

Л.А. Сорокин

(Академия ГПС МЧС России; e-mail: l.40in@ya.ru)

ПУТИ ОПТИМИЗАЦИИ СИСТЕМЫ ВИДЕОКОНТРОЛЯ С ВОЗМОЖНОСТЬЮ РАСПОЗНАВАНИЯ ЛИЧНОСТИ

Анализируется нагрузка системы видеоконтроля при распознавании лиц. Даны рекомендации и предложены методы повышения эффективности подобных систем.

Ключевые слова: распознавание личности, системы видеоконтроля.

L.A. Sorokin

OPTIMIZATION WAYS OF SAFETY SYSTEM WITH POSSIBILITY OF IDENTIFICATION

The load of the system video surveillance in face recognition was analyzed. Recommendations and proposed methods to improve the efficiency of such systems was given.

Key words: face recognition, video surveillance.

Статья поступила в редакцию Интернет-журнала 1 августа 2016 г.

Для повышения эффективности распознавания личности в интересах обеспечения безопасности предлагается создание автоматизированной системы, способной на основе биометрии лица идентифицировать человека в местах массового пребывания людей. Технологии распознавания на основе биометрии лица довольно широко используются правоохранительными и государственными органами в зарубежных странах. Известно об успешном применении данных технологий военной разведкой США в Афганистане при отслеживании перемещения террористов, полицией Нью-Йорка, Чикаго, Сан-Диего для поиска преступников [1].

Однако, существующие *системы безопасности (СБ)* с возможностью распознавания личности на основе уникальности биометрии лица [8] не предназначены для автоматической идентификации людей в режиме "реального" времени в местах с довольно большим человеческим потоком. Фактическое их назначение – оказание помощи сотрудникам подразделений безопасности в наблюдении за текущей обстановкой на объекте и автоматическое выявление интересантов по заданной заранее биометрии лица.

В данной статье описывается модель работы такой СБ, проанализирована её эффективность от числа и расположения видеокамер. Даны рекомендации и предложены методы повышения эффективности функционирования СБ.

Модель функционирования СБ с нагрузкой от сети видеоконтроля

В процессе функционирования крупная СБ, особенно при обеспечении массовых мероприятий, использует большое число источников данных. В качестве источников данных рассматриваются видеокамеры. Оператору необходимо анализировать большое количество видеоматериала. Используя современные системы, например, "Каскад-поток" [8], оператор часто вынужден отдавать команды системе на принудительное распознавание.

Ниже будет анализироваться СБ типовой архитектуры (рис. 1). **Модуль детектирования (МД)** обеспечивает автоматическое обнаружение изображений лиц в видеопотоке от источника видеосигнала. В **ядре идентификации (ЯИ)** сравниваются изображения лиц, полученных в результате детектирования и размещённых ранее в хранилище.

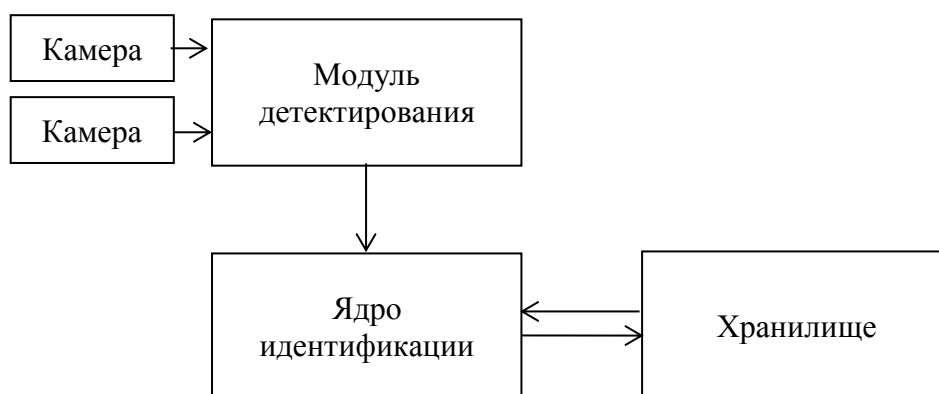


Рис. 1 Архитектура СБ

Выстроим математическую модель работы СБ, в которой будет видна зависимость скорости реакции СБ с оператором от числа и расположения камер, а также параметров регистрируемой информации. Будем строить модель, исходя из информации, используемой в существующих на сегодняшний день алгоритмах.

ЯИ и МД содержат программный код, основанный на наборе алгоритмов:

$$A_{\text{собр}} = \{A_1, A_2, \dots, A_p\}, \quad (1)$$

где p – число алгоритмов, используемых в МД и ЯИ СБ.

