

УДК 519.673

К.С. Смеляков

Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков

МЕТОДОЛОГИЧЕСКИЕ ОСНОВЫ МАШИННОГО САМООБУЧЕНИЯ ДЛЯ СЕГМЕНТАЦИИ ИЗОБРАЖЕНИЙ ОБЪЕКТОВ НЕРЕГУЛЯРНОГО ВИДА

С целью обеспечения эффективного решения задачи сегментации изображений объектов нерегулярного вида в автономных системах технического зрения в условиях неопределенности в работе рассматриваются методологические основы машинного самообучения, основанного на использовании прогностических функций и эталонов изображений рассматриваемых объектов и сцен.

Ключевые слова: сегментация, обучение, самообучение, прогноз, эталон.

Введение

Обучение – это процесс настройки системы технического зрения (СТЗ) на решение новых в определенном смысле задач обработки изображений, в результате которой СТЗ постепенно приобретает способность устойчивого функционирования. СТЗ может настраиваться на обработку изображений объектов новых классов, или на обработку изображений объектов известных классов в новых условиях получения изображений по масштабу, освещению, зашумлению. Целью машинного обучения является автоматизация процесса решения задач обработки изображений определенного типа [1].

Построение СТЗ возможно только на основе априорной информации о классах рассматриваемых объектов, их свойствах, а также условиях получения изображений объектов рассматриваемых классов. Когда имеется полная априорная информация, можно построить систему моделей изображений, произвести настройку критериев и определить стратегии анализа изображений, которые обеспечивают заданные характеристики СТЗ в отношении уровня автоматизации, вычислительной эффективности и качества получаемых решений [2 – 5]; в этом случае нет необходимости в обучении. В противном случае единственной возможностью создания СТЗ с заданными характеристиками является внедрение в систему модуля обучения с тем, чтобы производить настройку СТЗ в ходе ее функционирования.

Таким образом, можно выделить два основных подхода к построению СТЗ: 1 – построение по результатам идентификации модели и настройки критериев с реализацией обучения в качестве вспомогательного модуля; 2 – построение с реализацией обучения как основного модуля, предназначенного для настройки СТЗ на особенности изображений объектов и условия получения изображений [1].

Первый (классический) подход является предпочтительным потому, что он позволяет произвести не только настройку параметров системы, но также позволяет оценить стратегии и построить адаптивные процедуры обработки изображений, привлекая эксперта для решения некоторых нетривиальных за-

дач в условиях неопределенности. Второй подход в настоящее время используется в СТЗ, построенных на основе нейронных сетей, и используется на практике лишь для настройки параметров системы [6].

Обучение может выполняться в автоматическом режиме, в полуавтоматическом режиме с задержкой, в интерактивном режиме с ожиданием.

В первом случае обучение, выполняемое СТЗ в автоматическом режиме без участия эксперта – это самообучение, цель которого состоит в том, чтобы предварительно обработать полученное изображение произвести корректировку своих настроек.

Во втором случае нетиповое изображение переправляется к эксперту, а на вход СТЗ поступает новое изображение; эксперт производит обучение и корректирует настройки СТЗ; после этого рассмотренное экспертом изображение снова поступает на вход СТЗ и обрабатывается теперь как типовое.

В третьем случае нетиповое изображение переправляется к эксперту, СТЗ останавливает свою работу, эксперт производит обучение и корректирует настройки СТЗ, после этого СТЗ возобновляет свою работу и обрабатывает текущее изображение как типовое с новыми настройками.

Обучение может выполняться до сегментации, или после нее. В первом случае обучение выполняется в том случае, если установлено, что поступившее на вход СТЗ изображение является не типовым.

Во втором случае, обучение может использоваться для корректировки настроек СТЗ после сегментации (распознавания) для учета не типовых снимков, которые нельзя отнести к таковым лишь по интегральным оценкам входного изображения. Необходимость такой корректировки может устанавливаться экспертом, или самой СТЗ; основой для принятия решения о необходимости корректировки служат оценки качества результатов сегментации.

