

М. Н. Фаворская

РАСПОЗНАВАНИЕ ДИНАМИЧЕСКИХ ОБРАЗОВ НА ОСНОВЕ ПРЕДСКАЗЫВАЮЩИХ ФИЛЬТРОВ

Построена модель распознавания динамических образов из видеопоследовательностей на основе допустимых отображений в пространстве изображений и пространстве признаков. Введено понятие максимального динамического инварианта. Рассмотрен метод определения максимальных динамических инвариантов на основе фильтров, предсказывающих поведение траектории в многомерном пространстве признаков.

Ключевые слова: распознавание, динамические образы, инвариантные функции.

Постановка задачи распознавания видеопоследовательностей несколько отличается от распознавания изображений тем, что видеопоследовательность представляет собой некоторую совокупность последовательных во времени изображений (кадров) с возможно изменяющимися сценами наблюдения, иными словами само понятие «изображение» теряет статичность, обладает предысторией и становится возможным прогнозирование поведения интересующих нас объектов. В этом случае производится анализ некоторого множества идеальных последовательных изображений FV , на котором задано разбиение на подмножества $V^{fv_1}, V^{fv_2}, \dots, V^{fv_m}$, называемые классами разбиения (см. таблицу). При этом оператор $\phi_i^b: P \rightarrow O$, представляющий собой отображение множества объектов наблюдения P в различных проекциях на множество нормализованных объектов O приобретает качественно иную структуру и становится оператором $\phi_i^b: PV \rightarrow OV$, отражающим динамику поведения объекта наблюдения. Оператор ϕ_i^b играет существенную роль и при отображении ϕ_i^b множества нормализованных объектов OV на множество эталонов ZV , так как помимо традиционных признаков статических изображений используется информация о признаках движения, что повышает точность распознавания объектов и характеризует их поведенческие аспекты. Особенностью распознавания видеопоследовательностей является то, что они могут состоять как из плавно изменяющихся сцен (задачи сопровождения объектов), так и из множества не связанных между собой сцен (анализ видеофильмов). При анализе видеофильмов требуется возобновлять процедуру распознавания объектов в каждой сцене. Сложность отображения ϕ_i^b также связана с понятием многоуровневого движения в кадре. Задача автоматического разбиения сцены последовательных кадров на уровни движения является очень полезной с точки зрения практичес-

кого применения, но весьма далекой от решения на уровне модели.

Изображения объектов в видеопоследовательности представляют собой либо результат движения, либо процесс их развития. При этом объекты переходят из одного состояния в другое, последовательность или совокупность которых определяет принадлежность объекта тому или иному образу. Каждое состояние объекта определяется набором признаков (в том числе мгновенными значениями скоростей и ускорений и их направлениями), а вектор, соответствующий описанию объекта, может изменяться при различных ракурсах съемки и одновременно формировать определенную траекторию в пространстве признаков. Таким образом, задачу распознавания динамических образов можно определить как классификацию совокупностей состояний, как классификацию траекторий, так и то, и другое одновременно. При смене ракурсов изменяется описание объекта, но всегда существует некоторая закономерность смены его состояний, по которой и определяется образ, которому принадлежит наблюдаемый объект. Иногда для распознавания достаточно определить некую совокупность состояний, которую принимал объект относительно наблюдателя, иногда этих сведений оказывается недостаточно и приходится учитывать не только совокупность наблюдаемых состояний, но и динамику их проявления (предыстория поведения). Для случая учета предыстории поведения объекта окончательное решение о принадлежности объекта тому или иному образу формируется в результате продолжительного наблюдения за поведением объекта. Такую процедуру в некоторых случаях можно рассматривать как известную последовательную процедуру принятия решений [1].

Функции $\phi_i^b: OV \rightarrow ZV$, $\phi_i^b: PV \rightarrow OV$, $\phi_i^b: FV \rightarrow PV$, $\psi_i^b: RV \rightarrow FV$ характеризуют процессы порождения мно-

