

Интеллектуальный формат сжатого представления цветных графических изображений

Л.И. Лебедев¹, М.С. Вирясов¹, В.П. Громов¹
lebedev@pmk.unn.ru|mix3micha@mail.ru|GVP1941@bk.ru
¹ННГУ им. Н.И. Лобачевского, Нижний Новгород, Россия

В работе описывается интеллектуальный формат CIFc сжатого представления цветных графических изображений (ГИ). Предлагаемый формат является расширением эффективного формата IFc, базирующегося на корреляционно-экстремальных контурных методах (КЭКМ) распознавания символно-текстовой информации. Технология получения описания в формате CIFc основана на декомпозиции исходного ГИ с выделением текстового слоя и цветового слоя. Текстовый слой представлен черно-белым изображением, полученным в результате бинаризации обесцвеченного ГИ, и является маской переднего плана. Цветовой слой включает задний план и цветовую информацию о маске переднего плана и сжимается с использованием вейвлет-преобразования. Сжатие текстового слоя после его векторизации осуществляется двухкритериальным алгоритмом КЭКМ в режиме самообучения. Восстановление изображения осуществляется в обратном порядке на основе реставрации каждого из слоев. Приводятся теоретическое обоснование эффективности предлагаемого формата CIFc в сравнении с наиболее близким аналогом DjVu и результаты экспериментов на модельных и реальных цветных ГИ.

Ключевые слова: цветное графическое изображение, формат представления изображения, сжатие, текстовый слой, эффективность представления, алгоритм распознавания.

An intelligent format for compressed representation of colour graphic images

L.I. Lebedev¹, M.S. Virjasov¹, V.P. Gromov¹
lebedev@pmk.unn.ru|mix3micha@mail.ru|GVP1941@bk.ru
¹Lobachevsky State University of Nizhni Novgorod, Nizhny Novgorod, Russia

The paper describes the intelligent format CIFc for compressed representation of color graphic images (GI). The proposed format is an extension of the effective IFc format, which is based on the correlation-extremal contour methods (CECM) of character-text information recognition. The technology for obtaining descriptions in CIFc format is based on the decomposition of the original GI with the selection of the text layer and the color layer. The text layer is represented by a black-and-white image obtained as a result of binarization of the bleached GI, and is a foreground mask. The color layer includes background and color information about the foreground mask and is compressed using wavelet transform. Compression of the text layer after its vectorization is carried out by the two-criteria KECM algorithm in the self-learning mode. Image restoration is carried out in reverse order on the basis of the restoration of each of the layers. A theoretical justification is given for the effectiveness of the representation of color GI in the proposed CIFc format in comparison with the closest analogue of DjVu. The theoretical justification of the effectiveness of the proposed CIFc format in comparison with the closest analogue of DjVu and the results of experiments on model and real color GI's are given.

Keywords: color graphic image, image format, compression, text layer, presentation efficiency, recognition algorithm.

1. Введение

Для сжатия графических изображений (ГИ) черно-белых документов был разработан алгоритм, основанный на распознавании объектов корреляционно-экстремальными контурными методами (КЭКМ), которые относятся к методам сравнения с эталоном. Идея эффективного сжатия состоит в замене метрического описания распознанных объектов на параметры, полученные при совмещении с эталоном, объем которых значительно меньше. Для хранения получаемых параметров распознавания была предложена структура постоянной длины, которая являлась основой описания предложенного формата IFc [6]. Для повышения качества распознавания кроме базового среднеквадратичного критерия сходства был дополнительно введен критерий на основе оценки расстояния Хаусдорфа [1]. Полученный двухкритериальный алгоритм обеспечивает качество распознавания с любой заданной достоверностью. Такая же идея использования результатов распознавания методом сравнения с прототипом была предложена Яном Лекуном в технологии получения формата DjVu [5, 7]. В форматах DjVu и IFc используются однотипные методы распознавания, основанные на сравнении объекта с эталоном, однако вычисляемые в