Проведение обучения с экспертом по своей сути аналогично решению задач идентификации и настройки; важнейшей особенностью облегчающей обучение с экспертом является возможность использования эталонов изображений. При этом основную проблему в настоящее время представляет собой

обучение без эксперта, т.е. самообучение, поскольку лишь при решении некоторых тривиальных задач мы можем построить для целей обучения эталоны рассматриваемых изображений.

Основная задача работы, таким образом, состоит в том, чтобы описать методологические основы машинного самообучения для эффективного решения задачи сегментации изображений объектов нерегулярного вида в автономных системах технического зрения в условиях неопределенности.

1. Основы самообучения

Самообучающаяся СТЗ – это такая СТЗ, которая способна в ходе обучения самостоятельно изменять алгоритм своей работы и настройки в соответствии с условиями функционирования так, что с течением времени она улучшает свои характеристики и качество решений. В такой общей постановке задача самообучения относится к области искусственного интеллекта и в полной мере в работе решена быть не может. Поэтому, предполагая неизменность структуры моделей изображений, а также стратегий анализа изображений, рассмотрим задачу самообучения в смысле настройки параметров.

Самообучение необходимо в том случае, когда на вход СТЗ поступают не типовые данные, т.е. такие данные, на анализ которых СТЗ не была ориентирована. Этот факт, как правило, устанавливается по тем, или иным оценкам получаемых изображений (например, по средней яркости снимка), предварительно их анализу. Задача машинного самообучения – настроить параметры системы с целью автоматизации последующего процесса анализа изображений с новыми скорректированными параметрами.

Блок самообучения должен предшествовать каждой операции анализа изображения (предобработке, локализации и сегментации). Включаться в работу этот блок будет по факту в том случае, если оценка того или иного свойства изображения будет не типовой. Так, например, если на входе средняя яркость \bar{f} изображения не содержится в таблице настроек, это означает, что для такого уровня яркости идентификация модели и настройка не производились; следовательно, для устойчивой работы СТЗ эти операции необходимо выполнить.

Построение методов самообучения задача очень сложная. Поэтому для минимизации вероятности возникновения ситуации, в которой необходимо самообучение, на этапах идентификации и настройки необходимо рассматривать не только типовые условия, но также и возможные вариации условий получения изображений.

Для обеспечения адекватной и устойчивой сегментации, а также эффективного самообучения необходима системность моделирования [7]. Это свойство выражается в том, чтобы комплексно рассматривать изображения с топологической, геометрической, яркостной, контрастностной, цветовой и мультиспектральной точек зрения.

При этом (для делимости объектов и фона друг от друга в сложных условиях получения данных) возникает необходимость использовать структурные модели, а также мощные критерии и методы, устойчивые к определенным вариациям данных за счет инвариантности и адаптивности. Основными в этом отношении являются статистико-оптимизационные гистограммные и регрессионные яркостные модели, критерии и, соответствующие им методы сегментации изображений.

Методы самообучения в главном можно подразделить на методы использующие эталоны, и методы, не использующие эталоны. В условиях наличия эталонов изображений рассматриваемых объектов, самообучение выполняется аналогично решению задач идентификации модели и настройки критериев анализа данных в автоматическом режиме. Однако в большинстве случаев СТЗ не имеет эталонов, что для объектов нерегулярного вида являет собой наибольшую проблему при самообучении.

2. Построение прогностической модели

В отсутствие эталонов наиболее очевидным способом обучения является использование дедуктивных правил вывода, а именно – прогнозирования, суть которого состоит в том, чтобы по таблицам настроек построить систему уравнений регрессии с целью распространения полученных закономерностей на нерассмотренную область изменения аргумента.

Хотя производить такое прогнозирование можно сразу же после построения таблиц настроек, тем не менее, построение таких прогнозов является основой адекватной сегментации в не типовых условиях получения данных.

