Описание изображений сложных наземных объектов в задаче распознавания образов

11, ноябрь 2012

DOI: 10.7463/1112.0500490

Гулевич С. П., Веселов Ю. Г., Прядкин С. П.

УДК 778.35:629.7

Россия, МГТУ имени Н.Э. Баумана

Россия, ОАО «Камов»

Россия, ВУНЦ ВВС «Военно-воздушная академия имени профессора Н.Е. Жуковского и

Ю.А. Гагарина»

vesel_foto@mail.ru

Введение

Эффективность применения современных беспилотных летательных аппаратов (БЛА) против террористических, контрабандных, наркотических и других угроз, связанных с несанкционированным проникновением на охраняемые территории во многом зависит от качества систем наблюдения, сбора и обработки информации, управления и контроля, а также систем принятия решений. В качестве одной из основных систем сбора и обработки информации на борту БЛА используется оптико-электронная система (ОЭС). Проектирование ОЭС БЛА должно исходить из требований, предъявляемых условиями применения БЛА, комплексом алгоритмов, подлежащих реализации на борту БЛА, точностью представления входной и выходной информации, а также динамикой управления БЛА. Задача распознавания изображений наземных объектов в БЛА является ключевой для приближения БЛА к возможностям пилотируемой авиации. Эффективное решение указанной задачи повышает вероятность обнаружения наземных объектов с использованием ОЭС.

Успешность выполнения задачи распознавания сложных трехмерных объектов по их двумерному изображению, формируемому ОЭС,

характеризуется показателями эффективности. В большинстве систем технического зрения в качестве показателей используются вероятность правильного распознавания (классификации) или достоверность распознавания. Очевидно, что показатели эффективности зависят от дальности, маскировки, метеоусловий и целого ряда случайных факторов. Поскольку при измерениях сигнала от регистрирующего датчика (тепловизора, телекамеры или фотокамеры) возникают шумы, измеряемые величины носят случайный характер. Случайные факторы можно в зависимости от причин их возникновения разделить на группы. Первая группа факторов определяется самим излучающим или отражающим объектом: характером отражения и рассеивания сигнала от его поверхности. Ввиду структурной сложности изображений наземного объекта, многовариантности его ракурсов, окрасок, наконец, внешнего оборудования и средств маскировки эта группа имеет детерминированную и случайную составляющие. Вторая группа связана с воздействием среды на формирование изображения и описывается характером рассеивания (поглощения) оптического излучения в атмосфере. Оно зависит от метеоусловий, наличия турбулентности и тепловых потоков в атмосфере на участке луча “БЛА - объект”. Третья группа случайных факторов связана с приемником изображения ОЭС. Здесь происходит оцифровка изображения - квантование по уровню яркости сигнала и квантование в пространстве. В процессе квантования вносятся погрешности, обусловленные чувствительностью элементов матрицы приемника оптического излучения и конечными размерами каждого элемента. К этой же группе относятся помехи: естественные (ракурс солнца, зеленые насаждения, блики от водных поверхностей, стекла, крыш и других поверхностей) и искусственные (ослепление лазерами и другими мощными источниками).

Особенностями ОЭС БЛА, отличающими их от других систем технического зрения, являются: устойчивость к высокому уровню помех, обусловленных нестационарными условиями экспозиции сцены; работа в условиях реального времени; ограничения на вычислительные ресурсы и

ресурсы энергоснабжения. Особо следует выделить возможность адаптации ОЭС к новым объектам, хотя обычно ОЭС работают по образцу. В этой связи, работу системы технического зрения можно рассматривать на 2 этапах:

- проектирование и первичное обучение,
- эксплуатация и дообучение с целью распознавания новых объектов.

Рассмотрим некоторые принципы построения систем технического зрения БЛА, которые закладываются на первом этапе и определяют потенциальные возможности по распознаванию наземных объектов. Общая схема решения задачи распознавания приведена на рисунке 1.

Рисунок 1 – Схема решения задачи распознавания

Изображение цели, сформированное в фокальной плоскости оптическим каналом и оцифрованное матричным приемником ОЭС, будем обозначать $\mathbf{Z}$. По-существу, $\mathbf{Z}$ представляет собой двумерную матрицу размером $N \times M$ (размер фотоприемника в пикселях). Последовательно считывая строки матрицы $\mathbf{Z}$, преобразуем её в вектор $\vec{\mathbf{Z}}$ размерности $M \times N$. Вектор $\vec{\mathbf{Z}}$ будем называть образом наблюдаемого объекта $\mathbf{Z}$. Будем считать, что образ объекта может принадлежать к классу ω_s множества классов целей $\Omega = \{\omega_s, s=1 \dots K\}$. Множество Ω счетно. Каждый класс определяется конкретной моделью объекта. Алгоритм соотнесения образа объекта к одному из классов ω_s будем называть классификатором. Образу $\vec{\mathbf{Z}}$ в L -мерном метрическом пространстве

признаков соответствует вектор $\vec{X}$. Из рисунка 1 видно, что выбор информационных признаков перед первичным обучением системы технического зрения, после обучения имеет определяющее влияние на эффективность процесса распознавания наземного объекта.