алгоритмах оценки сходства имеют разный уровень инвариантности. Методы распознавания в IFc инвариантны относительно ортогональных преобразований и масштабирования, а в формате DjVu только к смещению. Поэтому, класс эквивалентности, порожденный одним эталоном, в формате IFc имеет большую мощность. Отсюда, сжатие изображения с хаотической ориентацией объектов и вариабельностью их масштаба в формате IFc будет выше, чем в формате DjVu, в котором потребуется хранить большее количество прототипов. Когда количество используемых прототипов приблизительно равно, эффективней будет формат DjVu, так как в нем не предусмотрена передача параметров об ориентации и масштабе объекта. Однако, наибольшим преимуществом формата DjVu является возможность сжатия цветных изображений. В то же время наличие огромного количества цветных графических документов и необходимость их более эффективного сжатия делает крайне актуальным решение задачи обработки цветового слоя в формате IFc.

2. Постановка задачи

Требуется разработать технологию представления цветных графических изображений на базе формата IFc. На основе анализа требуемого алгоритмического и

программного обеспечения с учетом векторной модели описания объектов, используемой в формате IFC, обработка цветных ГИ должна включать следующие этапы:

1. Декомпозиция изображения с выделением цветового и текстового слоев.
2. Вейвлет- сжатие цветового слоя.
3. Получение описания текстового слоя в формате векторной (контурной) модели представления изображения.
4. Представление текстового слоя в формате IFC.
5. Восстановление цветового слоя.
6. Восстановление текстового слоя в формате контурного описания изображения.
7. Растривание контурного описания изображения.
8. Слияние восстановленных слоев.

3. Методы решения

Рассмотрим методы решения, применяемые на каждом из этапов предложенной технологии представления цветных ГИ в формате IFC.

Для выделения текстового слоя из цветного трёхканального RGB изображения сначала осуществляется преобразование его в одноканальное изображение в оттенках серого. Для этого здесь для каждого пикселя выполним стандартное преобразование с получением полутонового изображения:

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B .$$

Далее, воспользовавшись методом Отцу, находим пороговое значение для бинаризации полутонового изображения. Пороговое значение для разделения по яркости Y всех пикселей на два класса находится на основе максимизации межклассовой дисперсии [3]. Для описания бинарного изображения вначале строится штриховая модель, на основании которой формируется контурная модель текстового слоя. Наконец, используя формат IFC, на базе контурной модели изображения получаем сжатое представление текстового слоя.

Получение цветового слоя можно описать следующей процедурой. Для формирования цветового слоя на исходном изображении вырезаются фрагменты, соответствующие текстовому слою. При этом на месте текстового слоя осуществляется восстановление цвета близкого к цвету фона в его окрестности. Эта аппроксимация необходима для обеспечения лучшего сжатия цветового слоя вейвлетным алгоритмом. Сжатие цветового слоя осуществляется на базе вейвлет- преобразования Баттерворта для изображений и преобразования Добеши 9/7 [2]. На заключительной стадии для кодирования коэффициентов вейвлетного разложения изображения применяется алгоритм SPIHT в комбинации с арифметическим кодированием. Основная идея SPIHT заключается в том, чтобы переупорядочить биты вейвлетного разложения по принципу выделения низко- и высокочастотной информации с последующим подавлением последней для лучшего восприятия изображения.

Восстановление изображения осуществляется в обратном порядке. Наиболее алгоритмически сложным представляется восстановление текстового слоя в части получения по контурной модели растрового изображения. Отметим, что при слиянии слоев восстановленное изображение наследует параметры цветового слоя.

4. Программная реализация

Изложенные выше алгоритмы для сжатия и восстановления изображений реализованы на языке C++ платформы Microsoft Visual Studio 2017 с использованием средств MFC для обеспечения взаимодействия с

пользователем [4]. Программа также предусматривает консольный запуск с параметрами.

Для получения сжатого представления, а также восстановления цветовой составляющей изображения используется свободная библиотека TILib (Tiny Library). Она предназначена для вейвлетного сжатия изображений с потерями, реализует биортогональный вейвлет-фильтр Добеши 9/7 и биортогональный вейвлет-фильтр Баттерворта. Для кодирования коэффициентов вейвлетного преобразования используется алгоритм SPIHT.

