

УДК 629.391

БАРАННИК В.В., д.т.н., профессор (ХУВС им. И. Кожедуба),
 ШИРЯЕВ А.В., к.т.н., доцент кафедры «Сети связи» (ХНУРЭ),
 КОРОЛЁВА Н.А., к.т.н., доцент (УкрГАЗТ)

Анализ существующих форматов видеоданных для семантической обработки

В статье рассматриваются вопросы, связанные с семантической обработкой видеоданных. Проанализированы достоинства и недостатки существующих форматов видеоданных в соответствии с их возможностями по семантической обработке и анализу. Вводятся понятия: наборы примитивов, функции построения примитивов, свойства формата для семантического анализа. Предложено развивать методы выделения и кодирования отдельного семантического слоя базовых информационных признаков видеоданных, адаптированных для семантического анализа в векторном формате представления.

Ключевые слова: видеоданные, методы, правила, параметры, формат, растр, вектор, базовые признаки, семантика, кодирование.

В настоящее время актуальными являются вопросы анализа и семантической обработки видеоданных для локализации и детектирования их смысловой составляющей. Такая обработка используется для различных сфер, включая семантический поиск видеоданных на основе содержимого, картографирование, семантическое маскирование видеоданных локализация и идентификация сцен, компенсация движения семантически значимых объектов.

В общем случае видеоданные представляются **форматами**, базирующиеся на многослойной иерархической системе, основанной на:

а) совокупности методов $(\Phi\{F(i, j), P(i, j), d(i, j)\})$, где $F(i, j)$ - методы получения ключевых объектов видеоданных, $P(i, j)$ - методы выделения базовых признаков, $d(i, j)$ - методы, определяющие соотношения между базовыми признаками);

б) правил $(E\{\phi(i, j), \phi(i, j)\})$, где $\phi(i, j)$ - правила обхода границ, $\phi(i, j)$ - правила преобразований растра);

в) способов и параметров $(G\{x_{i,j}, y_{i,j}, d_x, d_y, k_{res}\})$, где $x_{i,j}$ - входной и $y_{i,j}$ - выходной набор растровых данных, d_x и d_y - приращения соответственно по направляющим x и y , k_{res} - масштабирующий коэффициент) представления видеокадров в цифровом виде.

Форматы видеоданных в зависимости от вида представления и хранения графических данных делятся на **растровые, векторные и комплексные** (рис. 1).

Растровые форматы – базируются на способе представления видеоданных в виде матрицы пикселей заданного фиксированного размера со своими геометрическими параметрами. Основным отличием каждого из растровых форматов от других является то, что значения пикселей обрабатываются и сохраняются в виде отдельных битовых последовательностей, которые при восстановлении возвращаются на прежние свои позиции. Насчитывают более 20 различных способов представления растровых данных (BMP, JPG, PNG, TIF, GIF);

Для компактного представления видеоданных большинства растровых форматов используют методы сжатия, базирующиеся на алгоритмах LZW или RLE. Особенностью данных алгоритмов является то, что для сокращения числа бит используются некоторое подобие векторного описания кодовой последовательности сжатых данных. Поэтому можно говорить, что алгоритмы сжатия LZW и RLE обладают возможностью компактно сохранять растровые видеоданные в векторной форме при определенных ограничениях.

Векторные форматы - способ представления объектов и видеоданных, основанный на использовании элементарных геометрических наборов объектов (примитивов), которые используют совокупности простых и сложных математических функций, таких как: точки, линии, сплайны и многоугольники дуги, окружности, треугольники, прямоугольники и т.д. С помощью данных наборов функции построения примитивов осуществляют

© В.В. Баранник, А.В. Ширяев, Н.А. Королёва, 2014

аппроксимацию исходных видеоданных. Функции построения примитивов выполняют простейшие операции масштабирования и поворота примитивов для более точной аппроксимации.

Насчитывают несколько форматов представления векторных данных (SVG, EPS, EMF, POV, DXF, SK1, WMF и т.д.);

Комплексные форматы содержат в своей структуре элементы представления растровых и векторных форматов, которые сочетаются друг с другом посредством интегрального подхода к описанию объектов. Существует несколько форматов представления (PDF, DJVU, CGM и т.д.).