В качестве признаков, как правило, выбирают следующие признаки изображения:

- геометрические (отрезки прямых, дуги – обобщенный метод Хафа, детекторы углов –Харриса, Томаши, контуры и другие особенности изображения);
- спектральные (спектральные характеристики на основе Фурье или вейвлет преобразования);
- структурные (на основе морфологического анализа и других методов);
- энергетические или яркостные;
- статистические (на основе анализа распределений и статистик).

Без ограничения общности могут быть использованы любые признаки и их комбинации для обучения системы технического зрения. Будем предполагать, что при фиксированном выборе векторного пространства признаков классификатор обеспечивает оптимальное решение по распознаванию.

Постановка задачи.

Выбор наиболее информативных признаков, описывающих множество классов объектов $\Omega = \{\omega_j, j=1...K\}$, является ключевым моментом в задаче распознавания сложных по форме, спектру и другими признакам объектов. Требуется сформулировать принципы отбора информационных признаков образов, обеспечивающих решение задачи распознавания образов из K классов по критерию максимума вероятности распознавания $P_{\text{расп } X}$ заданным набором $\vec{X}$ любого объекта из K классов. Классы ω_j и ω_i являются не пересекающимися: если $\vec{X} \in \omega_j$ и $\vec{X} \in \omega_i$, то следует $\vec{X} \in \emptyset$.

Допущения:

1 в кадре присутствует информация только об одном нормализованном образе;

2 векторы признаков $\vec{X}$ для всех классов имеют одинаковый закон распределения.

Разделимость классов с точки зрения теории информации

Для оценки информативности признаков воспользуемся положениями теории информации - используем понятия *объем информации* $I(\vec{Z})$ и *энтропия* $H(\vec{Z})$. Ральф Хартли в работе [1] предложил объем информации $I(\vec{Z})$, содержащийся в образе и связанный с реализацией события $\vec{Z} = \vec{z}_k$ с вероятностью p_k , определять как логарифмическую функцию:

$$I(\vec{z}_k) = \log \left(\frac{1}{p_k} \right) = -\log p_k, \quad (1)$$

где основание логарифма, как правило, равно 2 (единица информации - bit) или e (единица информации - nat). Для определенности в работе выбрана единица e . Ввиду того, что разрешение приемника ОЭС определено – $M \times N$, тогда множество реализаций события $\vec{Z} = \vec{z}_k$ – счетно и ограничивается условием:

$$0 \leq I(\vec{z}_k) \leq M \times N.$$

Здесь предполагается, что минимальный объем информации соответствует достоверному событию ($p_k = 1$), а максимальный соответствует событию с наибольшей неопределенностью - реализация образа в виде 1 точки $p_k = 1/MN$).

В работе Клода Шеннона [2] введено среднее значение объема информации $I(\vec{z}_k)$ при всех возможных реализациях образа называемое *энтропией* $H(\vec{Z})$:

$$H(\vec{Z}) = E[I(\vec{z}_k)] = \sum_{k=1}^{MN} p_k \cdot I(\vec{z}_k) = -\sum_{k=1}^{MN} p_k \cdot \log p_k. \quad (2)$$

Считая, что $0 \cdot \log 0 = 0$ для энтропии $H(\vec{Z})$ получим диапазон:

$$0 \leq H(\vec{Z}) \leq \log(M \times N).$$

Прежде чем приступить к решению задачи по выбору наиболее информативных признаков ответим на вопрос: какой из двух наборов признаков F_X и G_Y с распределениями $f_X(x)$, $g_Y(y)$ несет больший объем информации об образе? Здесь $f_X(x)$ - гауссово L -мерное распределение случайного вектора признаков $\vec{X}$, $g_Y(y)$ – отличное от гауссового L -мерное распределение случайного вектора признаков $\vec{Y}$. Распределения $f_X(x)$, $g_Y(y)$ имеют одинаковые значения первого и второго статистических моментов.

В работе [3] был сформулирован принцип максимальной энтропии: “Если выводы основываются на неполной информации, она должна выбираться из распределения вероятности, максимизирующего энтропию при заданных ограничениях на распределение”. В [4] доказано, что принцип максимальной энтропии корректен и существует только одно распределение, обеспечивающее максимум энтропии, которое можно выбрать с помощью "аксиом согласованности":

- уникальность (результат должен быть единственным);
- инвариантность (выбор системы координат не должен влиять на результат);
- независимость системы (не должно иметь значения, будет ли независимая информация о независимых системах браться в расчет как в терминах различных плотностей отдельно, так и вместе, в терминах совместной плотности);
- независимость подмножеств (не должно иметь значения, будет ли независимое множество состояний системы рассматриваться в терминах условной плотности или полной плотности системы).

