

УДК 004.932.72'1

К ВОПРОСУ О ПОСТРОЕНИИ СИСТЕМЫ РАСПОЗНАВАНИЯ И ПОДСЧЕТА ЖИВОТНЫХ НА АЭРОФОТОСНИМКАХ

Часть 1: Анализ методов распознавания

В. В. Михайлов,

доктор техн. наук, ведущий научный сотрудник

Санкт-Петербургский институт информатики и автоматизации РАН

Я. В. Харин,

аспирант

Санкт-Петербургский государственный университет аэрокосмического приборостроения

Рассматриваются основные принципы и этапы построения системы подсчета и распознавания объектов на фотографиях. Проводится обзор методов сегментации изображений и распознавания. Разбираются их существенные достоинства и недостатки для решения задачи подсчета количества животных.

Ключевые слова — распознавание, сегментация, подсчет объектов.

Введение

В настоящее время задача автоматического распознавания и подсчета объектов является актуальной и востребованной. Ее решение способно автоматизировать труд человека и повысить производительность. Видеонаблюдение, регулирование движения транспорта, контроль качества деталей на конвейере, оценка численности животных — вот далеко не полный перечень областей человеческой деятельности, где необходимо решение данной задачи. Следует заметить, что создание каждой системы требует учета особенностей объектов распознавания, а также особенностей фотоснимков либо видеоряда для обоснованного выбора методов распознавания.

В настоящей статье рассматриваются основные принципы построения компьютерной системы распознавания и подсчета для определения количества животных на аэрофотоснимках. Приводятся результаты качественного анализа методов, подходящих для решения поставленной задачи. Конкретным объектом при построении прототипа системы выбраны дикие северные олени.

Материалы и методы

Северные олени являются важнейшей компонентой полярных экосистем, основным источником питания, благосостояния и этнической само-

бытности коренных народов Севера. Величина допустимого промыслового изъятия диких северных оленей рассчитывается на основе данных авиаучетов о численности и половозрастном составе популяции. Основные группировки оленей фотографируются во время их скоплений на летних пастбищах, и количество животных в них подсчитывается. Ручная обработка снимков при численности популяции 500–600 тыс. особей занимает около 3 мес. Для определения квоты к началу промыслового сезона время обработки снимков должно быть снижено до 10–15 дней. Автоматизация процесса обработки фотоснимков позволит, таким образом, решить две задачи: освободить специалистов от выполнения рутинной работы и повысить качество функционирования промысловой системы. При съемке олени находятся на различном удалении от камеры, поэтому их изображения на снимках будут видны под различными углами (от 45 до 90°), изображения будут иметь различные размеры и могут перекрывать друг друга. Помехи: камни, земляные бугры, впадины и т. п. — легко идентифицируются при ручной обработке снимков, но могут создать трудности при работе автоматической системы распознавания.

Система распознавания и подсчета животных должна решать следующие задачи.

1. Распознавать и подсчитывать общее число животных на снимках. При этом животные могут быть представлены как локальными объекта-

ми, так и неразделимыми группами. Неоднородный по цвету и фактуре природный фон может содержать помехи — камни, овраги и т. д.

2. Распознавать и подсчитывать количество животных, имеющих визуально различимые признаки. Для северных оленей — это телята и взрослые самцы. Условия распознавания по фону и помехам соответствуют п. 1.

В качестве первичной информации при разработке системы использованы фотоснимки групп и скоплений животных, сделанные во время авиачетов диких северных оленей на Таймыре в 2000, 2003 и 2009 гг. Кроме того, для получения количественных оценок правильности распознавания использовались автоматически сегментированные изображения стад животных.

Необходимо ввести ряд понятий и определений. Под *объектами* понимаются некоторые сущности, запечатленные на снимке, подлежащие подсчету. *Класс объектов* — некоторая совокупность объектов, называемых *элементами класса*, обладающих рядом близких свойств. Измеряемые или вычисляемые *свойства объектов*, позволяющие отличить классы друг от друга, называются *признаками*.

В общем случае в решении задачи подсчета объектов можно выделить следующие этапы [1]: предобработка снимков, сегментация, шумоподавление и фильтрация, отнесение сегментированных областей к классам объектов, дополнительная обработка некоторых классов объектов, подсчет количества найденных объектов.

