

Корреляционный алгоритм распознавания символьной информации в телевизионных изображениях

Д. Григорович

ЗАО «Московский научно-исследовательский телевизионный институт», Москва, Россия
greendm@mail.ru

Аннотация

Автоматизация распознавания и анализа полетных данных, представленных в виде символьной информации на телевизионных изображениях, является актуальной на данный момент для ряда задач управления космическими аппаратами.

В данной работе рассматриваются вопросы выделения, обработки и распознавания символьной информации, передаваемой в телевизионном формате, предлагается корреляционный алгоритм ее распознавания с использованием бинарных шаблонов и дается описание его работы, приводятся результаты экспериментов на реальных изображениях.

Ключевые слова: распознавание символьной информации, телевизионные изображения, корреляция.

1. ВВЕДЕНИЕ

При решении специальных задач мониторинга и управления космическими аппаратами (КА) осуществляют совместную передачу телевизионного (ТВ) сигнала с изображениями КА и телеметрической информации (ТМИ). Такая ТМИ представляется в виде символьной информации, которая характеризует отдельные параметры движения КА, и отображается на экране видеоконтрольных устройств для их анализа операторами [3], [4].

Требуется обеспечить высокую достоверность анализа поступающей информации для принятия соответствующего правильного решения по управлению процессом движения КА в реальном масштабе времени. Также необходимо проводить оперативные мероприятия послеполетного анализа, устранение и анализ возможных внештатных ситуаций. Вследствие быстрой смены символьной информации одновременно по нескольким параметрам движения КА на оператора возлагается обработка значительного объема данных, представленных в ТВ изображениях. В связи с этим возникает вопрос автоматизации процессов обработки символьной информации, и ее сохранения в удобной форме для последующего чтения и проведения анализа, т.е. в виде графиков, таблиц и т.д.

В данной статье рассматриваются вопросы выделения, обработки и распознавания символьной информации, передаваемой в ТВ формате, предлагается корреляционный алгоритм ее распознавания с использованием бинарных шаблонов и дается описание его работы.

2. ПОСТАНОВКА ЗАДАЧИ ИССЛЕДОВАНИЙ

Решение описанной задачи можно осуществить с помощью поиска и распознавания на ТВ изображениях фрагментов, соответствующих одному из множества шаблонов [2]. Общеизвестным подходом к решению подобных задач является корреляционное совмещение, которое сводится к

поиску максимума двумерной функции корреляции анализируемого (I) и эталонного (T) изображений [1], [5], [8]. Для получения корреляционной матрицы используется кросскорреляционная функция. Данная функция, записанная в терминах преобразования Фурье, будет иметь вид:

$$R(x, y) = \frac{J^{-1} [J[WT - WT] J^* [WI - WI]]}{\sqrt{\sigma^2(WT) \sigma^2(WI)}},$$

где J и J^{-1} – прямое и обратное преобразование Фурье, соответственно; $W = W(x, y)$ – весовая функция; σ^2 – дисперсия значений интенсивности внутри расчетной области. При вычислении преобразования Фурье используется алгоритм быстрого преобразования Фурье (БПФ), который дает преимущества во времени обработки перед прямым расчетом корреляционной функции.

Если размер изображения $W \times H$ и размер шаблона $w \times h$, то размер корреляционной матрицы $(W - w + 1) \times (H - h + 1)$. Оптимальный алгоритм сравнения, обеспечивающий наибольшую вероятность правильного распознавания, заключается в вычислении нормализованной корреляционной матрицы [9], причем $x' = 0 \dots w - 1$, $y' = 0 \dots h - 1$:

$$R(x, y) = \frac{\sum_{x', y'} (T'(x', y') \cdot I'(x + x', y + y'))}{\sqrt{\sum_{x', y'} T'^2(x', y') \cdot \sum_{x', y'} I'^2(x + x', y + y')}} ,$$

где $T'(x', y') = T(x', y') - 1 / (w \cdot h) \cdot \sum_{x'', y''} T(x'', y'')$,

$$I'(x + x', y + y') = I(x + x', y + y') - 1 / (w \cdot h) \cdot \sum_{x'', y''} I(x'', y'').$$

Нормализованная функция корреляции автоматически устраняет яркостные различия двух изображений. Идеальное совпадение анализируемого и шаблонного изображений будет при значении корреляции 1, полное различие изображений соответствует -1, значение 0 означает отсутствие корреляции между изображениями.

Далее после вычисления корреляционной функции производится поиск максимального значения, отвечающего наиболее вероятному смещению в области. Так как изображение представляется в виде конечного числа элементов, найденное положение максимума будет иметь целочисленные координаты в пикселях.