Допустимые отображения при распознавании объектов в видеопоследовательностях

Отображения в пространстве изображений	Отображения в пространстве признаков	Краткая характеристика
$\psi_i^b: RV \rightarrow FV$	$\psi_i^b: X^{RV} \rightarrow X^{FV}$	Низший уровень обработки. Пространственная и временная фильтрация видеопоследовательностей, выявление зон интереса
$\phi_i^b: FV \rightarrow PV$	$\phi_i^b: X^{FV} \rightarrow X^{PV}$	Средний уровень – сегментация. Вычисление глобальных и локальных признаков изображений проекций объектов с учетом предыстории
$\phi_i^b: PV \rightarrow OV$	$\phi_i^b: X^{PV} \rightarrow X^{OV}$	Средний уровень – сегментация. Вычисление глобальных и локальных признаков изображений объектов с учетом предыстории
$\phi_i^b: OV \rightarrow ZV$	$\phi_i^b: X^{OV} \rightarrow X^{ZV}$	Высший уровень – распознавание и понимание. Анализ глобальных и локальных признаков изображений объектов с учетом предыстории

жества изображений нормализованных объектов \mathbf{OV} из множества эталонов \mathbf{ZV} , множества объектов наблюдения \mathbf{PV} в произвольных проекциях из множества нормализованных объектов \mathbf{OV} , множества идеальных изображений \mathbf{FV} из множества изображений объектов наблюдения \mathbf{PV} и реальных изображений \mathbf{RV} из идеальных изображений \mathbf{FV} . Особенность рассматриваемого класса задач состоит в том, что в них отображение fv_i^t и соответствующее ему обратное отображение fv_i^b зависят не только от мешающего параметра β , но и от времени t . Обозначим функцию отображения fv_i^t , изменяющуюся во времени как $(fv_i^t)_{\beta,t}$, а изображение

$$RV_j = \psi v_j^t(FV_j, h), FV_j = (fv_i^t)_{\beta,t}(\phi v_i^t((\phi v_i^t)_\alpha(Z_j)))$$

называют динамическим изображением (видеопоследовательностью) объекта наблюдения, принадлежащего образу V_j . При этом $h \in \mathbf{H}$ – шум, возникающий в процессе отображения (\mathbf{H} – множество возможных значений h). Можно рассмотреть упрощенную модель при $\alpha = \text{const}$, однако, следует помнить, что преобразование ϕ становится намного сложнее из-за концепции многоуровневого движения в видеопоследовательности. Тогда

$$RV_j = \psi v_j^t(FV_j, h), FV_j = (fv_i^t)_{\beta,t}(\phi v_i^t(Z_j)). \quad (1)$$

Такая модель, отображающая динамические характеристики видео объектов, существенно отличается от модели статического изображения [2]. Предполагалось, что множество переходов $\beta_k \rightarrow \beta_v$ порождает множество \mathbf{B}^c , которое в свою очередь порождает множество преобразований $g \in \mathbf{G}$ пространства самого на себя, т. е. для любого перехода $\beta_k \rightarrow \beta_v$ найдется такое $g \in \mathbf{G}$, при котором выполняется равенство

$$gFV_k = g\phi v_i^t((fv_i^t)_{\beta_k}(PV_j)) = \phi v_i^t((fv_i^t)_{\beta_v}(PV_j)) = FV_v, \quad (2)$$

где $FV_k, FV_v \in \mathbf{FV}$.

Обобщенная модель (1) отличается тем, что любые переходы $(\beta_k \rightarrow \beta_v) \in \mathbf{B}^c$ обладают свойством (2), но переходы $t_n \rightarrow t_{n+1}$ этим свойством не обладают, т. е. между множествами переходов $(t_n \rightarrow t_{n+1}) \in \mathbf{t}^c$ и множеством порождаемых ими преобразований на множестве \mathbf{FV} нельзя установить соответствие вида (2). Именно эта особенность видеопоследовательностей не позволяет применять обычные методы распознавания для их классификации, хотя постановка задачи формально не изменяет своего содержания.

Если при распознавании статических объектов модель порождения изображений можно было рассматривать как совокупность траекторий, состоящих из всех точек множества \mathbf{X}^{FV} , окруженных облаком, расположение которого определяется распределением шума h , а сами эти точки (включая и точки $X^{FV}_k \in \mathbf{X}^{FV}$) образуют множество \mathbf{X}^{RV} , то при распознавании движущихся объектов эта модель существенно усложняется. Если при распознавании одного объекта модель динамического изображения представляет собой некую гиперплоскость, соответствующую одному образу и окруженную облаком, расположение которого зависит от характера шума, то при распознавании нескольких образов существует несколько таких гиперплоскостей, взаимное расположение которых практически не позволяет провести какие-либо границы между ними. Такие гиперплоскости строятся при значениях

$\beta = \text{var}, t = \text{var}$. Даже если существенно упростить задачу и считать, что $\beta = \text{const}$, то решение нельзя найти классическими методами теории распознавания образов. Дело в том, что переходы $t_n \rightarrow t_{n+1}$ порождают такие преобразования изображений \mathbf{FV} , относительно которых невозможно восстановить максимальные инварианты известными способами.