Для случая многомерного гауссового распределения $f_X(\vec{X})$ центрированного вектора признаков $\vec{X}$ известно выражение[7]:

$$f_X(\vec{X}) = \frac{1}{(2\pi)^{\frac{L}{2}}(\det(\Sigma))^{1/2}} \cdot \exp\left(-\frac{1}{2} \vec{X}^T \Sigma^{-1} \vec{X}\right). \quad (2a)$$

Тогда доопределим энтропию для непрерывного вектора признаков $\vec{X}$ следующим выражением [7]:

$$H(\vec{X}) = \int f_X(\vec{X}) \log(f_X(\vec{X})) d\vec{X}. \quad (26)$$

Последняя величина получается из (2) как предельный переход по количеству состояний для вектора признаков $\vec{X}$ и в литературе носит название дифференциальной энтропии. Поскольку на практике, как правило, используется квантованная оценка признаков и выборочные оценки вектора математического ожидания и матрицы ковариации, то нас будет интересовать только вывод некоторых выражений для оценки энтропии с переходом к дискретному представлению. Подставляя выражение (2а) в интеграл (26), для многомерного гауссового распределения $f_X(\vec{X})$ получено выражение для энтропии $H(\vec{X})$ [4]:

$$H(\vec{X}) = \frac{1}{2} [L + L \cdot \log(2\pi) + \log|\det(\Sigma)|], \quad (3)$$

где L - размерность вектора признаков $\vec{X}$, Σ - матрица ковариации случайного вектора признаков $\vec{X}$, $\det(\Sigma)$ - определитель матрицы Σ . Из (3) вытекает 2 вывода.

1. Если вектор признаков $\vec{X}$ описывается гауссовой статистикой, то энтропия ее имеет наибольшее значение среди прочих векторов признаков $\vec{Y}$, описываемых иной статистикой с таким же вектором средних значений, при этом выполняется неравенство: $H(\vec{X}) \geq H(\vec{Y})$, причем равенство достигается только при совпадении распределений $\vec{X}$ и $\vec{Y}$.

2. Энтропия $H(\vec{X})$ гауссового случайного вектора признаков $\vec{X}$ однозначно определяется его ковариационной матрицей Σ и не зависит от среднего значения вектора признаков $\vec{X}$.

На основании сделанных выводов может быть сформулирован *первый принцип*: в соответствии с принципом максимальной энтропии считаем, что вектор признаков $\vec{X}$ описывается гауссовой статистикой.

Поскольку энтропия $H(\vec{X})$ для наиболее информативного набора признаков $\vec{X}$ образа определяется его ковариационной матрицей Σ , то учитывая ее симметричный характер можно использовать разложение Карунена-Лоева [5]:

$$\Sigma = \Psi^T \cdot \Lambda \cdot \Psi, \quad (4)$$

$$\Lambda = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_L \end{bmatrix},$$

где Λ - диагональная матрица состоящая из собственных значений λ_i , а Ψ - матрица составленная из столбцов – собственных векторов, соответствующих λ_i . Заметим что ранг матрицы Σ т.е. $L^I = \text{rank}(\Sigma) \leq L$ определяет количество линейно независимых векторов признаков, отвечающих за классификацию $\vec{X}$. Выполняя процедуру нормализации для собственных значений и собственных векторов Σ , получим нормализованные собственные значения μ_i :

$$\mu_i = \frac{\lambda_i}{\text{tr}(\Sigma)}.$$

Учитывая тот факт, что Ψ соответствует ортогональному преобразованию, т.е. $\Psi^T \cdot \Psi = 1$, для $\det(\Sigma)$ получим:

$$\det(\Sigma) = \prod_{i=1}^{L^I} \mu_i. \quad (5)$$

Подставляя выражение (5) в (3), получим:

$$H(\vec{X}) = \frac{1}{2} \sum_{i=1}^{L^I} [1 + \log(2\pi) + \log \mu_i]. \quad (6)$$

В работе [5] показано, что максимальная погрешность разложения (4) определяется выражением:

$$\varepsilon^2 = E\{\|\Delta X(L)/X\|^2\} \leq \mu_L. \quad (7)$$

В (7) предполагается, что все собственные значения и их собственные вектора упорядочены в порядке убывания, так как значимость каждого компонента вектора признаков $\vec{X}$ определяется его собственным значением:

$$\mu_1 > \mu_2 > \dots > \mu_L. \quad (7a)$$

Сформулируем **второй принцип**: собственные значения разложения Карунена-Лоева (7a), упорядоченные в порядке убывания, определяют информационную значимость соответствующей компоненты вектора признаков $\vec{X}$.