Предобработка

Первый этап необходим для подготовки изображения к распознаванию. На этом этапе производится очистка изображений от помех и шумов. Под помехами и шумами понимаются сторонние возмущения, неселективные в отношении объектов и фона, действующие в системах создания, передачи и воспроизведения фотоснимков. Например, некоторые помехи могут быть результатом дефектных пикселей на матрице цифрового фотоаппарата или возникать в результате аппаратной дискретизации и квантования. При удалении помех важно выбрать такой способ очистки изображения, чтобы он не вызвал значительных искажений изображения, сохраняя объекты распознавания. В качестве фильтров для удаления помех и шумов служат различного рода усредняющие, частотные и пространственные фильтры [1, 2].

На этом этапе может быть увеличена яркость, повышена четкость изображения, могут быть применены операции усреднения и выравнивания гистограмм яркости. Если это требуется на последующих этапах, возможно снижение дисперсии яркости пикселей с сохранением резких

перепадов яркости. Выбор преобразований должен учитывать метод сегментации для сохранения признаков объектов. При необходимости может быть осуществлен переход из одной цветовой модели в другую. При выполнении сегментации по цветовым признакам бывает удобно преобразовать изображение в цветовую модель HSV.

Сегментация

Под сегментацией понимается процесс проверки каждого отдельного пикселя для того, чтобы выяснить, принадлежит ли он к интересующим объектам или нет. Результатом сегментирования изображения является бинарное изображение, в котором выделены области, обладающие признаками объектов в соответствии с критериями сегментации и признаками фона. Метод сегментации выбирается в зависимости от особенностей конкретной решаемой задачи.

Если объекты имеют четкие и стабильные границы, то, как правило, применяются методы выделения границ. Изображения рассматриваются как функция двух переменных, при этом производится поиск максимума градиента этой функции. Примерами таких методов служат фильтры Робертса, Кирша, Превита и Собеля. Главной проблемой этих методов является слабая устойчивость к помехам и шуму, поэтому их целесообразно применять, например, при сегментации объекта на монотонном фоне.

Если на изображении присутствуют стабильные различия в яркости (интегральной или спектральной) или различия в каком-либо другом значимом признаке отдельных областей, то целесообразно применять пороговые методы. Такие методы позволяют выделить области изображения, для которых значение выбранного параметра выше либо ниже определенного порога. Например, когда объект имеет яркость большую, чем остальная часть изображения, применение порогового фильтра даст хорошие результаты.

При наличии связности внутри отдельных сегментов применяются методы наращивания областей. Идея состоит в том, что выбираются стартовые точки, после чего производится анализ соседних с ними точек в соответствии с некоторым критерием однородности. Этим критерием, например, может служить яркость в некотором диапазоне [2]. Количество стартовых точек должно быть равно количеству однородных областей на изображении. Метод водоразделов является одним из эффективных способов практической реализации идеи наращивания областей. Он основан на поиске локальных минимумов с последующей группировкой вокруг них областей по связности.

Если связь между пикселями изображения в пространстве признаков задана в математиче-

ской форме, то для сегментации могут быть применены методы теории графов. Суть методов в следующем: изображение представляется в виде взвешенного графа, вершинами которого являются пиксели изображения. Вес ребра графа отражает близость точек в некотором пространстве признаков. Для снижения размерности, как правило, рассматриваются ребра графа, связывающие близлежащие пиксели. Затем производится решение задачи поиска минимальной стоимости разреза графа. Таким образом, изображение разбивается на однородные области, однородностью которых можно управлять, задавая вес ребра графа. Помимо однородности цвета и текстуры сегментов, можно управлять размером областей, их формой, сложностью и т. д.

Методы сегментации могут использоваться совместно, если это позволяет улучшить выделение искомым объектов на изображениях. Примером этому может служить совместная работа пороговых методов сегментации и методов наращивания областей. В этом случае пороговый метод может выделить яркостные минимумы изображения, а метод водораздела выделит весь объект, имеющий яркостный минимум.

Шумоподавление и фильтрация

Третий этап необходим для удаления помех, возникающих при сегментации. Для этого, как правило, используется обработка с помощью аппарата математической морфологии, поскольку изображение на данном этапе представляется бинарным [3, 4]. Может производиться дополнительная обработка сегментированного изображения, например операция сглаживания бинарных областей или удаление областей определенной формы.