Введем понятие динамического максимального инварианта. Пусть $\{\mathbf{RV}\}$ – последовательность кадров $\{RV_t, RV_{t+1}, \dots, RV_{t+N}\}$, где t – начальный, $t+1, t+2, \dots, t+N$ – последующие моменты наблюдения, причем t может быть выбрано в любой точке гиперповерхности. При этом любые две точки $X^{FV}_{k,t}$ и $X^{FV}_{v,t'}$ соответствующие одному и тому же образу, должны лежать на гиперповерхности, соответствующей заданному значению t . Тогда динамической инвариантной функцией будем называть любую функцию от $\{\mathbf{RV}\}$, если она не изменяется при изменении параметров β и t . Предположим, что для динамической инвариантной функции E при достаточно длинных последовательностях $\{\mathbf{RV}\}$ и $\{\mathbf{RV}\}^c$ равенство

$$E\{\mathbf{RV}\} = E\{\mathbf{RV}\}^c \quad (3)$$

выполняется только тогда, когда для каждой пары из серии наблюдений $\{\mathbf{RV}\}$ и $\{\mathbf{RV}\}^c$ существует такое преобразование $g \in \mathbf{G}$ и всегда найдутся такие значения t и t' , что

$$\begin{cases} RV_{k,t} = gRV_{v,t'} \\ RV_{k,t+1} = gRV_{v,t'+1} \\ \dots \\ RV_{k,t+N} = gRV_{v,t'+N} \end{cases} \quad (4)$$

Такую функцию назовем максимальным динамическим инвариантом, причем $g \in \mathbf{G}$ – преобразования, порожденные всевозможными переходами из \mathbf{B}^c . Условия (3) и (4) выполняются только тогда, когда либо все наблюдения $\{\mathbf{RV}\}$ и $\{\mathbf{RV}\}^c$ принадлежат гиперплоскости при отсутствии шума, либо все они сосредоточены в окрестности гиперплоскости, определяемой распределением шума h , и всегда существует возможность сдвинуть наблюдения во времени так, чтобы выполнялись равенства (4). Если помеха h постоянна на всем интервале наблюдения, то условия (4) будут выполняться на некоторой равноудаленной гиперплоскости. Отсюда следует, что некоторая функция от совокупности наблюдений обладает свойством максимального динамического инварианта только тогда, когда она постоянна на любой траектории или является совокупностью точек, принадлежащих гиперповерхности или лежащих в окрестности этой гиперповерхности.

Если наблюдения $\{\mathbf{RV}\}$ и $\{\mathbf{RV}\}^c$ достаточно продолжительны, то функции $E\{\mathbf{RV}\}$ и $E\{\mathbf{RV}\}^c$, обладающие свойствами максимального динамического инварианта, различаются только в том случае, если наблюдения $\{\mathbf{RV}\}$ и $\{\mathbf{RV}\}^c$ находятся на гиперповерхностях различных образов. Для решения задачи распознавания достаточно указать номер гиперповерхности, из которой произошла наблюдаемая последовательность, а максимальный динамический инвариант постоянен именно на тех точках, которые произошли из некоторой гиперповерхности. Тогда для построения решающей функции, инвариантной к параметрам β и t , достаточно построить такую функцию,

которая зависела бы от наблюдений только через максимальный динамический инвариант. Сформулируем задачу распознавания динамических образов в общем виде, используя свойства максимального динамического инварианта.

Предположим, что любое изображение формально описывается выражением (1). Задана обучающая выборка, состоящая из N^* последовательностей наблюдения. Для каждой последовательности наблюдений $\{\mathbf{RV}\}_k$, $k = 1, 2, \dots, N^*$, указана ее принадлежность определенному образу. Требуется по данным обучающей выборки восстановить любой максимальный инвариант $E_{\beta,t}\{\mathbf{RV}\}$ относительно параметров β и t по всем j и построить решающую функцию $U\{\mathbf{RV}\}$, зависящую от изображений \mathbf{RV} только через найденный максимальный инвариант (j – номер образа). Эта функция должна принимать одинаковые значения на множестве изображений одного образа и различные значения на множествах разных образов.