Определим время $t_{\text{обр}}$, затраченное на реализацию данных алгоритмов для распознавания одного человека и сопоставление с одним изображением лица:

$$t_{\text{обр}}^1 = F(A_{\text{собр}}) + t_0, \quad (2)$$

где F – функционал на множестве алгоритмов распознавания $A_{\text{собр}}$;

t_0 – начальный отрезок времени, не связанный с обработкой изображения, необходимый для его "захвата" и получения в нужном для обработки виде, одинаковый для всех камер.

Будем считать, что функционал F на наборе алгоритмов A_{cob} обладает свойством аддитивности, тогда

$$t_{\text{обp}}^1 = \sum_{i=0}^p F(A_i) + t_0, \quad (3)$$

где

$$A_i \in R^n, F(A_i) \rightarrow R. \quad (4)$$

Ряд алгоритмов, применяемых в СБ, используются в ЯИ, а ряд – в МД. Исходя из этого, уместно записать:

$$A_{\text{cob}} = \{A_{\text{яи}} \mid A_{\text{мд}}\}, \quad (5)$$

где $A_{\text{яи}}$ – множество алгоритмов, применяемых в ЯИ;

$A_{\text{мд}}$ – множество алгоритмов, применяемых в МД,

тогда справедливо следующее выражение:

$$t_{\text{обp}}^1 = F(A_{\text{яи}} \mid A_{\text{мд}}) + t_0 \quad (6)$$

или

$$t_{\text{обp}}^1 = F(A_{\text{яи}}) + F(A_{\text{мд}}) + t_0. \quad (7)$$

Если изображение искомого лица сопоставляется с m изображениями, хранимыми в БД, то время обработки определяется следующим образом:

$$t_{\text{обp}}^m = m F(A_{\text{яи}}) + F(A_{\text{мд}}) + t_0. \quad (8)$$

Далее будем считать, что в один момент времени с одной камеры поступает одно изображение лица, тогда распознавание изображений лиц, поступающих из сети, состоящей из k видеокамер, можно выразить следующим образом:

$$t_k^m = k (m F(A_{\text{яи}}) + F(A_{\text{мд}})) + t_0. \quad (9)$$

Поскольку t_0 – случайная величина, то в выражении (9) можно использовать формулу математического ожидания:

$$t_0 = M[T_0] = \int_0^{+\infty} t_0 dF(t_0), \quad (10)$$

где $F(t_0)$ – функция распределения случайной величины.

Для простоты будем оценивать t_0 как

$$t_0 = \frac{1}{2}(t_0^{\text{ниж}} + t_0^{\text{вpx}}), \quad (11)$$

где $t_0^{\text{ниж}}$ и $t_0^{\text{вpx}}$ – статистически рассчитанные нижний и верхний пределы для отрезка времени t_0 .

В формуле (9) отражена зависимость характеристики СБ с распознаванием лиц от числа видеокамер и объёма хранилища.

На основании формулы (9) можно получить асимптотическую временную сложность работы СБ ($f(k, m)$) с возможностью распознавания человека на основе уникальности биометрии лица:

$$f(k, m) = O(k(m + 1)), \quad (12)$$

где m – количество изображений искомого лица в хранилище;

k – число камер в сети видеоконтроля.

Однако выражение (12) представляет худший случай, то есть когда изображения лиц поступают со всех камер в сети видеоконтроля, а изображения в хранилище не отсортированы и сравнение осуществляется в произвольном порядке.

Оценим среднее время работы СБ ($g(k, m)$).

Распознавание интересанта представляется элементарным событием ω_j , где j является идентификационным номером личности, изображение которого содержится в хранилище ($j \in \{1, \dots, m\}$). Множество элементарных событий образуют пространство элементарных событий $\Omega = \{\omega_1, \dots, \omega_m\}$. В нашем случае алгебра событий A совпадает с множеством элементарных событий, то есть $A = \Omega$. Так как сравнение изображения распознаваемого искомого происходит со случайно выбранным изображением из хранилища, то будем считать, что элементарные события равновероятны, то есть $\forall \omega_j \in \{\omega_1, \dots, \omega_m\}$:

$$P(\omega_j) = \frac{1}{m}. \quad (13)$$

Таким образом, на вероятностном пространстве $(\Omega, A, P(\cdot))$ зададим случайную величину ζ как число необходимых сравнений для поиска личности в хранилище.