Рассмотрим ортогональное преобразование исходного вектора признаков $\vec{X}$:

$$\vec{Y} = \Psi^T \cdot \vec{X}, \quad \Psi \cdot \Psi^T \cdot \vec{X} = \Psi \cdot \vec{Y}, \quad \vec{X} = \Psi \cdot \vec{Y}, \quad (8)$$

т.е. матрица Ψ состоит из L^I линейно независимых векторов-столбцов ($|\Psi| \neq 0$) и эти столбцы ортонормированны:

$$\Psi_i^T \Psi_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}$$

Тогда для матрицы ковариации Σ_X получим:

$$\begin{aligned} \Sigma_X &\equiv E[(\vec{X} - E\{\vec{X}\}) \cdot (\vec{X} - E\{\vec{X}\})^T] = E[(\Psi \cdot \vec{Y} - E\{\Psi \cdot \vec{Y}\}) \cdot (\Psi \cdot \vec{Y} - E\{\Psi \cdot \vec{Y}\})^T] = \\ &= \Psi \cdot E[(\vec{Y} - E\{\vec{Y}\}) \cdot (\vec{Y} - E\{\vec{Y}\})^T] \cdot \Psi^T = \Psi \cdot \Sigma_Y \cdot \Psi^T. \end{aligned} \quad (9)$$

На основании выражения (9) и с учетом (4) получим:

$$\det(\Sigma_X) = \det(\Psi \Psi^T) \cdot \det(\Sigma_Y) = \det(\Sigma_Y). \quad (10)$$

Можно сделать вывод о том, что ортогональное преобразование пространства признаков не меняет энтропии (3) для вектора $\vec{Y}$. Из свойств

ортогонального преобразования вытекает, что оно не изменяет расстояний в пространстве признаков и эквивалентно повороту относительно центра рассеивания. Рассеивание, в нашем случае, описывается матрицей ковариации вектора признаков Σ_X . Для многомерного гауссового распределения $f_X(\vec{X})$ из (2a) видно, что поверхность постоянной плотности вероятности вектора признаков $\vec{X}$ является поверхностью второго порядка. Учитывая симметричный характер гауссового распределения относительно вектора математического ожидания $\vec{M}_X = E\{\vec{X}\} = (m_{x1}, m_{x2}, \dots, m_{xL})^T$, установлено, что эта поверхность - L-мерный эллипсоид рассеивания [7]. Поскольку в результате выполнения ортогонального преобразования матрица ковариации вектора признаков Σ_Y приведена к диагональному виду (4), то это означает, что полуоси эллипсоида указывают направления максимального рассеивания в L-мерном пространстве признаков. Смысл ортогонального преобразования $\vec{Y} = \Psi^T \cdot \vec{X}$ заключается в повороте системы координат относительно центра рассеивания $\vec{M}_X$ до совмещения координатных осей $\vec{OY}$ с главными полуосями рассеивания в L-мерном пространстве признаков. Так как каждый ω_s -й класс образов имеет свой эллипсоид рассеивания, то задача выбора набора признаков, обеспечивающих решение задачи распознавания с максимальным уровнем вероятности распознавания $P_{\text{расп}} \omega_i$ сводится к рассмотрению геометрической задаче с K эллипсоидами постоянной плотности вероятности.

Геометрический подход к задаче разделимости классов

Для определенности будем рассматривать эллипсоиды постоянной плотности вероятности с полуосями равными $3\sigma_{yi}^s = 3/\sqrt{\lambda_i}$, где λ_i - собственные значения выборочной ковариационной матрицы (4). Заметим, что эту цифру (3) далее в статье мы уточним! Тогда при условии независимости

компонент вектора признаков $\vec{Y}$ (соответствует диагональному виду Σ_Y), на основании правила «3-х СКО» для гауссового закона распределения 99,7% всех реализаций вектора признаков для соответствующего класса образов будет находиться внутри эллипсоида рассеивания. При этом правильность выбора вектора признаков $\vec{X}$ будет интерпретироваться как отсутствие пересечений эллипсоидов рассеивания различных классов.

На основании приведенных выше рассуждений сформулируем **третий принцип**: описание каждого класса в L -мерном пространстве признаков определяется его эллипсоидом рассеивания с центром $\vec{M}_X$ и полуосями равными $3/\sqrt{\lambda_i}$, где λ_i - собственные значения выборочной ковариационной матрицы вектора признаков по обучающей выборке класса.

В работе [8] приведена теорема.

Теорема 1. Пусть $\vec{X}^T \mathbf{A}_1 \vec{X} = 1$ и $\vec{X}^T \mathbf{A}_2 \vec{X} = 1$ – квадрики в R^L , причем первая является эллипсоидом. Квадрики не пересекаются тогда и только тогда, когда матрица $\mathbf{A}_1 - \mathbf{A}_2$ является знакоопределенной.

Данная теорема может быть использована для матриц $\mathbf{A}_1 = \Sigma_1^{-1}$ и $\mathbf{A}_2 = \Sigma_2^{-1}$. Применяя критерий Сильвестра[9] для матрицы $\Delta_{12} \equiv \Sigma_1^{-1} - \Sigma_2^{-1}$, необходимо проверить, что все главные миноры матрицы Δ_{12} положительны, либо все главные миноры отрицательны.