Распознавание

Входными данными для распознавания объектов являются изображения, полученные в результате процессов сегментации и шумоподавления. Помимо этого, здесь могут использоваться любые изображения, полученные на предыдущих этапах, и исходное изображение.

Широкое распространение при обнаружении и распознавании получили корреляционные методы, работающие с объектами в пространстве изображений или с признаками объекта в пространстве признаков [5, 6]. При работе с объектами задается эталон объекта, после чего производится многошаговая корреляция. По сути, данный метод реализует полный перебор в пространстве изображений (пространстве сигналов).

Методы, основанные на пространстве признаков, обладают значительно меньшей размерностью по сравнению с пространством сигналов.

Признаки могут сравниваться как с использованием порогов по величине сходства, так и без порога. При этом решение о принадлежности к тому или иному классу может приниматься на основе разнотипных признаков: метрических, статистических, логических, текстурных, структурно-лингвистических. При необходимости выполняется корреляционная обработка признаков, полученных от эталона и входного изображения.

Главной задачей при этом является выбор признаков. Набор признаков, используемых для распознавания объектов, должен удовлетворять следующим условиям:

- близости значений признака для объектов одного класса, существенное различие значений признака для объектов разных классов;
- набор признаков должен быть полным, т. е. в совокупности должен обеспечивать идентификацию объектов любого из классов;
- общее количество признаков должно быть минимальным.

Свойства природных объектов в значительной мере варьируют, объекты могут иметь разные размеры, изображения объектов могут перекрывать друг друга. На изображениях могут быть помехи, близкие по цветовой гамме и форме к искомым объектам. По этой причине можно говорить не о строгом распознавании, а о распознавании с некоторой вероятностью. При этом для уменьшения вероятности ошибок в минимальный набор могут вводиться добавочные, избыточные признаки.

Дополнительная обработка

Дополнительная обработка классов объектов после распознавания проводится для подготовки к последующим действиям над ними, которые требует решаемая задача. Например, на этом этапе может производиться оценка расстояния до распознанных объектов, проверка правильности распознавания пользователем или с помощью логических, синтаксических и прочих методов. Возможен итерационный возврат к предшествующим этапам обработки изображений и распознавания. Так, если были выявлены ошибки при распознавании, информация об этом может поступать на предшествующие этапы для их исправления [7].

Результаты анализа методов распознавания

Предобработка

На данном этапе в связи с необходимостью поиска признаков объектов был выбран медианный фильтр [2]. Выбранный фильтр показал лучшие результаты удаления помех по сравнению с линейными сглаживающими фильтрами, сохранив при этом четкость изображения.



■ Рис. 1. Удаление шумов с помощью медианного фильтра: слева — исходное изображение; справа — результат фильтрации

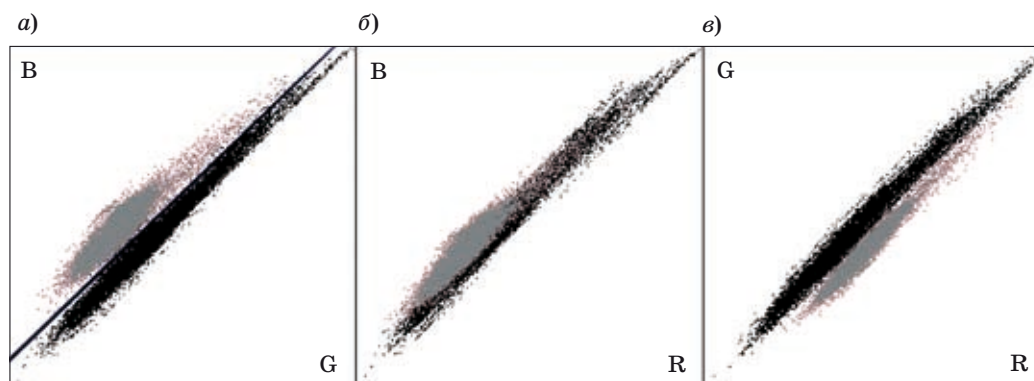
При использовании медианного фильтра важно определить размер окна фильтра. Если окно имеет слишком большой размер, то происходит снижение четкости изображения объектов. При оптимальном размере окна фильтр удаляет мелкие шумы и помехи, не снижая в общем четкости изображения. Практическим путем было установлено, что для удаления помех на представленных снимках размер окна должен составлять 0,1–0,5 от среднего размера объекта. Пример работы медианного фильтра показан на рис. 1. Как видно из рисунка, практически все мелкие аддитивные помехи в результате фильтрации со снимка удалены.

