

ИССЛЕДОВАНИЕ ВОЗМОЖНОСТИ УЛУЧШЕНИЯ АЛГОРИТМА БИЛИНЕЙНОЙ ИНТЕРПОЛЯЦИИ ДЛЯ КОРРЕКТИРОВКИ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ ПРИМЕНЕНИЕМ ТЕОРИИ ПОЛЕЙ ОРИЕНТАЦИИ

Н. Е. Тимофеева, А. С. Гераськин

*Саратовский национальный исследовательский государственный университет
им. Н. Г. Чернышевского*

Поступила в редакцию 01.11.2017 г.

Аннотация. Представлен улучшенный алгоритм билинейной интерполяции для корректировки цифрового изображения с использованием теории полей ориентации и последующего сравнения с аналогами. Дан алгоритм увеличения изображения с применением теории полей ориентации. Проведен сравнительный анализ алгоритмов увеличения изображений, используя метрику PSNR.

Ключевые слова: цифровое изображение, пиксель, алгоритм увеличения изображения, билинейная интерполяция, теория полей ориентации, метрика PSNR.

Annotation. An improved bilinear interpolation algorithm for correcting a digital image using the theory of orientation fields and subsequent comparison with analogs is presented. An algorithm for increasing the image using the theory of orientation fields is presented. A comparative analysis of image enhancement algorithms is carried out using the PSNR metrics.

Keywords: digital image, pixel, image expansion algorithm, bilinear interpolation, orientation field theory, metrics PSNR.

ВВЕДЕНИЕ

В современном мире информационные технологии используются во всех сферах человеческой деятельности. Особенно бурно развиваются такие направления как: системы визуального наблюдения, системы технического зрения, интеллектуальные системы обнаружения вторжений и др. Все эти системы построены на цифровой обработке изображений. Любое цифровое изображение имеет определенное количество пикселей, и как следствие определенное разрешение. Для анализа и дальнейшей обработки таких изображений требуется высокое разрешение, так как при обработке необходимо учитывать множество мелких деталей. Многие интеллектуальные системы компьютерного зрения изначально полагают, что на вход подаются изображение высоко разрешения. Например, алгоритмы обработки изображе-

ний, поступающих с камер видеофиксации нарушений, будут работать намного устойчивее, если на вход они будут получать изображение с высоким разрешением. Повысит информативность таких изображений можно техническим способом, однако это приводит к большей стоимости такого оборудования, что является не всегда приемлемым. Поэтому способ воздействия на изображение программными средствами наиболее интересен, так как не требует замены оборудования. Исходя из этого, реализация повышения информативности цифрового изображения программным путем является актуальной задачей. Под повышением информативности цифрового изображения понимается увеличение масштаба изображения, повышения разрешения изображения для того, чтобы на нём было возможным видеть то, что не видно на кадрах низкого разрешения.

Известно множество подходов к решению задачи повышения информативности цифрового изображения. Как правило, они от-

личаются друг от друга используемыми алгоритмами увеличения и механизмом борьбы с артефактами, что приводит к значительному увеличению времени обработки цифрового изображения.

Наиболее распространенными являются два алгоритма увеличения цифрового изображения: билинейная и бикубическая интерполяции. Так как в данных алгоритмах цветовые координаты пикселя вычисляются путём интерполяции четырёх соседних, обработанное изображение получается размытым. Поэтому применение этих алгоритмов в системах видеофиксации и компьютерном зрении не представляется возможным. Для применения данных алгоритмов в вышеперечисленных системах необходимо проводить корректировки изображения [1].

В данной работе описан метод, основанный на теории полей ориентации и применяемый для корректировки изображений, увеличенных в два раза с помощью алгоритма билинейной интерполяции. Обработка пикселей методом полей ориентации позволяет повысить точность изображения после работы алгоритма и уменьшить уровень артефактов.