Непересечение эллипсоидов рассеивания различных классов является необходимым условием при выборе эффективного набора признаков классификатора образов. В случае зашумления вектора признаков среднеквадратические отклонения ошибок (СКО) компонент вектора признаков, а, следовательно, и главные оси эллипсоидов рассеивания будут расти, что, в конце концов, приведет к их пересечению и возникновению состояния неопределенности. Для контроля выполнения необходимого критерия выбора признаков и оценки устойчивости классификатора к шумам

можно использовать расстояние между близкими классами (эллипсоидами рассеивания). Для этого воспользуемся следующей теоремой [8].

Теорема 2. Если выполняется условие теоремы 1, то квадрат расстояния между эллипсоидом $\vec{X}^T \mathbf{A}_1 \vec{X} = 1$ и квадрикой $\vec{X}^T \mathbf{A}_2 \vec{X} = 1$ совпадает с минимальным положительным корнем полинома

$$F(z) = D_\rho(\det[\rho \mathbf{A}_1 + (z - \rho) \mathbf{A}_2 - \rho(z - \rho) \mathbf{A}_1 \mathbf{A}_2]),$$

в предположении, что этот корень не является кратным. Здесь D_ρ — дискриминант полинома, рассматриваемого относительно переменной ρ .

Таким образом, для каждой пары близких эллипсоидов рассеивания с матрицами $\mathbf{A}_1 = \Sigma_1^{-1}$ и $\mathbf{A}_2 = \Sigma_2^{-1}$ необходимо составить матрицу размерности $L \times L$: $\mathbf{D}(z, \rho) = \rho \mathbf{A}_1 + (z - \rho) \mathbf{A}_2 - \rho(z - \rho) \mathbf{A}_1 \mathbf{A}_2$. Далее, разложив детерминант $\det[\mathbf{D}(z, \rho)]$ в виде характеристического многочлена по степеням ρ , получим полином порядка $n = L + 2$: $d(z) \equiv \det[\mathbf{D}(z, \rho)] = d_0 + d_1 \rho^1 + \dots + d_n \rho^n$, где $d_i = d_i(z)$. Отметим, что дискриминант полинома равен нулю тогда и только тогда, когда полином имеет кратные корни. Поскольку машинное определение выборочной оценки ковариационных матриц выполняется с точностью до погрешностей округления, то событие получения кратных корней практически является маловероятным. Выражение для дискриминанта имеет вид:

$$F(z) = ((-1)^{n(n-1)/2} / d_n) \cdot R(d, d'_\rho),$$

где $R(d, d'_\rho)$ — результат полинома $d(z)$ и его производной $d'_\rho(z)$

$$d'_\rho(z) = d_1 + 2d_2 \rho^1 + \dots + n d_n \rho^{n-1}.$$

Результант $R(d, d'_p)$ равен определителю следующей $(2n-1) \cdot (2n-1)$ матрицы:

d_n	d_{n-1}	d_{n-2}	.	.	.	d_0	0	.	.	0
0	d_n	d_{n-1}	d_{n-2}	.	.	.	d_0	0	.	0
.	.	.	.	.	.	.	.	.	.	.
0	0	0	0	0	d_n	d_{n-1}	d_{n-2}	.	.	d_0
nd_n	$(n-1)d_{n-1}$	$(n-2)d_{n-2}$	.	.	d_1	0	0	.	.	0
0	nd_n	$(n-1)d_{n-1}$	$(n-2)d_{n-2}$	.	.	d_1	0	0	.	0
0	0	nd_n	$(n-1)d_{n-1}$	$(n-2)d_{n-2}$	.	.	d_1	0	.	0
.	.	.	.	.	.	.	.	.	.	.
0	0	0	0	0	0	nd_{n-1}	$(n-1)d_{n-2}$	$(n-2)d_{n-2}$	.	d_1

Тогда, выбирая минимальный положительный корень - $\tilde{z}$ полинома $d(z)$, получим евклидово расстояние между двумя эллипсоидами в L -мерном пространстве признаков:

$$D_{12} = \sqrt{\tilde{z}}.$$

Отметим, что указанная процедура оценки расстояний может быть рассчитана на этапе проектирования классификатора. Выполняя пропорциональное расширение (k -коэффициент пропорциональности) главных полуосей эллипсоида рассеивания

$$\sigma_{yi}^* \rightarrow k \cdot \sigma_{yi}, \quad k > 1$$

получим таблицу соответствий $k \rightarrow k\sigma_{yi} \rightarrow D_{js}$ расстояний между классами j и s . Максимальное значение k при котором $D_{js} > 0$, определит запас устойчивости выбранного набора L признаков к воздействию шума.

На основании утверждений теоремы 1 и 2 сформулируем **четвертый принцип**: расстояния D_{js} между классами j и s для обучающих выборок этих классов определяет запас устойчивости $\gamma = D_{js} / \sqrt{\sigma_{yj}^2 + \sigma_{ys}^2}$ выбранной пары классов к воздействию шума.