Основой для построения прогноза метрических параметров изображений является многомасштабное представление (рис. 1). Основой для построения прогноза уровней яркости является представление изображений объекта в условиях вариаций яркости (рис. 2); аналогичным образом подготавливаются данные для оценивания иных элементов яркостной и контрастностной моделей. При этом обязательно должны быть учтены искажения изображения и, в том числе, уровни зашумления изображения (рис. 3).

На этой основе составляются таблицы настроек и по ним строятся уравнения регрессии, используемые для целей прогнозирования. На основе полученных функций прогноза составляются расширенные таблицы настроек, содержащие как истинные, так и прогнозируемые значения.

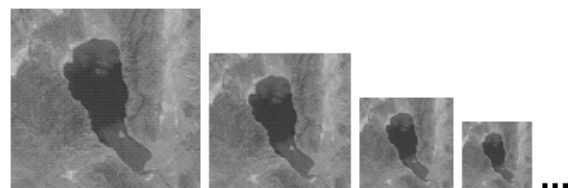


Рис. 1. Многомасштабное представление озера

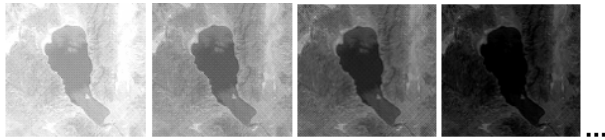


Рис. 2. Снимки озера в различных условиях освещения

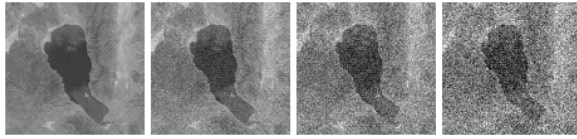


Рис. 3. Снимки озера при различных уровнях зашумления

Так, например, для учета вариаций условий получения изображений может быть выбрана средняя яркость снимка \bar{f} . В таких условиях область изменения средней яркости \bar{f} разбивается на некоторое количество диапазонов $\{\Delta\bar{f}_i\}_{i=1,n}$ и, соответственно этому разбиению, находят характеристики изображения в зависимости от средней яркости снимка. Результат такой идентификации для оценки яркости объекта в интервальном виде представляется в виде таблицы (табл. 1).

Таблица 1

Яркостная модель в интервальном виде

Средняя яркость снимка	$\Delta\bar{f}_1$...	$\Delta\bar{f}_n$
Яркость объекта	$[f_{\min}, f_{\max}]_1$...	$[f_{\min}, f_{\max}]_n$

Для решения задач прогнозирования на основе таблицы настроек строятся уравнения регрессии границ яркости объекта $f_{\min}(\bar{f})$, $f_{\max}(\bar{f})$ (рис. 4).

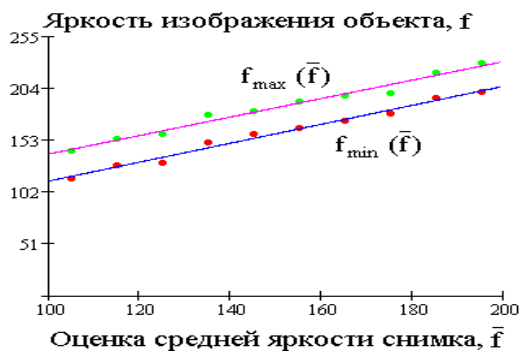


Рис. 4. Регрессия границ яркости объекта

При построении регрессии решаются следующие основные задачи: 1) выбор модели регрессии, 2) оценка параметров в выбранной модели регрессии, 3) проверка статистической гипотезы о значимости параметров, 4) проверка статистической гипотезы об адекватности полученного уравнения регрессии. В результате получаем искомую модель регрессии и оценки ее параметров [8, 9].