3. ИСХОДНЫЕ ДАННЫЕ

Исходные ТВ изображения имеют формат 720×576 пикселей (элементов), и представляются в виде 8-ми битных полутоновых изображений с 256-ю градациями яркости. Отдельное изображение является кадром, т. к. фоновая составляющая получается с внешней по отношению к КА ТВ

камеры. ТМИ в ТВ изображениях представлена в виде символьной информации в 52-х прямоугольных областях, называемых знакоместами. Указанные области имеют фиксированное расположение в ТВ кадре. Размерность знакоместа, состоящего из пяти символов, составляет 100×35 пикселей. Необходимо распознавать цифры, русские, некоторые английские и греческие буквы, и набор других символов. Символьная информация представлена в кадре определенным шрифтом на произвольном фоне КА. Отдельный символ занимает 14×23 пикселей.

При передаче символьной информации может ухудшаться ее качество за счет того, что:

- ТВ изображения могут подвергаться различным случайным искажениям, маскирующим характерное отличие символьной информации;
- ТВ сигнал может иметь малое значение отношения сигнал-шум;
- может изменяться фоновая составляющая ТВ изображения в местах расположения символьной информации и др.

Распознаваемые фрагменты ТВ изображений приведены на рис. 1.

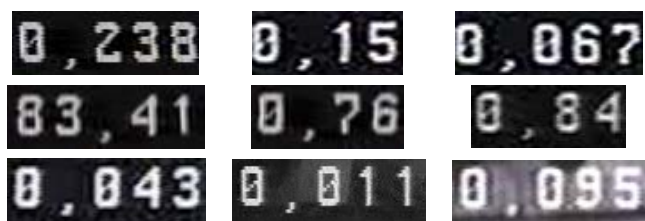


Рис. 1. Реальные фрагменты ТВ изображений с символьной информацией

Исходные данные о каждом символе должны быть представлены в отдельном файле, структура и формат которого определяется алгоритмом распознавания. Для каждого параметра движения КА задан диапазон допустимых значений.

4. ОПИСАНИЕ АЛГОРИТМА

В разработанном алгоритме используются специально подготовленные матрицы бинарных шаблонов символов, встречающихся в ТВ изображениях. Использование подобных эталонов позволяет повысить быстродействие и снизить влияние шумов.

Образцы символьной информации в виде бинарных шаблонов цифр представлены на рис. 2.



Рис. 2. Образцы бинарных шаблонов цифр

Корреляционный алгоритм распознавания символьной информации, представленной в ТВ изображениях, с использованием бинарных шаблонов (см. Алгоритм 1) включает несколько этапов обработки информации и работает следующим образом.

Вначале осуществляется получение ТВ изображения. Фиксированная структура ТВ кадра значительно облегчает и ускоряет сегментацию областей с отдельными параметрами движения КА. Координаты верхней левой вершины и размеры областей содержатся в априорных данных. С

помощью этой информации полученный кадр разбивается на отдельные интересующие области (ROI – Region of Interest).

Далее производится пороговое преобразование. На этом этапе обработки информации происходит преобразование полутонового изображения в бинарное изображение (рис. 3) с помощью порогового преобразования с фиксированной величиной порога.



Рис. 3. Бинарные фрагменты ТВ изображений с символьной информацией

Корреляционный алгоритм распознавания символьной информации в ТВ изображениях с использованием бинарных шаблонов

Начало

1. **Цикл:** Для каждого полученного изображения
2. **Цикл:** Для каждой интересующей области
3. Пороговое преобразование
4. Поиск контуров на изображении
5. Аппроксимация контуров полигонами
6. **Цикл:** Для каждого найденного контура
7. Создание минимального ограничивающего прямоугольника
8. Сегментация символа
9. Инвертирование матрицы символа
10. Нормализация символа к размеру шаблона
11. **Цикл:** Для каждого шаблона
12. Вычисление корреляционной функции между анализируемым символом и шаблоном
13. Определение максимального значения корреляционной функции
14. **Конец цикла**
15. Определение максимума из найденных максимальных значений корреляционных функций. Максимум соответствует лучшему совпадению анализируемого символа с шаблоном
16. **Конец цикла**
17. Формирование результата из распознанных символов
18. Анализ полученного результата на основе априорных данных
19. Сохранение корректного результата в базе данных
20. Построение графика и обновление таблицы параметра ТМИ
21. **Конец цикла**
22. **Конец цикла**

Конец алгоритма

Алгоритм 1. Корреляционное распознавание символьной информации

Одной из важных задач обработки и анализа ТВ изображений является сегментация, т.е. разделение его на области, для которых выполняется определенный критерий однородности, в данном случае, выделение на изображении областей приблизительно одинаковой яркости. Один из основных, простых и надежных при малом разрешении исходных изображений способов – это построение сегментации с помощью порогового преобразования. Порог – это признак, позволяющий разделить исходное ТВ изображение на два уровня яркости, каждый из которых будет отображать свой

класс объектов. Яркость пиксела бинарного изображения принимает два значения, черный и белый цвета. При уменьшении числа значений яркости исчезают слабоконтрастные детали, плавные переходы тона становятся более резкими.