Можно показать, что случайная величина ζ является дискретной, равномерно распределённой на множестве $\{1, \dots, m\}$. Тогда математическое ожидание случайной величины ζ будет следующим [3]:

$$M[\zeta] = \sum_{i=1}^m x_i p_i, \quad (14)$$

где

$$\zeta(\omega_i) = x_i \text{ и } P(\zeta(\omega_i) = x_i) = p_i. \quad (15)$$

Путём преобразований (15) можно получить следующее выражение для математического ожидания:

$$M[\zeta] = \frac{m+1}{2}. \quad (16)$$

Математическое ожидание (16) отражает среднюю вычислительную сложность работы СБ при распознавании одного искомого. Вычислительная сложность обработки видеопотока с k камер выражается следующим образом:

$$g(k, m) = O\left(k \left(\frac{m+1}{2}\right)\right). \quad (17)$$

Таким образом, если изображения в хранилище не отсортированы и сравнение осуществляется в произвольном порядке, средняя сложность имеет вид выражения (17).

Но в случае наличия хранилища, содержащего достаточно большое количество изображений интересантов, основной составляющей времени распознавания является время, затраченное на запрос и получение информации об изображении лица из хранилища и сравнение с ним [6]. Далее повысим эффективность работы СБ путём уменьшения количества операций сравнений.

Повышение эффективности СБ

Повышение эффективности СБ, а именно – повышение скорости реагирования предлагается достигать за счет распараллеливания вычислений и уменьшения количества операций сравнения. Предлагается выделять группы камер, внутри которых обработка будет происходить последовательно, а вне группы – параллельно. При этом время распознавания в группе будет иметь ту же линейную зависимость. А для всей СБ время распознавания будет определяться временем распознавания в группе, в котором максимально

$$t_k^m = \max_i t(i), \quad (18)$$

где

$$t(i) = r_i (l_i F(A_{\text{ян}}) + F(A_{\text{мд}})) + t_0 \text{ и } r_i < k, l_i < m \quad (19)$$

i – индекс группы;

l_i – количество анализируемых из хранилища изображений лиц интересных в i -й группе;

r_i – число камер в i -й группе сети видеоконтроля.

Таким образом, для эффективной работы СБ должна быть настроена так, чтобы

$$r_i (l_i F(A_{\text{ян}}) + F(A_{\text{мд}})) + t_0 \rightarrow \min. \quad (20)$$

При этом необходимо, чтобы соблюдался ряд дополнительных условий. Проведем анализ ошибок первого и второго рода при работе СБ. Оценим, как часто будут возникать пропуски и ложные совпадения (используем расчёты 3).

Пусть система идентификации установлена на пункте контроля. Вероятность пропуска лица из хранилища равна FRR. Если принять допустимой одну ошибку пропуска интересанта при использовании информации о l_i человек, которые анализируются в блоке видеокамер, то

$$\text{FRR } l_i = 1. \quad (21)$$

Выразим l_i

$$l_i = \frac{1}{\text{FRR}}. \quad (22)$$

Вероятность возникновения ошибки ложного совпадения для одного человека равна p_{fls}^0 :

$$p_{\text{fls}}^0 = \text{FAR } l_i. \quad (23)$$

А вероятность возникновения ошибки ложного совпадения для потока изображений с r_i видеокамер равна p_{fls}^1 :

$$p_{\text{fls}}^1 = \text{FAR } r_i l_i. \quad (24)$$

Если принять допустимой одну ошибку ложного совпадения, то число ложных совпадений можно вычислять по формуле:

$$l_i = \frac{1}{(\text{FAR} * r_i)} \quad (25)$$

или с учётом (20)

$$l_i = \frac{\text{FRR}}{\text{FAR}}. \quad (26)$$

Этот набор условий позволяет определять эффективность распознавания внутри группы видеокамер с учётом появления ложных срабатываний.

Другой набор условий может определять количество камер, которые необходимо использовать для поиска искомого. В общем случае, когда входные условия поиска всегда одинаковые, то для поиска и регистрации привлекаются все камеры и все кластеры. В этом случае (9) имеет максимальное значение.

Теперь предположим, что выбор камер, которые будут использоваться для идентификации искомого, зависит от группы из h условий:

$$G = \{g_1, g_2, \dots, g_h\}, \quad (27)$$

тогда количество камер в блоке

$$r_{i0} = \varphi^i(G) \quad (28)$$

и при этом

$$r_{i0} \leq r_i. \quad (29)$$

С математической точки зрения это означает, что в (20) параметр r_i будет уменьшаться, и, следовательно, будет уменьшаться и искомое значение t_k^m .