Особенность построения регрессии состоит в том, что различные элементы модели представляют-

ся с использованием различных шкал [7]. В этом отношении регрессия в классическом смысле строится лишь для числовых элементов моделей. Для не числовых (структурных) элементов моделей, например, для топологического типа, вместо регрессии строится функция соответствия топологического типа масштабу представления изображения вида:

$$F_{co}(s) = \begin{cases} \text{type } T_1, & \text{if } M_1 \leq s < M_2, \\ \text{type } T_2, & \text{if } M_2 \leq s < M_3, \\ \dots & \\ \text{type } T_n, & \text{if } M_n \leq s < M_{n+1}, \end{cases}$$

где $\text{type } T_i$ – топологический тип; s – площадь; M_i – ограничения по масштабу. При этом если изображениям объектов в некотором масштабе могут соответствовать различные типы, они задаются перечислением вида: $F_{co}(s) = (\text{type } T_1, \text{type } T_3)$.

Таким образом, задаются все не числовые элементы модели. Поскольку общих методов построения регрессии и прогнозирования для структурных элементов модели не существует, функция соответствия должна составляться таким образом, что бы заменить собой и регрессию и прогноз.

Регрессии в широком смысле слова играют преобладающую роль в прогнозировании статистическими методами. Это обстоятельство подтверждается опытом прогностических разработок. Когда регрессия специфицирована и оценены ее параметры, она может быть применена для прогнозирования. Оценка прогностических величин получается с помощью простой операции подстановки в уравнение регрессии значений независимых переменных.

Если рассматривать прогноз за пределы диапазона наблюдаемых значений, что представляет для нас наибольший интерес, то получаемые оценки имеют ценность в той мере, в которой есть основания полагать, что принятая форма регрессии может быть распространена на некоторую область за пределами диапазона наблюдаемых значений. Это предположение может быть верно, или не верно, в зависимости от специфики решаемой задачи. Однако в отсутствие информации о поведении исследуемой зависимости за пределами диапазона наблюдаемых значений использование регрессии является важнейшим инструментом прогнозирования [9].

Один из наиболее распространенных методов прогнозирования заключается в экстраполяции, т.е. в продлении регрессии за пределы диапазона наблюдаемых значений. Экстраполяция базируется на следующих допущениях: 1) развитие явления может быть с достаточным основанием охарактеризовано плавной (эволюторной) траекторией; 2) общие условия, определяющие тенденцию развития внутри диапазона наблюдаемых значений, не претерпят существенных изменений за его пределами.

Экстраполяция дает точечную прогностическую оценку. Однако из-за наличия ошибки при-

ближения наблюдаемого распределения регрессией естественным является желание оценить величину такой ошибки. Для этих целей внутри диапазона наблюдаемых значений мы можем использовать дисперсию. Однако распространить эти оценки погрешности на нерассмотренный диапазон наблюдаемых значений мы не можем, поскольку вероятность совпадения экстраполяции и регрессии на нерассмотренном диапазоне значений – явление маловероятное, что означает рост ошибки прогнозирования тем больше, чем дальше берутся значения от границ диапазона наблюдаемых значений. Поэтому одна из важнейших задач, возникающих при экстраполяции тренда, заключается в определении доверительных интервалов прогноза. В общем случае доверительный интервал регрессии определяется в виде полосы: $f_x \pm t_\alpha \sigma_f$, где f_x – расчетное значение по регрессии, t_α – значение t-статистики Стьюдента, σ_f – СКО регрессии [9].

3. Настройка СТЗ с использованием прогностических функций

Для целей самообучения могут использоваться эталоны и, следовательно, классические методы идентификации и настройки в автоматическом режиме. Однако на практике получение таких эталонов СТЗ для объектов нерегулярного вида представляет собой большую проблему при самообучении.

Использование эталонов в детерминированном случае. При решении некоторых задач типы объектов поля зрения регламентированы; кроме того, может производиться позиционирование камеры на объекте. В таких условиях применение методов локализации и сегментации, например, вычитания из фона, вырезания из фона, или ряда иных позволяет нам получить изображение требуемого объекта, построить его фотореалистичный эталон и применить методы идентификации и настройки по эталону в автоматическом режиме.

