лиц предлагается представлять в виде двумерных матриц с применением методов предобработки изображений. Идентификация человека по изображению лица осуществляется путем применения математического аппарата иммунокомпьютинга, который позволяет вычислять энергию связи между двумя объектами. Применительно к поставленной задаче энергия связи характеризует близости входного изображения лица к лицам, находящимся в обучающей выборке.

Представленное в статье решение задачи по идентификации человека по изображению лица имеет ряд ярко выраженных достоинств. Одно из главных достоинств – простота реализации, которая требует низких вычислительных затрат благодаря наличию простых математических операций.

Также стоит отметить, что предлагаемый в данной статье метод идентификации по изображению лица отличается от метода на основе фильтров Габора и метода EigenFace следующими особенностями:

1. Работает с набором неструктурированных данных, прост в реализации и не требует дополнительных настроек.
2. В качестве меры невязки между входным изображением и изображением из обучающей выборки используется энергия связи.
3. Присутствуют операции, связанные с предобработкой исходных изображений с целью нормализации световых и контрастных характеристик исходных изображений.
4. Присутствуют дополнительные преобразования для расчета порогового значения для каждого класса.
5. Уменьшение размерности исходных данных осуществляется за счет сингулярного разложения матриц, характеризующих исходные изображения лиц.

Программная реализация, полученная в рамках решения частных задач, позволяет использовать предложенное решение как в режиме анализа записанного видеофайла, так и анализа видеопотока в режиме реального времени.

На данный момент остались незатронутыми такие темы как влияние разных световых характеристик на качество идентификации и зависимость количества классов в обучающей выборке на скорость идентификации человека по изображению лица. Также в рамках дальнейших научных исследований предполагается изучить работу предложенного метода на выборке изображений, полученных в неконтролируемых условиях.

Литература

1. Лукьяница, А. А. Цифровая обработка видеоизображений / А.А. Лукьяница, А.Г. Шишкин. – Москва. Изд-во: Ай-Эс-Эс Пресс, 2009. – 518 с.

2. Цыпкин, Я. З. Информационная теория идентификации / Я.З. Цыпкин. – Москва. Изд-во: Физико-математическая литература РАН, 1995. – 336 с.

3. Viola, P. Rapid Object Detection using a Boosted Cascade of Simple Features / P. Viola, M.J. Jones. // IEEE Conf. on Computer Vision and Pattern Recognition. – 2001. – P. 9.

4. Viola, P. Robust real-time face detection / P. Viola, M.J. Jones // International Journal of Computer Vision. – Vol. 57. – 2004. – P. 137–154.

5. Standard Default Color Space for the Internet – sRGB / M. Stokes [et al.] [Электронный ресурс]. – Режим доступа: http://www.w3.org/Graphics/Color/sRGB, свободный. – Загл. с экрана.

6. Gonzalez, R. C. Digital Image Processing / R. C. Gonzalez, R. E. Woods. – Vol. 3, 2008. – 730 p.

7. Tarakanov, A. Immunocomputing: Principles and Applications / A. Tarakanov, V. Skormin, S. Sokolova. – NY: Springer-Verlang New York Inc., 2003. – 193 p.