Для получения сжатого представления, а также восстановления текстового слоя используются программные средства, разработанные авторами.

Для выделения текстового и цветового слоёв, а также для формирования и сохранения восстановленного изображения используются средства библиотеки компьютерного зрения с открытым исходным кодом OpenCV [8].

Сжатое представление изображения является совокупностью трёх файлов: цветовой слой (.Ti), текстовый слой (.bds + .etl). Для удобства и большей эффективности на финальном этапе сжатия полученные файлы объединяются в архив с помощью программы 7z.

5. Результаты экспериментов

Ниже приводятся результаты представления цветных изображений в форматах CIFc и DjVu. На рис.1 приведено исходное полноцветное изображение «Радуга цифр» размером 1320×1320 (41817600 байт). В формате tiff файл занимает 5620668 байт.

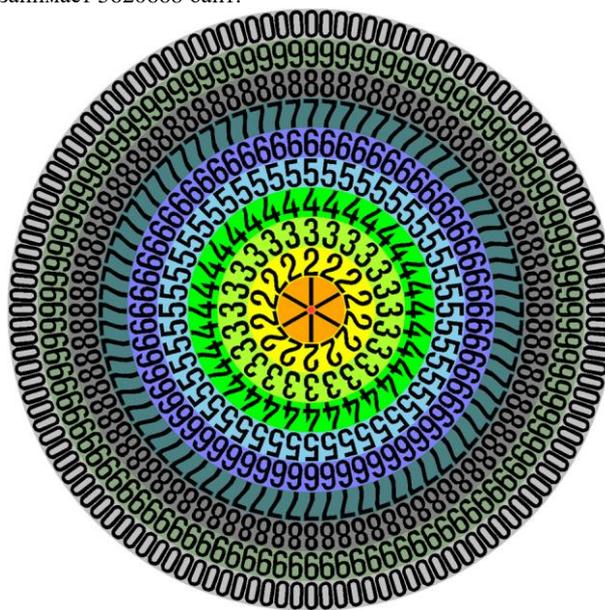


Рис. 1. Цветное изображение «Радуга цифр».

На рис. 2 приведено восстановленное изображение «Радуга цифр», полученное из формата CIFc. Цветовой слой на данном изображении сжат в 200 раз и представлен файлом Ti объемом 26136 байт. Текстовый слой в формате CIFc представлен двумя файлами: параметрического описания объектов bds (8414 байт) и эталонов etl (476 байт). Арифметическое кодирование этих трех файлов дает конечный объем файла в формате CIFc, равный 31124 байт. В формате DjVu цветное изображение «Радуга цифр» имеет размер 665885 байт, что более чем в 21 раз больше относительно формата CIFc.

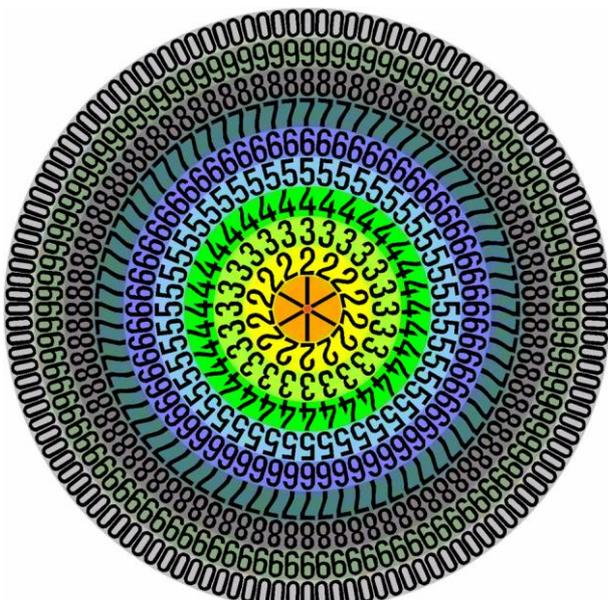


Рис. 2. Восстановленное изображение «Радуга цифр».

На рис. 3 представлено изображение P2, являющееся сканом страницы книги об условных знаках на топографических картах.

