

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Математико-механический факультет

Системного программирования

Улитин Александр

Повышение разрешения изображений автомобильных номеров

Бакалаврская работа

Допущена к защите.

Зав. кафедрой:

д. ф.-м. н., профессор Терехов А. Н.

Научный руководитель:

к. ф.-м. н., доцент Вахитов А. А.

Рецензент:

инженер-программист Пименов А. А.

Санкт-Петербург

2013

SAINT-PETERSBURG STATE UNIVERSITY

Mathematics & Mechanics Faculty

Chair of Software Engineering

Ulitin Alexander

ИЗМЕНИТЬ ТУТ

Bachelor's Thesis

Admitted for defence.

Head of the chair:
professor Andrey Terkhov

Scientific supervisor:
Ph D Alexander Vakhitov

Reviewer:
Ph. D Alexander Pimenov

Saint-Petersburg

2013

Содержание

1. Введение	5
2. Обзор существующих алгоритмов	6
2.1. Обучаемые алгоритмы	6
2.2. Интерполяционные	6
2.2.1. Метод ближайшего соседа	6
2.2.2. Билинейная интерполяция	7
2.2.3. Бикубическая интерполяция	7
2.3. Спектральное представление	7
2.4. Регуляризация	8
3. Постановка задачи	9
3.1. Метрика PSNR	9
3.2. Интерполяция	10
3.3. Изображения как векторы	10
4. Задача	12
5. Используемые алгоритмы	13
5.1. Обучаемый алгоритм на словарях	13
5.2. Алгоритм с использованием регуляризации	14
6. Результаты	16
6.1. Обучаемый алгоритм на словарях	16
6.2. Алгоритм с использованием регуляризации	16
7. Заключение	17
7.1. какие средства использовались	17
7.2. что было сделано	17

7.3. Дальнейшее исследование	17
7.3.1. исследование других подходов	17
7.3.2. автоматическое детектирование	17

1 Введение

Цифровое изображение имеет конечное количество пикселей. Количество этих пикселей деленное на единицу площади называется разрешением цифрового изображения.

Для анализа и обработки изображений в большинстве случаев используются изображения высокого разрешения, так как они позволяют увидеть детали, которые не различимы, или плохо различимы на изображениях с низким разрешением. Многие задачи компьютерного зрения изначально полагаются на то, что изображение на входе в хорошем разрешении. Например, алгоритмы распознавания автомобильного номера будут работать намного устойчивее, если на вход они будут получать изображение с высоким разрешением. С первого взгляда, задача увеличения разрешения чисто аппаратная – чтобы увеличить разрешение необходимо просто взять фотоаппарат с большей разрешающей способностью. Но часто уже имеется некоторая фотография или набор фотографий, и нет возможности повторно произвести съемку. Примером может послужить низкокачественные картинки с изображений камер наблюдения. В этом случае уместно использовать программные алгоритмы повышения разрешения (Super-resolution методы). Эти алгоритмы используют некоторое дополнительное знание о изображении для качественного повышения разрешения. Примером дополнительных знаний может послужить информация о том, как двигался объект во время съемки, или оптические параметры камеры. Зная эти параметры можно алгоритмическим путем получить более качественное изображение.

В этой работе рассмотрены два алгоритма для повышения разрешения на примере автомобильных номеров.

2 Обзор существующих алгоритмов

Классические алгоритмы Super-Resolution делятся на четыре категории¹. Существует несколько обзорных статей по этим алгоритмам [1, 2]

2.1 Обучаемые алгоритмы

Обучаемые алгоритмы используют для восстановления изображений используют знание о том, что в действительности изображено на картинке. Такие алгоритмы хорошо подходят для повышения разрешения однотипных изображений, таких как лица или автомобильные номера.

2.2 Интерполяционные

Подход интерполяции наиболее интуитивно понятный метод решения задачи SR. Решение задачи, в общем случае, разбивается на три шага: определение относительного движения (Motion estimation) для каждого изображения, интерполяция LR изображений на HR сетку, удаление смазывания и шума для полученного HR изображения.