Использование эталонов в условиях неопределенности. При решении не тривиальных задач объекты поля зрения не регламентированы. В таких условиях применение методов вычитания из фона, или вырезания из фона может дать нам изображение, лишь предположительно того класса объектов, которое мы ищем. В таких условиях, необходимо оценить к какому классу объектов принадлежит полученное изображение; после этого можно использовать полученное изображение в качестве эталона, как и в детерминированном случае.

Таким образом, важнейшей в условиях неопределенности является задача классификации изображений по системе заданных классов, которая может ставиться в смысле дискриминации, или в смысле идентификации. Очевидно, что решение в этом случае будет носить вероятностный характер.

Для целей классификации изображений в условиях неопределенности предлагается метод класси-

фикации с прогнозированием, основные шаги которого состоят в следующем.

Фильтрация. Оцениваются модельные характеристики полученного изображения, и производится их сравнение с аналогами изображений каждого класса объектов с целью фильтрации (отсева) классов, которым изображение принадлежать не может. Для этих целей для структурных элементов модели используются функции соответствия, а для параметрических – прогностические функции и их доверительные интервалы. В результате фильтрации получаем множество классов $\{k_j\}_j$, которым может принадлежать рассматриваемое изображение.

Классификация. После выполнения фильтрации необходимо произвести дискриминацию (идентификацию) рассматриваемого изображения по системе классов $\{k_j\}_j$, $j=1, \dots, J$ (в отношении определенного класса k_*).

Дискриминация. Для целей дискриминации получаемых изображений в настоящее время применяются критерии теории решений, основанные на анализе степени близости характеристик изображений в пространстве признаков; структурные признаки изображений после фильтрации не рассматриваются. Одними из наиболее распространенных критериев являются классификатор по минимуму расстояния, корреляционное сопоставление и, так называемые, статистически оптимальные классификаторы, например, байесовский классификатор [1].

Идентификация. Для целей идентификации получаемых изображений в настоящее время применяются пороговые критерии, основанные на оценках степени близости характеристик изображений в пространстве признаков; структурные признаки изображений после фильтрации не рассматриваются. В качестве оценок степени близости изображений в пространстве признаков рассматривают функции расстояния, корреляционного сопоставления, а также дискриминантные функции. Множественная классификация. В условиях неопределенности может решаться задача множественного выбора. Суть решения этой задачи состоит в том, что не отвергается принадлежность рассматриваемого изображения ни одному из классов $\{k_j\}_j$. На основе этого предположения и выбранной функции сходства, рассматриваемое изображение относится к классу k_j с некоторой вероятностью p_j .

Множественная идентификация. Хотя задача идентификации ставится для проверки на принадлежность изображения определенному классу объектов, тем не менее, в условиях неопределенности, задача идентификации может решаться для всех классов k_j . В таких условиях, как и при множественной классификации, для каждого k_ξ класса, ($\xi=1, \dots, \Xi$), которому может принадлежать изображение, задается вероятность принадлежности p_ξ .

Множественная классификация (идентификация) в условиях неопределенности нужна для того, чтобы не потерять объект при низкой точности прогноза, или близости прогнозируемых характеристик изображений объектов различных классов.

Таким образом, в результате решения задачи классификации получаем изображение, которое с некоторой вероятностью можно отнести к определенному классу объектов; принимаем это изображение в качестве эталона, производим идентификацию модели и настройку критериев.

После самообучения корректируем (дополняем) таблицу настроек и уточняем прогностические функции. В случае множественного результата задача обучения и, затем, сегментации решается J (Ξ) раз для каждого класса объектов, после чего при распознавании (и, впоследствии, прикладном анализе объектов по их изображениям) производится принятие окончательного решения о верной классификации. После отсева ложных изображений производится корректировка таблиц настроек и прогностических функций: лишь зная объект можно верно скорректировать таблицы настроек и прогностические функции определенного класса.

Настройка без использования эталонов в условиях неопределенности. Если эталон не может быть получен, тогда для получения искомого настроек СТЗ мы можем использовать лишь прогностические функции (экстраполяции) того класса объектов, изображения которого стремимся сегментировать.