Таблица 1

Сравнение свойств форматов растровых и векторных видеоданных

Свойства формата	Растровая графика	Векторная графика
Способ представления видеоданных	Строится из множества пикселей	Описывается в виде последовательности команд
Реалистичность видеоданных	Эффективно используются для представления реальных образов	Нельзя получить видеоданные фотографического качества
Качество редактирования видеоданных	При масштабировании и вращении изображения возникают искажения	При масштабировании и вращении видеоданных не происходит потери качества
Особенности печати видеоданных	Изображения могут быть легко распечатаны на принтерах	В некоторых случаях печать невозможна или распечатанные видеоданные имеют существенные отличия с электронными копиями

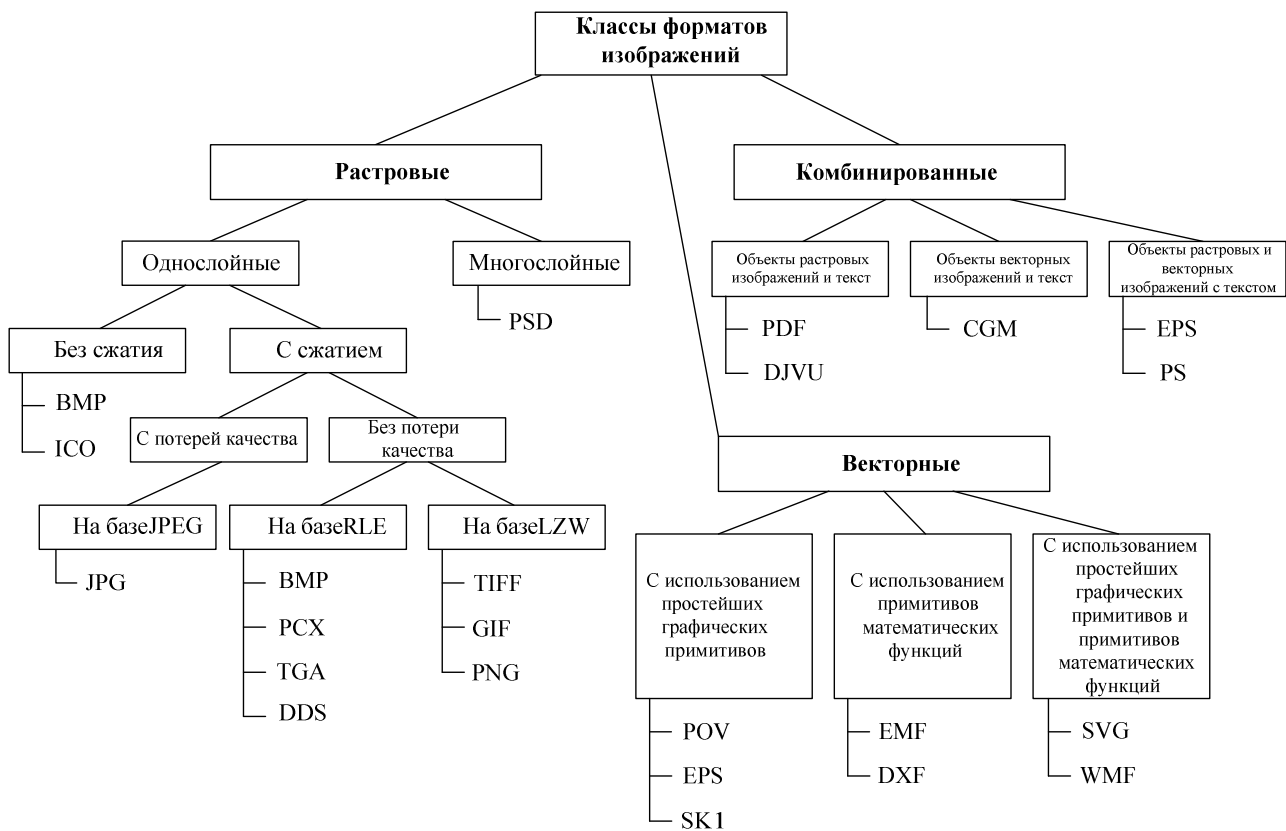


Рис. 1. Классы форматов видеоданных

Под способностью формата для семантического анализа подразумевается такая форма представления видеоданных, в которой семантическая основа выявляется и сохраняется в виде отдельного слоя многослойного формата.