Следующим этапом обработки ТВ изображения является поиск контуров символов на изображении знакоместа. Под контуром обычно понимается местоположение локального изменения или резкого перепада яркости на изображении. При этом подразумевается, что такие перепады возникают на границах объектов. Для кодирования контура используется цепной код Фримена [7], с помощью которого представляются границы в виде последовательности отрезков прямых линий определенной длины и определенного направления. Суть цепного кода заключается в том, что для любой связанной линии на изображении кодируются координаты начального пиксела, а для каждого последующего пиксела цепочки в качестве кода используется его приращение, определяющее переход на один из смежных пикселей. Поскольку таких смежных пикселей всего 8, то для кодирования каждого пикселя необходимо 3 бита информации. В данном алгоритме выполняется поиск только крайних внешних контуров символов (рис. 4).



Рис. 4. Контуров символов, найденные во фрагментах ТВ изображений

Далее найденные контуры аппроксимируются полигонами методом Дугласа-Пекера [6]. Идея обработки по указанному методу состоит в построении по контуру полигона с меньшим числом точек. Алгоритму задается желаемая точность аппроксимации, которая определяет максимальное расстояние между исходным контуром и полигоном.

Затем создается минимальный прямоугольник, ограничивающий контур, и осуществляется сегментация символа путем извлечения из изображения матрицы, соответствующей ограничивающему прямоугольнику [9] (рис. 5). Каждый бит матрицы символа инвертируется (рис. 6). Дополнительно производится нормализация символа, т.е. масштабирование каждого символа до определенного размера, для сравнения с шаблонами символов, хранящихся в файлах. Вопрос масштабирования требует отдельных исследований для определения оптимального масштаба шаблонов символов, при котором точность распознавания будет максимальной. При увеличении масштаба символов, с одной стороны, усиливаются их характерные отличия друг от друга, с другой стороны, увеличиваются искажения и дефекты символьной информации, подвергаемой распознаванию. Для этой задачи был взят размер шаблона символа 14×23 пикселей. При этом анализируемый символ масштабируется до того же размера.



Рис. 5. Фрагменты ТВ изображений с ограничивающими символами прямоугольниками



Рис. 6. Фрагменты инвертированных ТВ изображений с символьной информацией

Следующим этапом данного алгоритма является вычисление корреляционной матрицы для изображений символа и каждого шаблона в отдельности. Для каждой корреляционной функции определяется максимум, соответствующий наиболее вероятным координатам смещения в области анализируемого изображения. Затем определяется лучшее совпадение по максимуму из найденных максимальных значений корреляционной функции для каждого шаблона. Если размеры распознаваемого изображения I и шаблона T равны, то результирующая матрица состоит из одного элемента. В данном алгоритме за счет выделения каждого отдельного анализируемого символа с помощью минимального ограничивающего прямоугольника значительно сокращается объем вычислений, поскольку размер символа приводится к размеру шаблона, и уменьшается область поиска.

Из распознанных символов формируется результат в виде строки. Результат распознавания анализируется, является ли он корректным числом или строкой текста на основе априорных данных возможных значений параметра ТМИ.

В результате выполнения указанных операций по обработке ТВ изображений в каждом знакоместе осуществляется распознавание символьной информации, содержащей полетные данные КА. Распознанная информация сохраняется в базе данных. Впоследствии полученные данные используются для построения графиков и таблиц, характеризующих параметры движения КА за определенный отрезок времени.

5. ЭКСПЕРИМЕНТАЛЬНЫЕ РЕЗУЛЬТАТЫ

Для реализации указанного алгоритма обработки и распознавания ТВ изображений было написано программное обеспечение (ПО) с графическим интерфейсом на языке C++. Для разработки программы использовалась кросс-платформенная библиотека Qt [10]. Ряд этапов данного алгоритма был реализован с применением библиотеки алгоритмов компьютерного зрения, обработкой изображений и численных алгоритмов общего назначения с открытым кодом OpenCV [9]. Настройки графического интерфейса, априорная информация, и вспомогательные данные процесса распознавания ТВ изображений хранятся в XML-файлах, и считываются программой при ее загрузке. Корректные результаты распознавания символьной информации, содержащейся во всех подвергаемых распознаванию знакоместах ТВ изображений, записываются в базу данных SQL в реальном времени. Графики параметров движения КА строятся и отображаются с использованием набора Qt-виджетов и вспомогательных классов Qwt [11] в процессе работы ПО. В программе реализована многопоточная обработка всех знакомест, что позволяет распознавать большое количество символьной информации параллельно.