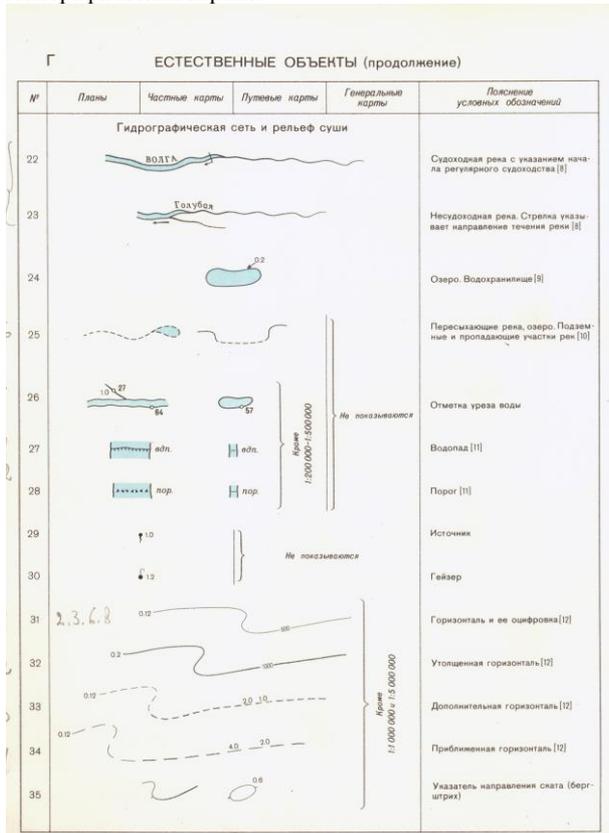


Рис. 3. Исходное изображение P2 в формате tif (62,8 МБ)

На рис. 4 представлена иллюстрация текстового слоя изображения P2, а на рис. 5 его цветовой слой. Описание изображения P2 в формате CIFs составило 34.8 Кб, а в формате DjVu 38.6 Кб. Отметим, что цветовой слой в формате CIFs был сжат в 3000 раз. На данном изображении P2 эффективность форматов в плане получаемых коэффициентов сжатия приблизительно одинакова. На рис. 6 представлено восстановленное изображение из формата CIFs, а на рис. 7 из формата DjVu.