Использование интеллектуальных систем поддержки принятия решений. При решении любой задачи в СТЗ (самообучение, предобработка, локализация, сегментация и др.) исходные данные могут представляться не одним лишь изображением и результатами его анализа на предыдущем этапе, но также и данными, полученными из сторонних источников (датчиков, или систем). Для целей самообучения такие данные требуются в условиях неопределенности с целью исключения двусмысленности при классификации. В этом отношении интеллектуальные системы как раз и нужны для обработки получаемых разнотипных данных с целью уstra-

нения неопределенности путем применения правил вывода на основе аппарата алгебры логики и реляционной алгебры.

Выводы

С целью обеспечения эффективного решения задачи сегментации изображений объектов нерегулярного вида в автономных системах технического зрения в условиях неопределенности в работе рассмотрены методологические основы машинного самообучения, основанного на использовании прогностических функций и эталонов изображений рассматриваемых объектов и сцен. При этом показано, что ключевым, для обеспечения автоматизации обучения и высокой точности результатов настройки и сегментации является использование эталонов рассматриваемых изображений.

Список литературы

1. Gonzalez R. *Digital Image Processing. Second Edition* / R. Gonzalez, R. Woods. – Prentice Hall, 2002. – 793 p.
2. Форсайт Д. *Компьютерное зрение. Современный подход: Пер. с англ / Д. Форсайт, Ж. Понс.* – М.: Вильямс, 2004. – 928 с.
3. Яне Б. *Цифровая обработка изображений / Б. Яне.* – М.: Техносфера, 2007. – 584с.
4. Шапиро Л. *Компьютерное зрение; пер. с англ. / Л. Шапиро, Дж. Стокман.* – М.: БИНОМ, 2006. – 752 с.
5. Sonka M. *Image processing, analysis, and machine vision / M. Sonka, V. Hlavak, R. Boyle.* – California (USA): Cole Publishing Company, 1999. – 770 p.
6. *Искусственный интеллект.* – В 3 кн. Кн. 1. *Системы общения и экспертные системы / Под ред. проф. Э.В. Попова.* – М.: Радио и связь, 1990. – 464 с.
7. Лямец В.И. *Системный анализ. Вводный курс: учебн. пособ. / В.И. Лямец, А.Д. Тевяшев.* – Х.: ХТУРЭ, 1998. – 252 с.
8. Дрейпер Н. *Прикладной регрессионный анализ: В 2-х кн. Кн. 1 / Пер. с англ. – 2-е изд., перераб. и доп. / Н. Дрейпер, Г. Смит.* – М.: Финансы и статистика, 1986. – 366 с.
9. Четыркин Е.М. *Статистические методы прогнозирования. Изд. 2-е, перераб. и доп. / Е.М. Четыркин.* – М.: Статистика, 1977. – 200 с.

Поступила в редколлегию 5.02.2009

Рецензент: д-р техн. наук, проф. И.В. Рубан, Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков.

МЕТОДОЛОГІЧНІ ОСНОВИ МАШИННОГО САМОНАВЧАННЯ ДЛЯ СЕГМЕНТАЦІЇ ЗОБРАЖЕНЬ ОБ'ЄКТІВ НЕРЕГУЛЯРНОГО ВИДУ

К.С. Смеляков

З метою забезпечення ефективного рішення задачі сегментації зображень об'єктів нерегулярного виду в автономних системах технічного зору в умовах невизначеності в роботі розглядаються методологічні основи машинного самонавчання, заснованого на використанні прогностичних функцій і еталонів зображень об'єктів і сцен, що розглядаються.

Ключові слова: сегментація, навчання, самонавчання, прогноз, еталон.

MACHINE LEARNING METHODOLOGICAL CONCEPTS FOR IRREGULAR OBJECT IMAGE SEGMENTATION

K.S. Smelyakov

In order to increase a stability of irregular object image segmentation in condition of uncertainty for autonomous vision systems the machine learning models are proposed which are based on forecasting functions and image templates for the considered objects and scenes.

Keywords: segmentation, teaching, self-training, prognosis, standard.