Проверка корреляционного алгоритма осуществлялась на компьютере с процессором Intel(R) Core(TM) i5-2410M 2.30 ГГц, оперативной памятью DDR3 800 МГц, под управлением 64-битной операционной системы Windows 7.

Результаты экспериментальных исследований показали, что время распознавания одного знакоместа в ТВ изображении, содержащего пять символов составило 10 мс. Было проведено распознавание 300 тестовых ТВ изображений с символьной информацией. Точность распознавания символьной информации в ТВ изображениях составила 85%.

6. ЗАКЛЮЧЕНИЕ

Для распознавания символьной информации в ТВ изображениях с использованием бинарных шаблонов разработан корреляционный алгоритм обработки ТВ изображений и написано ПО. Сегментация символов осуществляется с помощью контуров, что позволяет значительно сократить размер корреляционной матрицы и уменьшить время поиска лучшего совпадения с шаблоном. Для определения степени близости входного изображения и шаблонов используется вычисление кросс-корреляции с помощью алгоритма БПФ, при этом уменьшается общая вычислительная сложность алгоритма. Недостатком корреляционной меры сходства является ее чувствительность к геометрическим искажениям видимых размеров символов при изменении ракурса съемки.

Алгоритм способен распознавать символьную информацию в ТВ изображениях с малым значением отношения сигнал-шум, невысоким разрешением и работать в реальном масштабе времени с параллельной обработкой до 100 исходных изображений ТМИ.

По результатам экспериментальных исследований корреляционный алгоритм распознавания символов с использованием бинарных шаблонов обладает достаточно высокой точностью распознавания и высоким быстродействием, что делает его применение оправданным при распознавании символьной информации в ТВ изображениях. Дальнейшее улучшение точности распознавания связано с применением более гибкого порогового преобразования, а также использованием дополнительных признаков, обеспечивающих достоверную селекцию символьной информации на произвольном фоне.

7. ССЫЛКИ

- [1] Гонсалес Р. Цифровая обработка изображений. – М., 2005. – 1072 с.
- [2] Григорович Д.В. Распознавание символьной информации в телевизионных изображениях с использованием бинарных шаблонов. Сб. докл. XX международной научно-технической конференции «Радиолокация, навигация, связь» (RLNC-2014). – Т. 1. – С. 159 – 164, Воронеж, Россия, апрель 2014.
- [3] Григорович Д.В., Смирнов А.И. Оперативная селекция телеметрической информации в телевизионных изображениях. Сб. мат. XI Международной научно-технической конференции «Распознавание-2013». – Курск, Россия, сентябрь 2013. – С. 382 – 384
- [4] Кравец В.Г., Любинский В.Е. Основы управления космическими полетами. – М., 1983. – 224 с.
- [5] Сойфер В.А. Методы компьютерной обработки изображений. – М., 2001. – 784 с.
- [6] Douglas D.H., Peucker T.K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, Journal «The Canadian Cartographer», vol. 10, 1973. – No 2. – P. 112 – 122.
- [7] Freeman H. On the encoding of arbitrary geometric configurations, IEEE Electronic Computers, IRE Transactions, vol. 10, 1961. – No 2. – P. 260 – 268.
- [8] Nikhil Gupta, Rahul Gupta, Amardeep Singh, Matt Wytock. Object Recognition using Template Matching, December 2008. <http://cs229.stanford.edu/proj2008/GuptaGuptaSinghWytock-ObjectRecognitionUsingTemplateMatching.pdf>.
- [9] OpenCV API Reference. <http://docs.opencv.org/modules/refman.html>

[10] Qt Project Documentation. <http://qt-project.org/doc/>.

[11] Qwt User's Guide. <http://qwt.sourceforge.net/>.

Об авторе

Григорович Дмитрий – начальник сектора 333 ЗАО «МНИТИ».

E-mail: greendm@mail.ru.

Abstract

At present an automation of recognition and analysis of the flight data as the character information in TV images is an urgent issue for some control tasks of space vehicles. This article is devoted to the character information extraction, processing and recognition in TV format.

The correlation algorithm of character information recognition with use of binary patterns is considered, its operation description is performed, and the experimental results with real images are described.

Keywords: *character information recognition, TV images, correlation.*