Сегментация

Поскольку объекты распознавания находятся на неоднородном фоне и имеют разные оттенки цвета, то методы выделения границ не смогут дать хороший результат. Распознаваемые объекты, как правило, контрастируют на зеленом фоне. В связи с этим был выбран пороговый метод сегментации. В качестве порога используется отношение спектральной яркости одной составляющей цвета к другой. Для отбора спектральной пары был проделан эксперимент, в котором использовались участки изображений объектов

и фона, полученных из аэрофотоснимков. При проведении эксперимента были взяты фрагменты всех имеющихся в наличии типов фотографий. Другими словами, из множества снимков были перенесены объекты на одно изображение. На другое изображение были помещены фоновые цвета снимков. После чего оба изображения подверглись анализу: каждому пикселю изображения объектов и фона была поставлена в соответствие точка на координатной плоскости. Координатами точки являются значения яркости составляющих цвета. В результате получены три графика скопления точек, соответствующих различным спектральным парам (рис. 2). Черные точки на графиках соответствуют пикселям фона, серые — пикселям объектов.

Видно, что скопления точек на рис. 2, а не перекрывают друг друга, в отличие от рис. 2, б и в. Именно эта спектральная пара была использована нами для отделения объектов от фона. Для решения задачи сегментации проведем прямую, разграничивающую скопления точек. Преобразовав уравнение прямой в неравенство, можно выделить либо только верхнее скопление точек, либо только нижнее. На изображении будут выделяться пиксели объекта или фона в зависимо-



■ Рис. 2. Цветовые зависимости фона и объектов распознавания: а — для зеленой и голубой компонент цвета; б — для красной и голубой, в — для красной и зеленой

сти от поставленного в неравенстве знака. Коэффициенты уравнения прямой были рассчитаны из условия минимума суммы точек, попадающих в чужую область. При этом вероятность ошибочного отнесения области изображения фона к объекту составила около 2 %, тогда как вероятность ошибочного отнесения области изображения объекта к фону — 0,001 %. Приведенные числа являются результатом эксперимента на эталонных изображениях, по которым производился поиск порога. При сегментации других изображений вероятность ошибки может существенно возрасти. Помимо этого, при вычислениях не учитывалась возможность наличия на изображении предметов, например камней, похожих по цветовым характеристикам на животных.

Результат сегментации изображения пороговым методом представлен на рис. 3. Выбранный метод устойчив к сложным формам объектов, которые возникают в результате наложения изображений единичных объектов друг на друга.

Использованный нами алгоритм сегментации позволяет выделять объекты любой формы и обладает высоким быстродействием благодаря простому методу проверки принадлежности пикселя к объекту. Однако, как можно заметить на рис. 3, некоторые участки изображения были ошибочно сегментированы.



■ Рис. 3. Результаты сегментации изображения пороговым методом: сверху — исходное изображение; снизу — сегментированное

Были опробованы методы теории графов. Вначале был применен метод сегментации SWA (*Segmentation by Weighted Aggregation*) [8]. Его суть состоит в построении пирамиды взвешенных графов, где каждый верхний слой получен из нижнего путем объединения вершин графа, вес ребер между которыми минимален. В качестве параметра веса ребра графа была взята разница пикселей в цветовом пространстве. Результаты работы метода представлены на рис. 4.

Получившееся изображение можно обработать пороговым фильтром, чтобы привести его к бинарному виду (рис. 5).

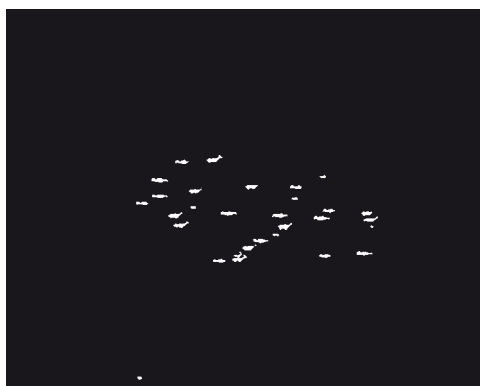
Этот алгоритм сегментации выделил только объекты. Его достоинством является высокая надежность. К недостаткам можно отнести сложность вычислений, что негативно сказывается на быстродействии системы.