При семантическом анализе ключевыми аспектами являются:

1) достоверность аппроксимации семантики видеоданных ($PSNR(\bar{A}, \bar{B})$)

$$PSNR(\bar{A}, \bar{B}) = 10 \log \left(\frac{2^b \cdot M \cdot N}{\sum_{x=0}^M \sum_{y=0}^N (A_{x,y} - B_{x,y})^2} \right),$$

где \bar{A} - вектор исходного набора семантически видеоданных; \bar{B} - вектор аппроксимированного набора семантически видеоданных; b - глубина вложенности видеоданных; M - число элементов видеоданных по горизонтали; N - число элементов видеоданных по вертикали;

2) сложность обработки ($M(N_+, N_-, N_{\times}, N_{\div}, N_{\otimes}, N_{rot}, N_{res})$,

где N_+ и N_- - число операций суммирования/вычитания, N_{\times} и N_{\div} - число операций умножения/деления, N_{\otimes} - число операций сдвига, N_{rot} и N_{res} - число операций поворота и масштабирования);

3) компактность представления кода ($N(W_{in}, W_{out})$), которая оценивается как

$$N(W_{in}, W_{out}) = \frac{W_{in}}{W_{out}},$$

где W_{in} - объем видеоданных до их компактного представления; W_{out} - объем видеоданных после компактного представления;

4) масштабируемость ($Z(N_{rot}, N_{res})$).

Существующие подходы к реализации семантического анализа базируются на выделении из общего объема видеоданных необходимого набора базовых признаков, достаточных для семантического анализа, которые в дальнейшем описываются с помощью преобразований блокового типа растрового и векторного представления базовых признаков.

Блочные преобразования растрового представления для получения базовых признаков - при блоковых преобразованиях исходные значения растровых видеоданных разбиваются на блоки размера, равному окну фильтрующей или скользящей функции, которая выполняет операции над растровыми видеоданными.

При операциях с блоковыми преобразованиями над растровым представлением базовых признаков видеоданных наблюдаются высокая сложность масштабирования и поворота. При этом присутствует высокий уровень пропуска и огрубления результатов при операциях семантического и структурного анализа для классов слабо и средне насыщенных видеоданных. Однако, следует отметить, что для класса насыщенных видеоданных операции с преобразованиями над растровым представлением обеспечивают более высокий процент соответствий и меньший процент пропуска и огрубления результатов [3].

Достоинства блоковых преобразований растрового представления:

- семантические признаки получаются при непосредственной работе с растром видеоданных;
- семантические признаки формируются для блоковых растровых наборов, что позволяет проводить анализ соответствия семантики блоковому содержанию раstra;
- отсутствие различия скорости обработки от типа насыщенности видеоданных;
- высокая достоверность данных, которая достигается за счет представления семантики непосредственно из блокового набора раstra видеоданных;
- симметричность обработки;

Недостатки блоковых преобразований растрового представления:

- низкая скорость обработки слабо- и средне-насыщенных видеоданных за счет применения однотипного, не ориентированного на класс насыщенности видеоданных математического аппарата;
- масштабируемость видеоданных и его семантических признаков осуществляется с искажениями, характерными для растровых данных;
- нет математического аппарата для описания компактного представления семантических данных;

Векторное представление базовых признаков - базовые признаки изображений аппроксимируются набором различных математических функций с заданной погрешностью.

Векторное представление базовых признаков [2] обладает большей гибкостью для различного рода преобразованиями. Это позволяет значительно сократить время выполнения преобразований и время семантического анализа для классов слабо- и средне-насыщенных видеоданных. Для класса сильно

насыщенных изображений процесс векторизации является трудоемким. Поэтому существенная доля ресурсов процессорного времени затрачивается на выполнение всех видов операций.