АЛГОРИТМЫ УВЕЛИЧЕНИЯ ИЗОБРАЖЕНИЙ

Общепринятые алгоритмы увеличения можно поделить на две категории: адаптивные и неадаптивные.

Неадаптивные алгоритмы включают: метод ближайшего соседа, билинейный, бикубический, сплайны, функция кардинального синуса (sinc), метод Ланцоша и другие. В зависимости от сложности, они используют от 0 до 256 (или более) смежных пикселей для интерполяции. Чем больше смежных пикселей они включают, тем более точными могут оказаться, но это достигается за счёт значительного прироста времени обработки. Эти алгоритмы могут использоваться как для развёртки, так и для масштабирования изображения [2].

Адаптивные алгоритмы включают в себя многие коммерческие алгоритмы в лицензированных программах, таких как Qimage,

PhotoZoom Pro, Genuine Fractals и другие. Многие из них применяют различные версии своих алгоритмов (как правило, на основе по-пиксельного анализа), т.е. когда обнаруживают наличие границы, производится попытка минимизировать неприглядные дефекты интерполяции в местах, где они наиболее видны. Эти алгоритмы в первую очередь разработаны для максимизации бездефектной детальности увеличенных изображений, так что некоторые из них для вращения или изменения перспективы изображения непригодны [2].

Адаптивные методы изменяются в зависимости от предмета интерполяции (резкие границы, гладкая текстура), тогда как неадаптивные методы обрабатывают все пиксели одинаково.

Наиболее часто используемым алгоритмом при увеличении цифровых изображений является алгоритм билинейной интерполяции. Это связано с низкой ресурсоемкостью алгоритма. Однако при применении данного алгоритма возникает сильная пикселизация изображения.

Билинейная интерполяция рассматривает квадрат 2×2 известных пикселя, окружающих неизвестный. В качестве интерполированного значения используется взвешенное усреднение этих четырёх пикселей. В результате изображения выглядят значительно более гладко, чем результат работы метода ближайшего соседа [2].

Билинейную интерполяцию, как правило, рассматривают как расширение линейной интерполяции для функций двух переменных, то есть это линейная интерполяция первого порядка в двумерном случае (рис. 1):

$$f(x, y) = (1-x)(1-y)f(0,0) + x(1-y)f(1,0) + (1-x)yf(0,1) + xyf(1,1)$$

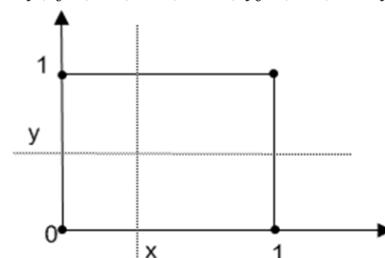


Рис. 1. Геометрическая интерпретация линейной интерполяции первого порядка в двумерном случае

Таким образом, можно провести линейную интерполяцию сначала в одном направлении, а затем в другом.

Рассмотрим эту задачу, более подробно используя рис. 2.

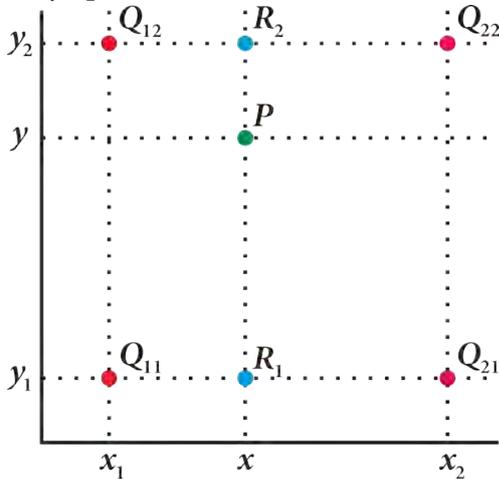


Рис. 2. Билинейная интерполяция

Допустим, что необходимо интерполировать значение функции f в точке $P = (x, y)$. Для этого необходимо знать значения функций в точках $Q_{11} = (x_1, y_1)$, $Q_{12} = (x_1, y_2)$, $Q_{21} = (x_2, y_1)$ и $Q_{22} = (x_2, y_2)$, которые окружают P . На рис. 2 изображены четыре красные точки, которые являются известными значениями функции, а значение в зеленой точке должно быть интерполировано. Для этого необходимо выполнить следующие шаги:

1. интерполируется значение вспомогательных точек R_1 и R_2 вдоль оси абсцисс, где

$$R_1 = (x, y_1) = (x, y_2)$$

2. проводится линейная интерполяция между вспомогательными точками R_1 и R_2

$$f(R_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21})$$

$$f(R_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22})$$

3. вычисляем значение функции в точке P .

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2).$$

Таким образом, значение функции в этой точке можно вычислить следующей формулой:

$$f(x, y) \approx \frac{f(Q_{11})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y_2 - y) +$$

$$+ \frac{f(Q_{21})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y_2 - y) +$$

$$+ \frac{f(Q_{12})}{(x_2 - x_1)(y_2 - y_1)} (x_2 - x)(y - y_1) +$$

$$+ \frac{f(Q_{22})}{(x_2 - x_1)(y_2 - y_1)} (x - x_1)(y - y_1).$$

Результат билинейной интерполяции не зависит от порядка шагов [3].

Если данный алгоритм применить к задаче увеличения цифровых изображений, то будет наблюдаться сильная пикселизация картинки. Для решения данной проблемы применим теорию полей ориентации.

ПОЛЯ ОРИЕНТАЦИИ ИЗОБРАЖЕНИЯ

Цифровое изображение в растровом представлении представляет собой набор пикселей. Каждому пикселю можно сопоставить некоторую монотонную функцию, которая будет отражать изменение его яркости. Так как значение пикселя, в данном представлении, отражает низкоуровневые значения изображения, то для качественной обработки цифрового изображения этой информации не хватает.

Рассмотрим задачу увеличения цифрового изображения как задачу изменения масштаба цифрового изображения, а именно вычисление значения яркости в точках, которые не принадлежат исходному множеству, в котором эти значения известны.

Как правило, для решения данной задачи применяют алгоритмы интерполяции, которые основываются на сопоставлении пикселям некоторых функций с последующим их сложением. Данный подход при больших увеличениях создает эффект ступенчатости.

Для избавления от ступенчатости при обработке цифровых изображений вводя анизотропии, причем анизотропия должна зависеть от локальных особенностей изображения [4]. В качестве такого введения может выступать ориентация – направление, вдоль которого изображение меняется меньше всего.

Ориентация являются более высокоуровневыми признаками изображения, чем цвета пикселей, поэтому могут быть эффективно

применены для задачи увеличения цифрового изображения [4].

Если рассмотреть ориентацию в непрерывном случае, то она будет определять направление изолиний в точке, перпендикулярно вектору градиента. Таким образом, ориентацию можно представить, как вектор, который известен до знака. Назовем их векторами ориентаций. Для работы с данными векторами их необходимо привести к виду обычных векторов [4].

Это можно провести, проведя удвоение угла поворота, после чего вектор ориентации превращается в обычный вектор. Операция удвоения угла обладает следующими свойствами:

- после удвоения угла векторы \vec{x} и $-\vec{x}$ становятся равными;
- максимально различающиеся ориентации – перпендикулярные превращаются в максимально различающиеся векторы – противоположно направленные.

Таким образом, удвоение угла осуществляет взаимно-однозначное отображение пространства векторов ориентации в пространство обычных векторов.

Для удвоения угла векторов можно рассмотреть вектор $(x, y)^T$ как комплексное число $x + iy$, и возвести его в квадрат:

$$(x + iy)^2 = (x^2 - y^2) + i * (2xy).$$

Для соответствующего преобразования вектора введем операцию sqr :

$$\text{sqr}((x, y)^T) = (x^2 - y^2, 2xy)^T.$$

Это выражение может быть вычислено целочисленно.