Если ввести параметры нумерации камер и кластеров СБ, то в качестве условий (27) можно задавать номера камер следующим образом:

- с привязкой к предварительной информации о местах пребывания иско-
мых;
- с привязкой к "следам" их пребывания на основе анализа видеоданных
из камер на искомого за предыдущие периоды работы СБ;
- с привязкой к местам, где они чаще всего появляются, с заданием частоты
появления в качестве весового коэффициента для вершины графа.

Тогда, решая задачу (20) одним из известных методов обхода графа, можно в разы сократить параметр t_k^m , отражающий суммарное время работы распознающей части и существенно снизить нагрузку на вычислительные средства обрабатывающего центра СБ.

Далее в качестве условий объединения камер в группы рассмотрим задание возможных мест появления физических лиц. Данную кластеризацию можно осуществить следующими способами:

1. Отслеживание движения. При попадании физического лица в объектив камеры, распознавать его, а затем отслеживать его передвижение. Но почти у всех методов отслеживания движения есть недостатки: принятие элементов фона в качестве движущихся объектов, низкая производительность системы трекинга, многочисленные ошибки при перекрытии объектов, игнорирование объектов при их малом размере или малой амплитуде движения и т. д.

2. Построение вероятностного графа маршрутов (рис. 2). Сеть видеоконтроля с установленными в разных местах видеокамерами можно условно рассматривать в качестве нагруженного графа, ребра которого несут определенную нагрузку. В качестве такой нагрузки будем рассматривать значение вероятности (q_{ij}) того, что искомый пришёл в конечную вершину (j) из начальной (i). Следовательно, рассмотрим ориентированный граф возможных маршрутов передвижения искомого. Вершинами графа являются места видеофиксации (вход, касса и прочее). Каждое ребро графа помечено меткой, значение которой является вероятностью того, что искомый в предыдущий момент времени находился в вершине (i), при условии, что сейчас он находится в вершине (j).

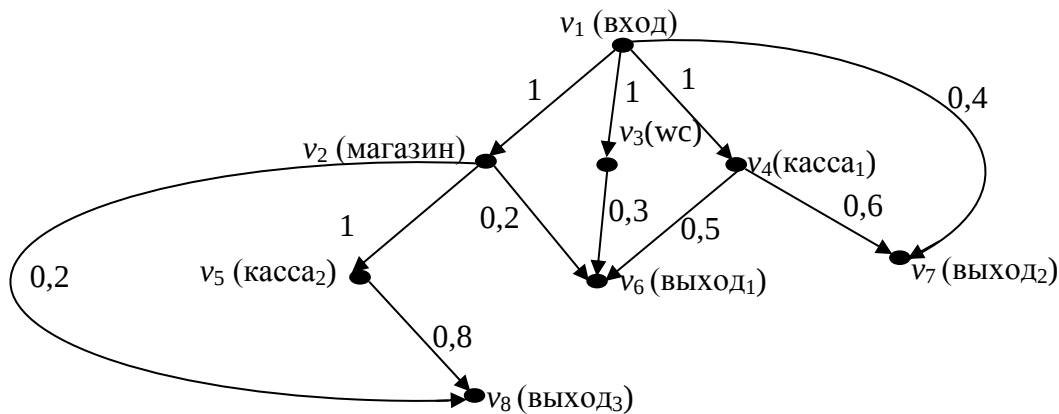


Рис. 2. Пример вероятностного графа маршрутов, построенный на основе схемы Савёловского вокзала

Для построения графа вероятностных маршрутов можно воспользоваться теорией случайных процессов в части, касающейся конечных однородных Марковских цепей.

Рассмотрим следующий случайный процесс на фазовом пространстве E

$$\{\xi_n, n \in N\}, \quad (30)$$

где ξ_n – случайная величина, равная идентификатору камеры, которая зафиксировала искомого;

n – порядковый номер камеры, которая зафиксировала искомого (момент времени);

$E = \{1, \dots, m\}$ – фазовое пространство. Элементами данного множества являются идентификаторы камер.

Проверим, что случайный процесс, описанный в выражении (30), является **конечной однородной Марковской цепью (КОМЦ)**:

1. Данный процесс является цепью, поскольку временное пространство $T = N$ – дискретно.

2. Будем считать, что цепь, описанная в выражении (30) обладает Марковским свойством, то есть $\forall n \in N, \forall i, j, i_{n-1}, \dots, i_0 \in E$ верно равенство:

$$P\{\xi_{n+1} = j \mid \xi_n = i, \xi_{n-1} = i_{n-1}, \dots, \xi_0 = i_0\} = P\{\xi_{n+1} = j \mid \xi_n = i\}. \quad (31)$$

Данное допущение вполне уместно, так как то в поле видимости, какой камеры окажется искомый в момент времени $n + 1$, зависит только от того, где он был в момент времени n .