Достоинства векторного представления базовых признаков:

- высокая степень соответствия векторизируемых признаков семантике видеоданных за счет характерного выбора базовых признаков;

- за счет использования малого числа примитивов для аппроксимации слабо- и средненасыщенных видеоданных достигаются: высокая скорость обработки; высокая достоверность данных; высокая скорость масштабирования; компактное представление;

- симметричность обработки для операций преобразований растровое изображение - векторное изображение и обратно.

Недостатки векторного представления базовых признаков формируются за счет использования большого числа примитивов для аппроксимации сильно-насыщенных мелкими объектами видеоданных: существенно низкая скорость обработки видеоданных; существенно низкая достоверность данных; существенно низкая скорость масштабирования; отсутствие компактного представления.

Выводы

1. Блочные преобразования и векторное представление базовых признаков имеют ряд недостатков для семантического анализа;

- низкая скорость получения набора базовых признаков видеоданных;

- низкая скорость масштабирования видеоданных при обеспечении заданной достоверности;

- отсутствует математический аппарат для описания компактного представления семантических данных.

2. Описаны основные ключевые параметры, отвечающие за основные характеристики семантического анализа такие как: время получения базовых признаков; сложность обработки видеоданных; достоверность аппроксимации семантики видеоданных; компактность представления кода; масштабируемость.

3. Предложено развивать методы формирования и кодирования отдельного семантического слоя базовых информационных признаков для создания нового формата видеоданных, адаптированного для семантического анализа в компактной форме.

Литература

1. Lalonde M., Gagnony L., Boucherz M. - C. Non-recursive paired tracking for vessel extraction from retinal images. - Centre de recherche informatique de Montreal, 2000.
2. Дуда Р., Харт П. Распознавание образов и анализ сцен. - М.: ДИАЛОГ-МИФИ, 2001. - 521 с.
3. Шикин А.В., Боресков А.В. Компьютерная графика. Полигональные модели. - М.: ДИАЛОГ-МИФИ, 2001. - 464 с
4. Баранник В.В. Метод квадратурного сжатия трансформант вейвлет-преобразования в двумерном полиадическом пространстве / В.В. Баранник, А.В. Ширяев // Сучасна спеціальна техніка, К.: №2(25). - 2011. - С. 73-80.

Баранник В.В., Ширяев А.В., Корольова Н.А. АНАЛІЗ ІСНУЮЧИХ ФОРМАТІВ ВІДЕОДАНИХ ДЛЯ СЕМАНТИЧНОЇ ОБРОБКИ. У статті розглядаються питання, пов'язані з семантичною обробкою відеоданих. Проаналізовано переваги та недоліки існуючих форматів відеоданих відповідно до їх можливостями по семантичній обробці та аналізу. Вводяться поняття: набори примітивів, функції побудови примітивів, властивості формату для семантичного аналізу. Запропоновано розвивати методи виділення та кодування окремого семантичного шару базових інформаційних ознак відеоданих, які адаптовані для семантичного аналізу у векторному форматі представлення.

Ключові слова: відеодані, методи, правила, параметри, формат, растр, вектор, базові ознаки, семантика, кодування.

Barannik V.V., Shiryaev A.V., Korolyova N.A. ANALYSIS OF EXISTING FORMATS OF VIDEO DATA FOR SEMANTIC TREATMENT. The article discusses issues connected with semantic processing of video data. The advantages and disadvantages of existing video data formats in accordance with their capabilities on semantic processing and analysis have been analyzed. The concepts such as: a set of primitives, a function of primitive modeling, format properties for semantic analysis are introduced. It has been proposed to develop methods for the selection and encoding of semantic layer of basic information video data features adapted for semantic analysis in vector format of presentation.

Key words: video data, methods, rules, parameters, format, raster, vector, basic features, semantic, coding.

Рецензент д.т.н., профессор Безрук В.М. (ХНУРЭ)

Поступила 31.03.2014г.