2.2.1 Метод ближайшего соседа

При печати этого текста компьютером тоже была решена задача повышения разрешения, т.к. принтер печатает изображения с намного более большим разрешением, чем была исходная картинка. В этом случае, задача решается не самым лучшим образом - компьютер просто просчитывает недостающие точки используя значения пикселя изображения, который находится ближе

¹Стоит заметить, что существуют алгоритмы, которые будут плохо ложиться на эти категории. Но такие алгоритмы не сыскали большой известности, поэтому из рассмотрения убраны.

всего. Такая примитивная интерполяция называется «метод ближайшего соседа».

2.2.2 Билинейная интерполяция

Билинейная интерполяция чуть более продвинутый метод интерполяции. Для интерполяции используется взвешенные суммы линейных интерполяций по вертикальной и горизонтальной оси. Более подробное описание и формулу билинейной интерполяции можно найти в [3, стр. 64].

2.2.3 Бикубическая интерполяция

Бикубическая интерполяция обладает одним из лучших сочетаний качества и скорости. Бикубическая интерполяция является обобщением интерполяции кубическими сплайнами на двумерный случай. Алгоритм кубической интерполяции реализован во многих редакторах цифровых изображений. В нашей статье мы будем считать алгоритм бикубической интерполяции отправной точкой. По сравнению результатов бикубической интерполяции и предложенного метода можно судить том, улучшает алгоритм изображение, или нет. Более подробно алгоритм описан в [4].

2.3 Спектральное представление

В алгоритмах этого класса используется разложение изображения в некоторый другой базис (примером может являться двумерное дискретное преобразование Фурье, или дискретное косинусное преобразование), и обратное к нему. Этот метод удобен тем, что мелкие детали соответствуют высокочастотным компонентам базиса, и за счет этого становится возможным вос-

становливать мелкие детали изображения, которые возможно получить за счет интерполяции с нескольких LR изображений.

2.4 Регуляризация

Алгоритмы с использованием регуляризации используют некоторые знания о природе изображения. Например, в алгоритме MAP используют предположение о вероятности распределений точек.

3 Постановка задачи

Для рассуждений необходимо некоторым образом формально записать то, с чем имеем дело. Для этого понадобится ввести несколько дополнительных понятий.

3.1 Метрика PSNR

Для того, чтобы сравнивать несколько алгоритмов повышения разрешения необходимо ввести какую-нибудь количественную метрику. Чтобы иметь возможность посчитать метрику необходимо знать эталонное изображение, то есть такое изображение, которое должен выдать идеальный алгоритм повышения разрешения. Очевидно, что создать эталонный алгоритм невозможно – это отображение из $\mathbb{N}^{n \cdot m} \rightarrow \mathbb{N}^{nk \cdot mk}, k \geq 2$. Знание эталонного изображения несколько выходит за рамки поставленной во введении задачи – в реальных условиях истинное изображение будет неизвестно. Для тестирования алгоритмов были использована стандартная модель получения изображений низкого разрешения из высокого.

Пусть x – истинное изображение, \tilde{x} – найденное изображение из изображений низкого разрешения. Посчитаем среднеквадратичную ошибку между пикселями этих изображений.

$$\text{MSE}(\tilde{x}, x) = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [\tilde{x}(i, j) - x(i, j)]^2$$

И выразим отношение MAX_I – максимальное значение яркости изобра-

жения к среднеквадратичной ошибке в децибелах $\text{PSNR}(\tilde{x}, x)$.

$$\text{PSNR}(\tilde{x}, x) = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}(\tilde{x}, x)} \right)$$

Это и будет нашей метрикой. Задачей будет достигнуть максимального значения PSNR из имеющихся данных.