При некотором упрощении задачи может быть получена оценка вероятности распознавания. Выполняя процедуру нормализации и

ранжирования (7а) собственных значений для выборочной ковариационной матрицы каждого класса, можно оставить только три значимых компоненты вектора признака:

$$\mu_1 > \mu_2 > \mu_3.$$

В работе [7] найдено аналитическое выражение для вероятности попадания случайного трехмерного вектора признаков с гауссовой плотностью вероятности в эллипсоид рассеивания:

$$\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2} = k^2.$$

Здесь, как и при составлении таблицы расстояний между классами, полуоси эллипсоида (a, b, c) пропорциональны СКО компонент вектора признаков:

$$a = k \cdot \sigma_x, \quad b = k \cdot \sigma_y, \quad c = k \cdot \sigma_z. \quad (11)$$

В соответствии с выражением (2а) при $L=3$, выполняя замену

$$\vec{X}^T \Sigma^{-1} \vec{X} = \frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2},$$

получим выражение для вероятности попадания вектора признаков $\vec{X}=(x,y,z)^T$ в эллипсоид B_k :

$$P(\vec{X} \in B_k) = \frac{1}{(2\pi)^{3/2} \sigma_x \sigma_y \sigma_z} \iiint_{B_k} e^{-\frac{1}{2}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} + \frac{z^2}{\sigma_z^2})} dx dy dz.$$

Переходя к сферическим координатам и интегрируя по частям, получим:

$$P(\vec{X} \in B_k) = 2 \cdot \Phi(k) - 1 - \sqrt{\frac{2}{\pi}} \cdot k \cdot e^{-\frac{k^2}{2}}, \quad (12)$$

где $\Phi(k) = \frac{1}{\sqrt{2\pi}} \cdot \int_0^k e^{-\frac{t^2}{2}} dt$ - функция Лапласа[7].

В [5] показано, что уровень значимости признаков, как правило, резко падает с увеличением ранжированного номера собственного значения выборочной матрицы ковариации и на основании (7) относительная ошибка представления вектора признаков для размерности $L>3$ составит единицы процента. Реализация вектора признаков $\vec{X}=(x,y,z)^T$ позволяет выразить

отклонение от центра эллипсоида рассеивания $\vec{M}_X = (m_x, m_y, m_z)^T$ через расстояние, выраженное в среднеквадратических отклонениях, аналогично (11):

$$x = k \cdot \sigma_x, \quad y = k \cdot \sigma_y, \quad z = k \cdot \sigma_z.$$

Значения СКО здесь получены на этапе первичного обучения, поэтому для вектора признака $\vec{X}=(x,y,z)^T$ могут быть найдены такие значения $k = k_j$, которые для заданного класса ω_j (с вектором $\vec{M}_{X_j}$ матрицей Σ_j) определяют расстояние до центра эллипсоида, выраженное в СКО. Подставляя k_j в (12), получим вероятность попадания в эллипсоид $P(\vec{X} \in B_j)$. На рисунке 2 представлены 3 зависимости вероятности попадания $P(\vec{X} \in B_j)$, при этом B_j представляет собой двухмерный эллипс, трехмерный эллипсоид и трехмерный куб. Для трехмерного эллипсоида при $P(\vec{X} \in B_j) = 0.997$ в соответствии с (12) вычислено $k=3.9$, что превышает оценку для трехмерного куба $k=3.4$ (обобщенное «правило трех сигм» на пространство R^3). Это факт свидетельствует о том, что не правомерно использовать аппроксимацию многомерного эллипсоида соответствующей размерности параллелепипедом, так как в этом случае получается завышенная оценка вероятности, а доверительные интервалы имеют заниженную оценку. Для эллипсоида рассеивания доверительный интервал составил $\pm 3,9$ СКО, что на 15 % больше, чем у куба.

Рисунок 2 – Зависимость $P(\vec{X} \subset B_j)$ от СКО, где B_j - эллипсоид, куб и эллипс

Сформулируем **пятый принцип** селекции информационных признаков образов: эллипсоиды рассеивания для всех K классов строятся на основе обучающих выборок этих классов, полуоси эллипсоидов определяются в соответствии с выражением (11), где $k=3,9$. При этом корректный набор признаков должен обеспечить высокий запас устойчивости:

$$\gamma = \min_{\Omega} \gamma_{js} \gg 1.$$

Оценка вероятности распознавания образа каждого класса

Найдем теперь вероятность того, что реализация вектора признаков $\vec{X}$ принадлежит классу ω_j из множества классов $\Omega = \{\omega_j, j=1 \dots K\}$. Эта вероятность является апостериорной, поскольку вычисляется на основе

реализации вектора признаков $\vec{X} = (x_1, x_2, \dots, x_L)^T$. Тогда в соответствии с формулой Байеса [5] получим:

Вероятность того, что реализация вектора признаков $\vec{X}$ принадлежит классу ω_j , определяется выражением:

$$\tilde{P}(\vec{X} \in B_j | \omega_j) = \frac{P(\vec{X} \in \omega_j) \cdot P(\vec{X} \in B_j | \omega_j)}{\sum_{s=1}^K (P(\vec{X} \in \omega_s) \cdot P(\vec{X} \in B_s | \omega_s))}, \quad (13)$$