В общем случае по заданной обучающей выборке требуется определить множество преобразований \mathbf{G} , соответствующее множеству возможных переходов \mathbf{V}^c или множеству возможных значений $\beta \in \mathbf{B}$. Такая задача очень сложна. Ее можно упростить, считая, что заранее задается множество преобразования \mathbf{G} или восстанавливается такая функция $E\{\mathbf{RV}\}$, которая обладает свойствами максимального динамического инварианта только относительно последовательностей, входящих в обучающую выборку. При обучении распознаванию динамических образов задача определения полного множества преобразований существенно усложняется и практически не может быть решена без дополнительных ограничений. Процесс определения преобразований \mathbf{G} , восстановления максимального динамического инварианта и выбора решающей функции, зависящей от наблюдения только через найденный максимальный динамический инвариант, называется процессом обучения. Применение полученной решающей функции для принятия решений о принадлежности образам новых наблюдений называется процессом распознавания.

Рассмотрим модель порождения изображений (1) более подробно. Предположим, что в этой модели отсутствуют помехи, т. е. $h = 0$ и $\beta = \text{const}$. Тогда каждому образу динамического объекта будет соответствовать единственная траектория в некотором пространстве признаков. Траектории, соответствующие различным интервалам наблюдения за объектами одного и того же образа, будут представлять собой отрезки траектории этого образа. Процесс распознавания заключается в определении принадлежности наблюдаемого объекта тому или иному образу, т. е. по серии наблюдений требуется указать номер той траектории, на которой лежат наблюдаемые точки. При этом траектории различных образов могут пересекаться. Однако задача усложняется тем, что проекции движущегося объекта, отображаемые во фронтальной плоскости (плоскости реальных изображений или кадров), могут кардинально изменяться при движении объекта. В этом случае нахождение обобщенного максимального динамического инварианта становится еще более трудной, а в некоторых случаях и невозможной задачей, или

следует принять, что траектория объекта размывается. В последнем варианте можно говорить о принадлежности наблюдаемого объекта некоторому образу с некоторой долей вероятности.

Вернемся к более простому варианту, когда проекции объектов подвержены группе аффинных преобразований. Предположим, что точки траекторий образов установлены в процессе обучения. Однако мгновенного наблюдения (в результате которого будет получена одна точка, принадлежащая той или иной траектории) за движущимся объектом будет недостаточно, потому что траектории образов могут пересекаться, и возникает неоднозначность. Таким образом, для распознавания движущихся объектов требуется учитывать совокупность наблюдений, определяющих поведение объекта в предыдущие моменты времени, т. е. предысторию изображений $\{\mathbf{RV}\}_0^N$ на некотором интервале времени t_0, t_N . При появлении шума в модели (1) и при $\beta = \text{var}$ принцип учета предыстории приобретает еще большее значение, так как при этом для распознавания динамических образов требуется еще большее количество точек наблюдения из предыстории. Таким образом, максимальный динамический инвариант зависит от предыстории, а решающая функция должна зависеть от наблюдений только через максимальный динамический инвариант. По определению, максимальный динамический инвариант обладает особым свойством, заключающимся в том, что для достаточно длинных последовательностей наблюдений $\{\mathbf{RV}\}_0^N$ он принимает одно и то же значение только в том случае, когда эти последовательности принадлежат одной и той же гиперповерхности. Если же сравниваемые последовательности принадлежат различным гиперповерхностям, то их максимальные инварианты не равны между собой. Основная трудность при решении задачи распознавания динамических образов состоит в определении максимальных динамических инвариантов: если они будут найдены, то и задача распознавания динамических образов будет решена.