Наряду с направлением ориентация обладает еще одним важным свойством, которое называется «выраженность». «Выраженность» представляет собой число, которое характеризует насколько сильно, данная ориентация выделяется среди других. В рассматриваемом случае в качестве выраженности будет взята длина вектора градиента [4].

Для вычисления ориентации в непрерывном случае, воспользуемся алгоритмом структурного тензора. Для всех пикселей изображения вычисляются дискретные аналоги гра-

диентов, эти градиенты возводятся в квадрат по формуле $(x + iy)^2 = (x^2 - y^2) + i * (2xy)$, а затем квадраты усредняются по окрестности каждого пикселя. Этот алгоритм позволит получить выраженную ориентацию в областях с большими по модулю и параллельными градиентами. Если же векторы градиентов малы по модулю или перпендикулярны, то ориентация получается слабовыраженной. Реализация данного алгоритма представляется следующим образом:

1. Вычисляем дискретный аналог градиента для каждого пикселя. Это можно сделать, при помощи следующего разностного шаблона:

$$\vec{g}_{i,j} = \begin{pmatrix} 2c_{i,j+1} + c_{i-1,j+1} + c_{i+1,j+1} - \\ - 2c_{i,j-1} - c_{i-1,j-1} - c_{i+1,j-1} \\ 2c_{i+1,j} + c_{i+1,j-1} + c_{i+1,j+1} - \\ - 2c_{i-1,j} - c_{i-1,j-1} - c_{i-1,j+1} \end{pmatrix}.$$

2. Возводим векторы в квадрат:

$$\vec{a}_{i,j} = \text{sqr}(\vec{g}_{i,j}).$$

3. Вычисляем промежуточные вертикальные суммы по рекуррентной формуле:

$$\vec{b}_{i,j} = \vec{b}_{i-1,j} - \vec{a}_{i-N-1,j} + \vec{a}_{i+N,j}.$$

4. Вычисляем горизонтальные суммы вертикальных сумм и получаем требуемую ориентацию окрестности:

$$\vec{d}_{i,j-1} = \vec{d}_{i,j-1} - \vec{b}_{i,j-N-1} + \vec{b}_{i,j+N}.$$

Данный алгоритм в случае квадратной окрестности размера $(2N + 1) \times (2N + 1)$ позволяет вычислять ориентацию за $O(1)$.

АЛГОРИТМ УВЕЛИЧЕНИЯ ИЗОБРАЖЕНИЯ С ПРИМЕНЕНИЕМ ТЕОРИИ ПОЛЕЙ ОРИЕНТАЦИИ

Из проведенного анализа выше следует, что поле ориентаций гораздо более гладкое, чем исходное изображение, поэтому может быть хорошо проинтерполировано даже простой билинейной интерполяцией. Однако не очевидно, как поле ориентаций превратить обратно в изображение.

Для этого был разработан следующий алгоритм:

1. По исходному изображению вычисляем поле ориентаций В.

2. Увеличиваем разрешение исходного изображения алгоритмом билинейной интерполяции.

3. Увеличиваем поле ориентаций В удвоением угла по формуле $\text{sq}\left((x, y)^T\right) = \left(x^2 - y^2, 2xy\right)^T$, получаем поле ориентаций высокого разрешения.

4. Сокращаем вдвое углы поля ориентаций D.

5. Для увеличенного изображения вычисляем поле градиентов F.

6. Преобразуем поле градиентов F так, чтобы его вектора были коллинеарны векторам поля D.

7. По полю градиентов F восстанавливаем изображение.

СРАВНИТЕЛЬНЫЙ АНАЛИЗ АЛГОРИТМОВ УВЕЛИЧЕНИЯ ИЗОБРАЖЕНИЙ

Для тестирования работоспособности разработанного алгоритма использовались изображения с ресурса «ImageNet». ImageNet – это проект, обладающий массивной базой данных аннотированных изображений, которые используются для тестирования и отработки алгоритмов распознавания образов, алгоритмов увеличения и систем машинного зрения.