3. Конечность Марковской цепи следует из того, что мощность фазового пространства E равна m .

4. Будем считать, что условные вероятности $P\{\xi_{n+1} = j \mid \xi_n = i\}$ не зависят от n , то есть $\forall n \in N, \forall i, j \in E$:

$$P\{\xi_{n+1} = j \mid \xi_n = i\} = \dots = P\{\xi_1 = j \mid \xi_0 = i\}. \quad (32)$$

Тогда случайный процесс, описанный в выражении (30) является однородным.

Далее на основе статистической информации необходимо задать матрицу переходных вероятностей:

$$\Pi = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & \ddots & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix}, \quad (33)$$

где p_{ij} – вероятность того, что искомый будет находиться в поле видимости камеры с идентификатором j при условии, что до этого он находился в поле видимости камеры с идентификатором i .

Тогда метки рёбер q_{ij} вероятностного графа маршрутов могут быть получены следующим образом:

$$q_{ij} = P\{\xi_n = i | \xi_{n+1} = j\} = \frac{P\{\xi_n = i\} * P\{\xi_{n+1} = j | \xi_n = i\}}{P\{\xi_{n+1} = j\}}. \quad (34)$$

Описанная выше выкладка позволяет произвести распознавание искомого следующим образом: при распознавании искомого в узле сети видеоконтроля необходимо сравнивать лицо с уже идентифицированными лицами в смежных вершинах в порядке уменьшения значения метки ребра (например, при идентификации в узле v_6 необходимо сравнивать сначала изображения, зафиксированные в узле v_4 , далее если соответствие не установлено, то и в v_3 , и v_2).

Далее из свойств КОМЦ следует 4, что фазовое пространство $E = \{1, \dots, m\}$ разбивается в объединение не пересекающихся классов взаимно сообщающихся состояний:

$$E = E_1 \sqcup \dots \sqcup E_{n_1} \sqcup \{j_1\} \sqcup \dots \sqcup \{j_{n_2}\}, \quad (35)$$

где $E_1, \dots, E_{n_1}$ – непересекающиеся классы взаимно сообщающихся существенных непоглощающих состояний;

$\{j_1\}, \dots, \{j_{n_2}\}$ – существенные поглощающие состояния.

Таким образом, множество идентификаторов камер распадается на непересекающиеся классы. При распознавании искомого в разных группах осуществляется обращение к разным исходным данным. В результате применения параллельной обработки данных между группами будет увеличено быстродействие СБ в целом.

Вывод

Анализ СБ показал возможность создания модели работы алгоритмов СБ при распознавании лиц с нагрузкой от сети видеоконтроля. В общем виде она определяется формулой (9). Анализ условий возможного применения модели показал наличие ряда характеристик, фиксируемых и определяемых совместно с распознаванием, которые позволяют существенно уменьшить суммарное время обработки данных и улучшить точность распознавания. Это возможно обеспечить за счет привязки к номеру камеры данных о местах и частотах пребывания искомого в определённых местах в определённый период. Тогда объединение камер в группы, например с использованием вероятностного графа маршрутов, позволит обрабатывать информацию внутри групп – последовательно, а вне групп – параллельно.

Литература

1. **Melinda** Ham Face Recognition Technology. <http://www.uts.edu.au/about/faculty-law/news/face-recognition-technology>.
2. **FaceVACS-DBScan** // Cognitec Systems: [сайт]. URL: <http://www.cognitec.com/>
3. **Гнеденко Б.В.** Курс теории вероятности. М.: Едиториал УРСС, 2005.
4. **Кельберт М.Я., Сухов Ю.М.** Вероятность и статистика в примерах и задачах. Т. II: Марковские цепи как отправная точка теории случайных процессов и их приложения. М.: МЦНМО, 2009.
5. **Сорокин Л.А.** Автоматизированная система скрытой идентификации личности человека // Матер. 22-й междунар. науч.-техн. конф. "Системы безопасности – 2013". М.: Академия ГПС МЧС России, 2013. С. 246-249.
6. **Сорокин Л.А.** Распределенная модель хранения и поиска информации о личности на открытом множестве // Матер. IV Всерос.науч.-практ. конф. с междунар. уч. "Проблемы обеспечения безопасности при ликвидации последствий чрезвычайных ситуаций". Воронеж, Воронежский институт ГПС МЧС России, 2015. № 1. С. 436-439.
7. **Рост** преступности в России связали с экономическим кризисом. <http://www.rbc.ru/politics/01/12/2015/565db2189a7947e785881f88>.
8. **Описание** характеристик системы "Каскад-поток". www.technoserv.com.