Рассмотрим один из методов определения максимальных динамических инвариантов. Предположим, что известны аналитические зависимости

$$\mathbf{X}_{t+1}^{RV} = \Phi_j(\mathbf{X}_{t-p}^{RV}, \mathbf{X}_{t-1}^{RV}, \dots, \mathbf{X}_{t-N}^{RV}), j = 1, 2, \dots, m, \quad (5)$$

где $\mathbf{X}_{t+1}^{RV}, \mathbf{X}_{t-p}^{RV}, \mathbf{X}_{t-1}^{RV}, \dots, \mathbf{X}_{t-N}^{RV}$ – точки в пространстве, соответствующие наблюдениям за динамическим объектом в моменты времени $t+1, t, t-1, \dots, t-N$. Функцию Φ_j можно рассматривать как предсказывающий фильтр, который по предыдущим значениям \mathbf{X}^{RV} определяет его будущие значения. Кроме того, функция Φ_j является некоторой моделью, адекватной гиперповерхности j -го образа. Если модель (5) достаточно точная и если последовательность $\{\mathbf{X}^{RV}\}$ лежит в окрестности k -й поверхности, то

$$\sum_{t=0}^{\infty} \left(\mathbf{X}_{t+1}^{RV} - (\mathbf{X}_{t+1}^{RV})^k \right)^2 = \min_j \sum_{t=0}^{\infty} \left(\mathbf{X}_{t+1}^{RV} - (\mathbf{X}_{t+1}^{RV})^j \right)^2,$$

где $(\mathbf{X}_{t+1}^{RV})^k$ – координаты точки \mathbf{X}_{t+1}^{RV} , вычисленные k -м фильтром; $(\mathbf{X}_{t+1}^{RV})^j$ – координаты точки \mathbf{X}_{t+1}^{RV} , вычисленные j -м фильтром; \mathbf{X}_{t+1}^{RV} – действительные координаты точки. Равенство

$$\begin{aligned} & \arg \min_j \sum_{t=0}^{\infty} \left(X_{t+1}^{RV} - (X_{t+1}^{RV})^j \right)^2 = \\ & = \arg \min_j \sum_{t=0}^{\infty} \left((X_{t+1}^{RV})^c - (X_{t+1}^{RV})^j \right)^2 \end{aligned}$$

достигается тогда, когда последовательности $\{\mathbf{RV}\}$ и $\{\mathbf{RV}\}^c$ лежат в окрестности одной и той же поверхности, и только тогда, когда выполняются равенства (4). Значит,

$$\arg \min_j \sum_{t=0}^{\infty} \left(X_{t+1}^{RV} - (X_{t+1}^{RV})^j \right)^2 \quad (6)$$

обладает свойством максимального динамического инварианта. Построив решающее правило, зависящее от наблюдений через выражение (6), можно решить задачу распознавания динамических образов.

С точки зрения математической статистики задача распознавания образов полностью совпадает с задачей различения гипотез. Каждому образу ставится в соответствие вероятностная мера, а параметром, по которому различаются эти меры, является номер образа. На основании наблюдений следует предпочесть одну из гипотез. В классической постановке задачи различения гипотез считаются известными условные распределения вероятностей. В задаче распознавания эти распределения можно определить в процессе обучения. Предположим, что в процессе обучения определены значения вероятностей $P(\{\mathbf{X}^{PV}_v\}/V_j)$ для всех значений v и j ($v = 1, 2, \dots, S$ – номера состояний). Предположим, что существует возможность определить вероятности $P(\{\mathbf{X}^{PV}_v\}/V_j)$ для любой совокупности состояний $\{\mathbf{X}^{PV}\}$ и для всех j . Тогда, на каждом шаге наблюдения можно вычислить отношения вероятностей P_{jN}/P_{kN} , где $P_{kN} = (\{\mathbf{X}^{PV}_1\}, \{\mathbf{X}^{PV}_2\}, \dots, \{\mathbf{X}^{PV}_N\}/V_k)$; $P_{jN} = (\{\mathbf{X}^{PV}_1\}, \{\mathbf{X}^{PV}_2\}, \dots, \{\mathbf{X}^{PV}_N\}/V_j)$; $\{\mathbf{X}^{PV}_1\}$ – состояние, которое принял динамический объект в момент первого наблюдения; $\{\mathbf{X}^{PV}_2\}$ – состояние, которое принял динамический объект в момент второго наблюдения и т. д. Используя отношение P_{jN}/P_{kN} для решения задачи распознавания, можно применить известные последовательные процедуры. Но отношение P_{jN}/P_{kN} и есть тот функционал, который зависит от наблюдений через совокупности тех состояний, который принимает динамический объект в процессе наблюдения. Последовательные процедуры могут дать удовлетворительный результат, поскольку в них не просто используется перечисление состояний, а учитываются вероятности попадания представляющей точки в ту или иную область пространства признаков.