Для численного анализа сравнения работы алгоритмов на различных изображениях применяется метрика PSNR. Как правило, PSNR определяется через среднеквадратичное отклонение (MSE), которое для двух монохромных изображений I и K размера $m \times n$, одно из которых считается зашумленным приближением другого, вычисляется так:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - K(i, j)|^2$$

PSNR определяется так:

$$PSNR = 10 \log_{10} \left(\frac{MAX_I^2}{MSE} \right),$$

где MAX_I – это максимальное значение, принимаемое пикселем изображения.

Когда пиксели имеют разрядность 8 бит, $MAX_I = 255$. Вообще говоря, когда значения сигнала представлены линейно с B битами на значение, максимально возможное значение MAX_I будет $2^B - 1$ [5].

Для цветных изображений с тремя компонентами RGB на пиксель применяется такое же определение PSNR, но MSE считается по всем трем компонентам (и делится на утроенный размер изображения).

Таким образом, алгоритм сравнения изображения с его аналогом выглядит следующим образом:

– уменьшаем исходное цифровое изображение;

– увеличиваем уменьшенное цифровое изображение до размеров исходного;

– вычисляем значения метрики PSNR для исходного и увеличенного изображений.

Для осуществления сравнения алгоритмов билинейной и бикубической интерполяции с алгоритмом билинейной интерполяции с последующей корректировкой методом полей ориентации разобьем все изображения на четыре группы:

Группа 1 – изображения, имеющие несколько крупных объекты, с отсутствием резких цветовых переходов;

Группа 2 – изображение имеет множество крупных объектов с резкими цветовыми переходами;

Группа 3 – изображение имеет множество мелких объектов с плавными переходами;

Группа 4 – имеется много мелких объектов с различной цветовой палитрой и резкими переходами.

Результаты сравнения алгоритмов обработки изображений представлены на рис. 3.

Из полученного сравнения методов, представленных выше видно, что разработанный алгоритм осуществляет увеличение цифровых изображений практически аналогично методу бикубической интерполяции. Но при наличии в изображениях большого количества деталей и цветовых переходов разработанный алгоритм показывает наилучшие результаты.