где $P(\vec{X} \in \omega_j)$ - априорная вероятность принадлежности вектора $\vec{X}$ классу ω_j , $P(\vec{X} \in B_j | \omega_j)$ - условная вероятность события $\vec{X} \in B_j$ при условии, что произошло событие $\vec{X} \in \omega_j$. Если классификатор распознает 100 классов ($K=100$) и все события $\vec{X} \in \omega_j$ равновероятны, то $P(\vec{X} \in \omega_j) = 0,01$ и (13) упрощается:

$$\tilde{P}(\vec{X} \in B_j | \omega_j) = \frac{P(\vec{X} \in B_j | \omega_j)}{\sum_{s=1}^K P(\vec{X} \in B_s | \omega_j)}. \quad (14)$$

Из постановки задачи следует, что совместная вероятность независимых событий определяется произведением их вероятностей. Тогда вероятность распознавания класса ω_j определяется принадлежностью $\vec{X}$ к эллипсоиду B_j (с вероятностью $\tilde{P}(\vec{X} \in B_j | \omega_j)$) и непринадлежностью к другим эллипсоидам с вероятностью $(1 - \tilde{P}(\vec{X} \in B_j | \omega_j))$:

$$P_{\text{расп}\omega_j} = \tilde{P}(\vec{X} \in B_j | \omega_j) \cdot \prod_{\omega_s, s=1 \dots K, s \neq j} (1 - \tilde{P}(\vec{X} \in B_s | \omega_s)). \quad (15)$$

Подчеркнем тот факт, что выражение (15) позволяет оценивать максимальную вероятность распознавания объекта на множестве классов Ω , так как на этапе разработки (при обучении) системы учитываются СКО, обусловленные лишь ошибками квантования (оцифровки) сигнала изображения. В действительности СКО компонент вектора признаков всегда будет больше на реальном сигнале.

Таким образом, получена оценка вероятности распознавания для каждого класса ω_j . Найдем вероятность распознавания образа $P_{\text{расп } X}$ для заданного набора признаков $\vec{X}$, как минимальное значение вероятности $P_{\text{расп } \omega_j}$ для всех K классов:

$$P_{\text{расп } X} = \min_{\omega_j, j=1 \dots K} \{P_{\text{расп } \omega_j}\}. \quad (16)$$

В результате сформулированы ряд принципов и определен критерий выбора набора признаков по максимуму вероятности распознавания образов $P_{\text{расп } X}$, обеспечивающих решение задачи распознавания с определенным уровнем вероятности распознавания (16). Это значение $P_{\text{расп } X}$ соответствует минимальному искажению входного сигнала, так как учитывается только дискретная форма представления сигнала $\vec{X}$. При увеличении отношения сигнал/шум эллипсоид рассеивания расширяется (рост СКО). Последнее обстоятельство в соответствии с (12) приводит к увеличению k и снижению вероятностей $P_{\text{расп } \omega_j}$. В случае аномального роста погрешностей признаков эллипсоиды рассеивания разных классов пересекаются и это еще более резко снижает $P_{\text{расп } \omega_j}$.

Существенная размерность пространства признаков классов

Сформулируем еще один принцип по выбору размерности набора признаков L_{min} , обеспечивающих решение задачи распознавания с соответствующим ему максимальным уровнем вероятности распознавания.

С одной стороны в соответствии с методом величина $L_{\text{min}} = L^I$ определяется максимальным рангом матриц ковариаций $\Sigma_j, j = 1 \dots K$ то есть

$$L^I = \max_{j=1 \dots K} \text{rank}(\Sigma_j). \quad (17)$$

Заметим, что выражение (17) скорее является верхней границей для выбора размерности пространства признаков. Как правило, назначая уровень значимости признаков, удастся существенно понизить размерность этого пространства, например, используя упорядоченный набор собственных значений разложения Карунена-Лоева для ковариационной матрицы класса. Именно в этой связи был введен термин “существенная размерность” и на основании (15), (16) получим значение L^{II} .

С другой стороны, минимальное возможное число признаков описывается степенями свободы физической системы. Например, движение твердого тела в пространстве описывается 6 фазовыми координатами (3- линейных , 3-угловых). Считая, что вдоль каждой фазовой координаты необходимо иметь хотя бы 1 координату вектора признаков, получим 6 компонент. Если движения объекта можно считать плоским, то получим 5 компонент (одно уравнение связи координат движения объекта, например $y=0$). Обобщая, приходим к выводу: если на движение объекта наложено m уравнений связи, то получим $6-m$ компонент вектора признаков. Тогда получим $L^0 = 6 - m$ степеней свободы системы, значение L^0 определяет “истинную размерность” системы.

Тогда для размерности пространства признаков можно записать:

$$L^0 \leq L_{min} \leq \{L^I, L^{II}\}. \quad (18)$$

В этой связи может быть сформулирован **принцип 6-существенной размерности пространства признаков**: размерностью пространства признаков классификатора определяется выражением (18).