Простое перечисление совокупности пройденных состояний не всегда приводит к положительному результату: для классификации можно использовать более детальные свойства траекторий, а именно порядок смены состояний. Присвоим каждой области пространства определенный вес так, чтобы веса областей различались. Рассмотрим систему координат, в которой по оси абсцисс откладываются значения времени, а по оси ординат – вес той области, в которой находится представляющая точка в данный момент. Тогда в процессе наблюдения за объектом вычерчиваются кривые линии, каждая из которых соответствует определенной траектории в пространстве признаков. Траекториям, прошедшим через одни и те же области, но в различных последовательностях, соответствуют различные кривые в новой плоской системе координат. Таким образом, задачу классификации слож-

ных пространственных траекторий можно свести к более простой задаче классификации плоских кривых.

Используя новую систему координат, задачу классификации пространственных кривых можно решить и другим способом. Построим координатную сетку, где по оси абсцисс отложены определенные отрезки времени, а по оси ординат – веса, присвоенные различным областям. Если считать, что факт попадания представляющей точки в один из прямоугольников этой сетки является одним из множества несовместных событий, то достаточно провести классификацию траекторий как совокупности всех состояний, принадлежащих этой траектории. Такая процедура позволяет выносить решение о принадлежности наблюдаемого изображения объекта тому или иному образу по неполным наблюдениям.

Для обучения системы необходимо оценить вероятности вида $P(\{\mathbf{X}^{PV}_v\}/V_j)$, где $v = 1, 2, \dots, S, j = 1, 2, \dots, m$; $\{\mathbf{X}^{PV}_v\}$ – событие, характеризующееся попаданием представляющей точки в один из прямоугольников координатной сетки; V_j – образы, на которые необходимо разделить все входные изображения. Когда такие вероятности определены, задача классификации траекторий сводится к задаче классификации совокупностей состояний, т. е. такой метод позволяет перейти от учета порядка следования состояний к учету совокупности состояний, что значительно упрощает решение задачи. В данном случае специфическую сложность представляет выбор масштаба времени. Масштаб времени и интервалы на оси абсцисс должны выбираться в соответствии со скоростью изменений местоположения объекта.

Задачу распознавания динамических образов с учетом последовательности состояний можно решить с использованием максимально инвариантных свойств выражения (6). В этом случае задача распознавания динамических образов формулируется следующим образом. Предположим, что любое изображение может быть формально описано выражением (1). Задана обучающая выборка, состоящая из N^* последовательных наблюдений $\{\mathbf{RV}\}_k, k = 1, 2, \dots, N^*$, указана ее принадлежность определенному образу. Требуется для каждого образа построить такой предсказывающий фильтр, который на основании некоторой последовательности предыдущих состояний указывал бы будущее состояние динамического объекта. Решение должно зависеть от сравнения предсказанных и действительных состояний на $(t + 1)$ -м шаге наблюдения. Таким образом, задача распознавания динамических образов сводится к задаче построения фильтра, предсказывающего поведение траектории в многомерном пространстве. Методы, основанные на такой постановке задачи распознавания динамических образов, называются методами обучающихся предсказывающих фильтров.

Задачу распознавания динамических объектов с учетом порядка следования состояний можно свести к задаче классификации многомерных процессов. Общим функциональным выражением предыстории функции времени в ограниченном частотном диапазоне является зависимость

$$\Phi(x(t)) = \sum_{i=0}^N \alpha_i x_{t-i} + \sum_{i_1=0}^N \sum_{i_2=0}^N \alpha_{i_1, i_2} x_{t-i_1} x_{t-i_2} + \dots, \quad (7)$$

где x_{t-i} – значения функции в моменты времени $t, t - 1, t - 2, \dots$; α_i – соответствующие коэффициенты. Выраже-

ние (7), называемое оператором Колмогорова–Габо́ра, используется для прогноза временных последовательностей. Первая сумма в этом операторе представляет собой линейный предсказывающий фильтр. Коэффициент α_i определяет влияние дискретного значения функции x , которое она принимает на i -м интервале из предыстории. Вторая сумма состоит из квадратичных комбинаций дискретных значений x . Коэффициенты α_{i_1, i_2} определяют веса пар значений функции в моменты $t - i_1$ и $t - i_2$ из предыстории.