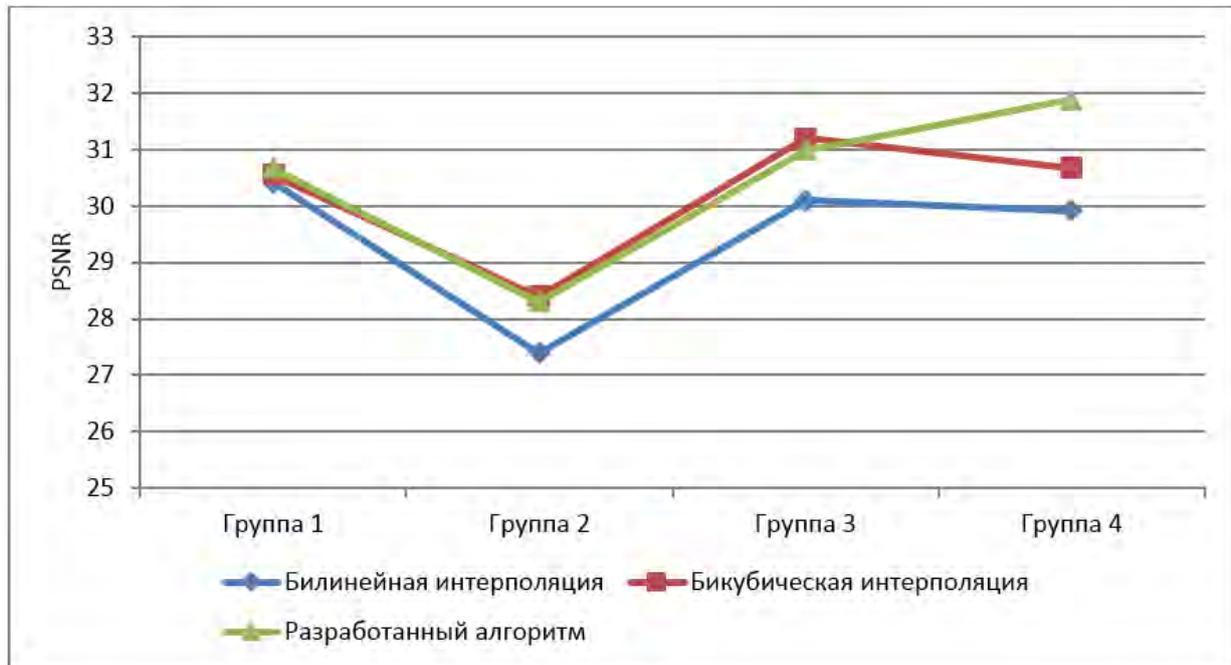


Рис. 3. График результатов сравнения

ЗАКЛЮЧЕНИЕ

Разработан новый метод корректировки алгоритма билинейной интерполяции с применением теории полей ориентации. Применение теории поля ориентаций при увеличении цифрового изображения, повысить точность изображения после работы алгоритма и уменьшить уровень артефактов. Особенно это заметно, если на изображении имеется большое количество мелких деталей и цветовых переходов.

СПИСОК ЛИТЕРАТУРЫ

1. *Абрамова, О. Ф.* Обзор алгоритмов масштабирования растровой графики / О. Ф. Абрамова, А. Е. Иванов, А. Н. Инкин // *European Student Scientific Journal*. – 2016. – № 2. – Режим доступа: <http://sjes.esrae.ru/ru/article/view?id=371>.

2. *Маркелов, К. С.* Модель повышения информативности цифровых изображений на базе метода суперразрешения / К. С. Маркелов // *Инженерный вестники* – М. : ФГБОУ ВПО «МГТУ им. Н.Э. Баумана». – 2013. – Вып. 3. – С. 525–542.

3. *Красильников, Н. Н.* Цифровая обработка 2D- и 3D-изображений: учеб. пособие / Н. Н. Красильников. – СПб. : БХВ-Петербург. – 2011. – 608 с.

4. *Сушинов А. А.* Новый метод вычисления поля ориентаций изображения / А. А. Сушинов, И. Н. Тетеревлев, В. В. Царевский // *Изв. ЮФУ*. – 2011. – Т. 121: Математика, вып. 8. – С. 187–199.

5. *Сидоров Д. В.* К вопросу оценки качества множества восстановленных изображений / Д. В. Сидоров // *Общество с ограниченной ответственностью «Синергия ПРЕСС»*. Прикладная информатика. – 2008. – Вып. 4. – С. 92–95.

Тимофеева Н. Е. – зав. лабораторией теоретических проблем информатики и ее приложений кафедры дискретной математики и информационных технологий, Саратовский национальный исследовательский государственный университет им. Н. Г. Чернышевского.

Тел.: +7 904 243 1815

E-mail: timofeevane@yandex.ru

Timofeeva N. E. – Chief of Laboratory of Theoretical Problems of computer science and its applications of the Chair of Discrete Mathematics and Informational Technologies, National Research Saratov State University.

Tel.: +7 04 243 1815

E-mail: imofeevane@yandex.ru

Гераськин А. С. – канд. пед. наук, доцент кафедры теоретических основ компьютерной безопасности и криптографии, Саратовский национальный исследовательский государственный университет им. Н. Г. Чернышевского.

Тел.: +7 905 381 6366

E-mail: gerascinas@mail.ru

Geraskin A. S. – Candidate of Pedagogical Sciences, Associate Professor of the Chair of Computer Science and Cryptography Theory, National Research Saratov State University.

Tel.: +7 905 381 6366

E-mail: gerascinas@mail.ru