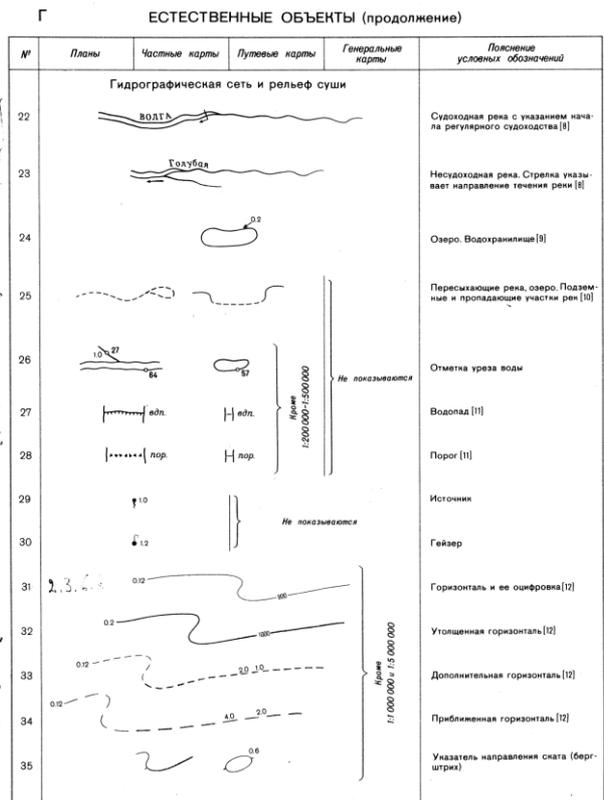


Рис. 4. Текстовый слой изображения P2

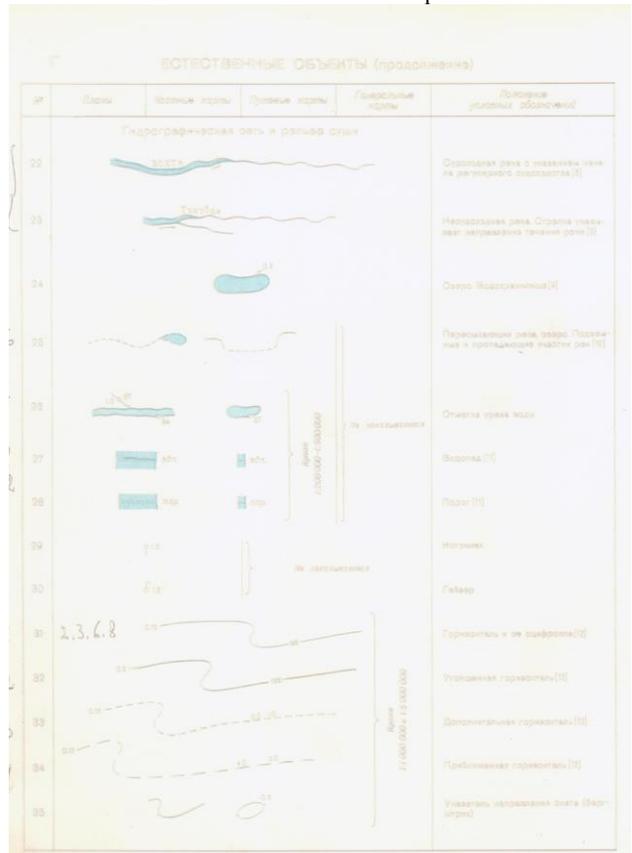


Рис. 5. Цветовой слой изображения P2

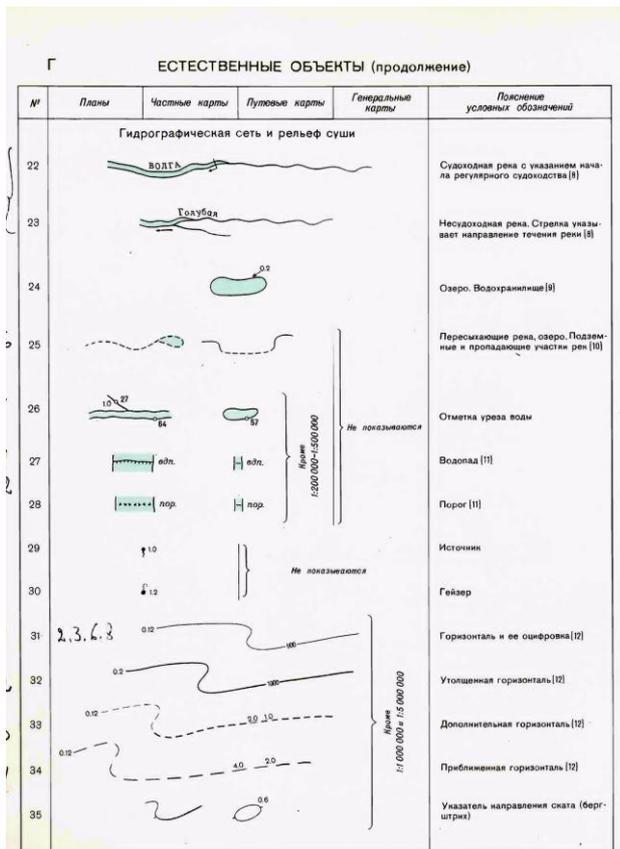


Рис. 6. Восстановленное изображение P2 из формата CIFc

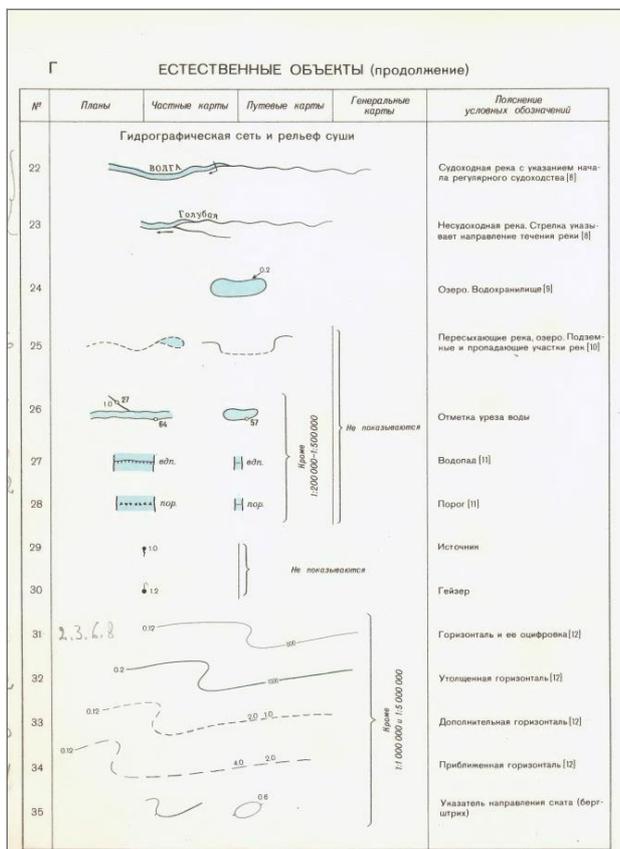


Рис. 7. Восстановленное изображение P2 из формата DjVu

6. Заключение

Разработана и реализована технология создания интеллектуального формата сжатого представления цветных графических изображений CIFc на базе формата IFc. Проведенные исследования и полученные результаты экспериментов показывают высокую эффективность сжатия изображений, представляемых в формате CIFc.

7. Благодарности

Работа выполнена при поддержке гранта РНФ № 16-11-00068-П и гранта РФФИ № 18-07-00715.

8. Литература

- [1] Лебедев Л.И. Двухкритериальный алгоритм распознавания объектов графических изображений на базе КЭКМ. / Л.И. Лебедев, Ю.Г. Васин // [ГРАФИКОН'2015](#): Труды Юбилейной 25-й Международной научной конференции. 2015. С. 112-114 (<http://elibrary.ru/item.asp?id=24366739>).
- [2] Симаков А.В. Прогрессивная передача изображений через Интернет / А.В. Симаков // https://www.compression.ru/download/articles/wavelets/simakov_2004_progressive_image_transmission_over_internet.pdf - (дата обращения 18.09.2019)