При наличии двух бинарных изображений появляется возможность учитывать результаты каждого из способов совместно. Например, можно применить операцию логического «И» для двух изображений и получить одно, являющееся результатом работы обоих методов (рис. 6).

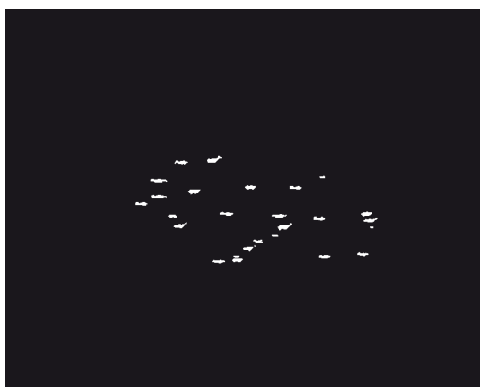
При совмещении методов увеличивается общее количество операций, а значит, ухудшается быстродействие системы. Одновременно с этим, когда применяется несколько методов сегмента-



■ Рис. 4. Результаты применения алгоритма SWA: сверху — до обработки; снизу — после обработки



■ Рис. 5. Результат работы порогового фильтра после применения сегментации методом SWA



■ Рис. 6. Логическое сложение результатов пороговой сегментации и сегментации с помощью метода SWA

ции, использующих различные признаки объекта, общая вероятность ошибки значительно снижается. В рассматриваемом случае время выполнения сегментации увеличивается незначительно. Это связано с небольшим количеством операций для выполнения порогового метода сегментации. Кроме того, оба метода дополняют друг друга, так как используют различные признаки объектов. Метод SWA основан на связности пикселей. Пороговый метод использует цветовые характеристики объектов. При ошибочной сегментации области изображения одним методом второй метод исправит ошибку первого. Таким образом, при совмещении двух методов получен значительный выигрыш в надежности системы при незначительном уменьшении ее быстродействия. Использование комбинации этих методов является хорошим решением для снижения погрешности распознавания. После объединения результатов двух методов уменьшилось также количество контуров, перекрывающих друг друга.

Для оценки погрешности системы сегментации был проведен эксперимент на автоматически сгенерированных изображениях стад [9]. При ге-

нерации использовался реальный, но одинаковый для всех изображений фон и фигурки животных с естественными цветовыми характеристиками. Число объектов на таких изображениях известно заранее. В результате проведения этого теста на 91 изображении стад было установлено, что погрешность подсчета с применением описанной системы составляет около 8 % при относительно небольшой дисперсии, равной 1,02. При вычитании из данной погрешности доли ошибок, связанных с наложением объектов друг на друга, погрешность составляет 3 %.

Шумоподавление и фильтрация

На этом этапе в описываемой системе производится сглаживание сегментированных областей и удаление мелких помех. Необходимость сглаживания вызвана тем, что границы объекта после применения сегментации могут быть крайне неровными, вследствие чего может быть затруднен анализ сегментированных областей на последующих этапах. Сглаживание выполняется с помощью набора операций математической морфологии — операции эрозии и масштабного преобразования [5]. Такой подход, помимо сглаживания, удаляет мелкие помехи, которые, как правило, присутствуют в большом количестве после проведения сегментации пороговыми методами (рис. 7).

Как можно заметить, в результате сглаживания границы областей стали более ровными, удалены мелкие шумы. Степень сглаживания и размер удаляемых помех зависит от выбора размера окна, операций эрозии и масштабного преобразования. Эксперименты показали, что размер окна должен составлять около 10 % от среднего размера объекта.

Распознавание

Для распознавания северных оленей на аэрофотоснимках были выбраны признаковые методы. Эти методы позволяют решить поставленную задачу в условиях, когда животные на снимках находятся под разным освещением, в разных позах, имеют разный размер, цвет. Другими словами, объекты имеют множество эталонов, определить каждый из которых не представляется воз-



■ Рис. 7. Сглаживание изображения после сегментации: слева — до сглаживания; справа — после сглаживания

возможным, что является причиной отказа от корреляционных методов.