В заключении сформулируем выводы по работе.

1 Рассмотренные принципы построения классификаторов основываются на самых общих положениях статистической теории оценивания и теории информации и не используют конкретных методов и схем построения классификаторов.

2 Изложенные в статье принципы могут лечь в основу научно-обоснованного подхода к методике выбора наиболее информативных признаков классификаторов.

*Работа поддержана Российским фондом фундаментальных исследований.
Проект № 11-08-00292.*

Список литературы

1. Hartley R.V.L. Transmission of Information // Bell System Technical Journal. July 1928. Vol. 7, no. 3. P. 535-563.
2. Шеннон К. Работы по теории информации и кибернетике: пер. с англ. М.: Издательство иностранной литературы, 1963. 832 с.
3. Shore J.E., Johnson R.W. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy // IEEE Transactions on Information Theory. 1980. Vol. IT-26. P. 26-37.
4. Jaynes E.T. On the rationale of maximum-entropy methods // Proceedings of the IEEE. 1982. Vol. 70. P. 939-952.
5. Фукунага К. Введение в статистическую теорию распознавания образов : пер. с англ. М.: Наука, 1979. 366 с.
6. Хайкин С. Нейронные сети: полный курс: пер. с англ. 2-е изд. М.: Издательский дом "Вильямс", 2006. 1104 с.
7. Вентцель Е.С. Теория вероятностей : учебн. для вузов. 5-е изд., стер. М.: Высшая школа, 1998. 576 с.
8. Uteshev A.Yu., Yashina M.V. Distance computation from an ellipsoid to a linear or a Quadric Surface in R^n // Computer Algebra in Scientific Computing, Springer Berlin Heidelberg, 2007, pp. 392-401. (Lecture Notes in Computer Science, vol. 4770). DOI: 10.1007/978-3-540-75187-8_31
9. Гантмахер Ф.Р. Теория матриц. М.: Наука, 1966. 576 с.

Description of complex images of ground objects in the problem of pattern recognition

11, November 2012

DOI: 10.7463/1112.0500490

Gulevitch S., P., Veselov Yu., G., Pryadkin S.P.

Russia, Bauman Moscow State Technical University

Russia, JSC «Kamov»

Russia, VUNTS Air Force «Air Force Academy named after Professor NE Zhukovsky

and Y. Gagarin»

vesel_foto@mail.ru

The article describes principles of selecting informative features for classification of optical images of complex surface facilities. The authors justify the selection of classes of objects using ellipsoidal estimation. They developed criteria for qualitative selection of the classifier attribute vector answering questions of resistance of classes to noise impact, of compact description of classes using ellipsoids of dispersion, of an essential dimension of the feature vector. The authors propose a science-based approach to selection of informative features of the signal classifier based on the concept of entropy of information theory and ellipsoidal estimation.

Publications with keywords: [pattern recognition](#), [selection of characteristics](#), [systems of machine vision](#), [information theory](#)

Publications with words: [pattern recognition](#), [selection of characteristics](#), [systems of machine vision](#), [information theory](#)

References

1. Hartley R.V.L. Transmission of Information. *Bell System Technical Journal*, July 1928, vol. 7, no. 3, pp. 535-563.
2. Shannon K. *Raboty po teorii informatsii i kibernetike*. [Work on information theory and cybernetics]. Collection of articles. Transl. from English. Moscow, Izdatel'stvo inostrannoi literatury, 1963. 832 p.
3. Shore J.E., Johnson R.W. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 1980, vol. IT-26, pp. 26-37.

4. Jaynes E.T. On the rationale of maximum-entropy methods. *Proceedings of the IEEE*, 1982, vol. 70, pp. 939-952.
5. Fukunaga K. *Introduction to Statistical Pattern Recognition*. Academic Press, New York, 1972. (Russ. ed.: Fukunaga K. *Vvedenie v statisticheskuiu teoriiu raspoznavaniia obrazov*. Moscow, Nauka, 1979. 366 p).
6. Haykin S. *Neural Networks: A Comprehensive Foundation*. 2nd edition. Prentice-Hall, 1998. (Russ. ed.: Khaikin S. *Neironnye seti : Polnyi kurs*. 2th ed. Moscow, Izdatel'skii dom "Vil'iamc", 2008. 1104 p.).
7. Venttsel' E.S. *Teoriia veroiatnostei* [Probability theory]. Moscow, Vysshaia shkola, 1998. 576 p.
8. Uteshev A.Yu., Yashina M.V. Distance computation from an ellipsoid to a linear or a Quadric Surface in R^n . *Computer Algebra in Scientific Computing*, Springer Berlin Heidelberg, 2007, pp. 392-401. (*Lecture Notes in Computer Science*, vol. 4770). DOI: 10.1007/978-3-540-75187-8_31
9. Gantmakher F.R. *Teoriia matrits* [Theory of matrices]. Moscow, Nauka, 1966. 576 p.