Для решения полинома (7) целесообразно использовать метод группового учета аргументов, основная особенность которого состоит в том, что нормальные уравнения составляются и решаются не относительно всего полинома (7), а относительно более простых полиномов с небольшим количеством членов, учитывающих только некоторые группы аргументов. При этом каждый упрощенный полином рассматривается как самостоятельный и независимый предсказывающий фильтр, коэффициенты которого определяются путем решения нормальных уравнений с небольшими матрицами коэффициентов. После настройки каждый фильтр (каждый упрощенный полином) проверяется на специальной проверочной последовательности, после чего отбираются лишь несколько наилучших в смысле результата предсказания фильтров и выход каждого из них используется в качестве аргумента для построения более сложного фильтра и т. д. После нескольких отборов образуется гибкая математическая структура, которая, с одной стороны, дает удовлетворительное предсказывание, а с другой стороны – может быть сравнительно легко реализована при сравнительно небольших обучающих последовательностях.

Рассмотрим метод группового учета аргументов на примере попарного учета аргументов. Пусть задана функция $x(t)$ на некотором интервале наблюдения T . Этот интервал разбивается на два интервала T_1 и T_2 . Пусть в качестве обучающей последовательности используется интервал T_1 , а в качестве проверочной последовательности – интервал T_2 . При этом учитывается только N точек из предыстории $x_t, x_{t-1}, \dots, x_{t-N+1}$. Для каждой пары аргументов x_{t-i_1} и x_{t-i_2} составляются упрощенные полиномы первого уровня типа $y_{j_1}(x_{t-i_1}, x_{t-i_2})$, и определяются неизвестные коэффициенты. После этого полиномы считаются обученными. С помощью проверочной последовательности оценивается качество предсказания каждого из обученных фильтров, определяемое, например, по величине среднеквадратичной ошибки. Из этих фильтров выбирается несколько лучших в смысле выбранного критерия, и составляются полиномы второго уровня вида $z_{j_2}(y_{j_1}, y_{j_1k})$. Далее составляются нормальные уравнения

относительно полиномов z_j , в результате решения которых определяются коэффициенты $\alpha_{j_1}^{(2)}$. Обученные фильтры (полиномы) проверяются на проверочной последовательности T_2 , и из них отбираются N^{**} лучших в смысле выбранного критерия качества. После этого значения, получаемые на выходах отобранных фильтров, используются для построения более сложных полиномов следующего ряда j_3 и т. д. Нарастивание сложности полиномов прекращается тогда, когда хотя бы один полином последнего ряда дает удовлетворительные предсказания. Этот полином используется в качестве предсказывающего фильтра.

Такая процедура может быть легко применена при распознавании динамических объектов. Для этого методом группового учета аргументов нужно построить столько предсказывающих фильтров по каждой координате, на сколько образов требуется разделить динамические объекты. В процессе распознавания предпочтение следует отдавать тому образу, для которого

$$\varepsilon_k = \min_j \sum_{i=1}^n \bar{\varepsilon}_{ij}^2,$$

где $\bar{\varepsilon}_{ij}^2$ – среднеквадратичная ошибка по i -й координате фильтра, обученного на объектах j -го образа.

Таким образом, в статье предложена упрощенная модель, отображающая динамические характеристики объектов в видеопоследовательности. Показано, что при распознавании нескольких динамических образов существует несколько гиперплоскостей в многомерном пространстве признаков, взаимное расположение которых не позволяет провести какие-либо границы между ними, что приводит к невозможности нахождения решения классическими методами теории распознавания образов. Сформулирована задача распознавания динамических образов с использованием функции максимального динамического инварианта и рассмотрен более простой вариант, когда проекции объектов подвержены группе аффинных преобразований. Показано, что метод определения максимальных динамических инвариантов на основе предсказывающих фильтров может быть реализован по схеме метода группового учета аргументов.

Библиографический список

1. Васильев, В. И. Распознающие системы : справ. / В. И. Васильев. Киев : Наукова думка, 1983.
2. Фаворская, М. Н. Инвариантные решающие функции в задачах распознавания статистических изображений / М. Н. Фаворская // Вестник СибГАУ. Красноярск, 2007. Вып. 1(14).

M. N. Favorskaya

THE DYNAMIC PATTERNS RECOGNITION BASED ON PREDICATING FILTERS

The model of dynamic patterns recognition from video sequences based on allowable mapping into images space and features space is realized. The term of maximum dynamic patterns recognition function is introduced. The method of maximum dynamic invariant functions based on filters, predicating a trajectory behavior in multivariate feature space is regarded.

Keywords: patterns recognition, dynamic patterns, invariant functions.