- [3] Метод Отсу, алгоритм и программная реализация // Википедия — свободная энциклопедия. – Режим доступа: <https://www.wikipedia.org>.
- [4] Хранилище документации Майкрософт для пользователей. – Режим доступа: <https://docs.microsoft.com>.
- [5] Haffner P. A General Segmentation Scheme for DjVu Document Compression / Patrick Haffner, Leon Bottou, Yann Lecun and Luc Vincent // Conference: Proceedings of the International Symposium on Mathematical Morphology (ISMM'02), 2002. – p. 3-4.
- [6] Vasin Yu.G. An effective format for representing graphic information / Yu.G. Vasin and L.I. Lebedev // Pattern recognition and image analysis, 2012, Vol. 22, No. 2, pp. 393–398. © Pleiades Publishing, Ltd., 2012..
- [7] <https://djvu.com/> - Официальный сайт.
- [8] <https://opencv.org/> - Официальный сайт.

Об авторах

Лебедев Леонид Иванович, к.ф.-м.н., ведущий научный сотрудник Центра интеллектуальных информационных технологий Института ИТММ Нижегородского государственного университета им. Н.И. Лобачевского. E-mail: lebedev@pmk.unn.ru.

Вирясов Михаил Сергеевич, студент 4-го курса кафедры Прикладной информатики Института ИТММ Нижегородского государственного университета им. Н.И. Лобачевского. E-mail: mix3micha@mail.ru.

Громов Владимир Петрович, ведущий программист Центра интеллектуальных информационных технологий Института ИТММ Нижегородского государственного университета им. Н.И. Лобачевского. E-mail: GVP1941@bk.ru.