Было выделено 3 класса объектов: одиночные животные; животные, перекрывающие друг друга; прочие объекты. В качестве признаков выбраны форма сегментированной области, ее площадь, вытянутость.

Вытянутость области определяется двумя параметрами: протяженностью области по осям X и Y . Вычисляются эти параметры путем нахождения разности между максимальными значениями координат, принадлежащих области, и минимальными. Площадь соответствует количеству пикселей в области. О форме области можно судить по такому параметру, как округлость, которая определяется соотношением

$$c = \frac{p^2}{S},$$

где p — периметр области; S — площадь области.

Округлость области является безразмерной величиной. Если область является окружностью, то тогда округлость принимает минимальное значение $4\pi = 12,57$. Для квадрата это значение равно 16. Как правило, округлость стремится к большим значениям для вытянутых объектов.

Затем были определены пороговые параметры для каждого класса статистическим методом. Некоторые параметры, такие как округлость сегментированной области, имеют постоянное пороговое значение на всех изображениях, другие, такие как площадь, требуют адаптивного подхода на каждом изображении в связи с различным масштабом объектов. Для объекта, находящегося в отдалении от других, параметр округлости лежит в диапазоне от 14 до 35. Округлость области сегментированного изображения, соответствующая скоплению животных, лежит в диапазоне от 35 до 300. Области, имеющие значение округло-

сти более 300, как правило, являются крупными помехами. Пороговые значения других параметров, соответствующие определенному классу, можно вычислить при обработке фотографий дискретного масштаба. Достичь этого можно, производя фотосъемку с определенной высоты. Например, значение площади области в пикселях, соответствующее одному животному, лежит в диапазоне от 75 до 200 при условии, что съемка была произведена на высоте 500 м. Группировка, состоящая из нескольких животных, может иметь размер на сегментированном изображении до 1600 пикселей. Области площадью более 2000 пикселей, как правило, являются крупными помехами.

Заключение

Анализ методов сегментации и распознавания объектов подтвердил принципиальную возможность автоматического распознавания и подсчета диких северных оленей на фоне летней тундры по реальным аэрофотоснимкам. Примененные методы показали весьма хорошие результаты при подсчете изображений животных на «простых» снимках (зеленый фон, отсутствие помех).

Однако были выявлены и недостатки методов, требующие доработки. При смене цветового баланса снимка результаты сегментации могут оказаться неудовлетворительными. Кроме того, темные участки на снимке (овраги, ущелья и т. п.) при выбранном методе сегментации идентифицируются как объекты, их надо выявлять на этапе шумоподавления и фильтрации. Для распознавания и подсчета объектов в скоплениях кроме геометрических характеристик должны быть использованы цветовые особенности окраски животных. Решению этих задач будет посвящена дальнейшая работа.

Литература

1. Ерош И. Л., Сергеев М. Б., Соловьев Н. В. Обработка и распознавание изображений в системах превентивной безопасности: учеб. пособие / СПбГУАП. — СПб., 2005. — 154 с.
2. Гонсалес Р., Вуде Р. Цифровая обработка изображений. — М.: Техносфера, 2005. — 1072 с.
3. Фурман Я. А., Юрьев А. Н., Яншин В. В. Цифровые методы обработки и распознавания бинарных изображений / КрасГУ. — Красноярск, 1992. — 248 с.
4. http://www.rusnauka.com/23_D_2009/Informatica/49967.doc.htm (дата обращения: 15.10.2010).
5. Яне Б. Цифровая обработка изображений. — М.: Техносфера, 2007. — 584 с.
6. http://www.ci.ru/inform06_06/p_24.htm (дата обращения: 17.10.2010).
7. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. — М.: Вильямс, 2004. — 928 с.
8. <http://cgm.computergraphics.ru/content/view/147> (дата обращения: 21.10.2010).
9. Михайлов В. В., Карташев Н. К. DEER COUNTER — программа-тренажер для выработки навыка визуальной оценки количества животных в группировке // Биологические ресурсы Крайнего Севера: перспективы охраны и рационального использования. — СПб.: ГУАП, 2010. С. 205–212.