3.2 Интерполяция

Одним из наиболее известных способов повышения разрешения является интерполяция. Для любого изображения можно бесконечно повышать разрешения, просто добавляя дополнительные значения между пикселями исходного изображения. Однако, такое повышение разрешения не всегда дает хорошие результаты. Поскольку значение PSNR лишь показывает разницу между двумя изображениями, то нет возможности численно оценить насколько алгоритм Super-resolution хорошо справился с задачей. Но если сравнивать этот же алгоритм со значением полученные каким-либо методом интерполяции то наглядно видно, где алгоритм справляется хорошо, а где плохо.

3.3 Изображения как векторы

Для возможности записывать все операции над изображением как матричное умножение, будем считать, что изображение высокого разрешения, которое мы хотим восстановить x размера $L_1 N_1 \times L_2 N_2$ записано в виде вектора в виде вектора $x = [x_1, x_2, \dots, x_N]^T$, где $N = L_1 N_1 \times L_2 N_2$. Другими словами, изображение x это то изображение, которое мы бы получили, если бы использовали матрицу фотоаппарата с более большим разрешением. Мы

будем рассматривать следующую стандартную модель искажения изображения

$$y_r = DB_r M_r x + \sigma_k \epsilon$$

где M_r матрица геометрического преобразования для конкретного изображения y_r , B_r обозначает матрицу размытия размером $L_1 N_1 L_2 N_n \times L_1 N_1 L_2 N_2$, D матрица понижения разрешения размера, $\sigma_k \epsilon$ – аддитивный шум.

4 Задача

$$y_r = DH_r W_r x + \sigma_k \epsilon, \quad 1 \leq r \leq m$$

где:

- x оригинальное изображение
- y_r наблюдение r
- D матрица понижение разрешения
- W_r матрица геометрического искажения
- H_r матрица размытия наблюдения r
- $\sigma_k \epsilon$ шум наблюдения r
- m количество наблюдений

Задача найти

$$\tilde{x} = \underset{\hat{x}}{\operatorname{argmax}} PSNR(\hat{x}, x)$$

5 Используемые алгоритмы

Для того, чтобы повысить разрешение автомобильного номера использовались два подхода: обучаемый [5] и интерполяционный с использованием регуляризации [6]. Этот выбор основывался исходя из природы изображения автомобильного номера.

5.1 Обучаемый алгоритм на словарях

В статье [5] авторы предлагают использовать для задачи повышения разрешения одного изображения использовать пару связанных словарей D_x и D_y – словари содержащие патчи из пространства изображений с высоким и низким разрешением.

Алгоритм требует предварительного обучения нейронной сети. Для этого была создана база с изображениями автомобильных номеров. Используя авторскую реализацию алгоритма была обучена пара словарей для восстановления изображений.

Обоснованием, для применения этого метода были следующие:

- Автомобильные номера содержат конечный набор патчей. Автомобильные номера очень похожи между собой. И явно видно, что между двумя изображениями автомобильных номеров есть общие кусочки изображения.
- Автомобильные номера имеют простую структуру – у них нет сложных переходов, большинство смены цветов это просто переход от белого к черному.

5.2 Алгоритм с использованием регуляризации

В статье [7] авторы предлагают улучшение алгоритма МАР для задачи повышения изображения. Изображение моделируется как Марковская сеть и используется алгоритм шаговой оптимизации в сочетании с адаптивным регуляризатором для сохранения границ.

Прежде, чем говорить о самом алгоритме требуется сделать предварительный шаг – установление движения (motion estimation).

Общий вид алгоритма следующий:

1. посчитать среднее от выровненных и билинейно интерполированных изображений $X^{(0)}$. Будем называть полученные на этом шаге изображение «начальным приближением»
2. положить $\gamma = 2v$, где v – максимальное значение градиента по осям x и y в начальном приближении $X^{(0)}$
3. Do:

$$(a) \quad X^{(n+1)} = X^{(n)} - \alpha \cdot \text{grad}(X^{(N)}, \gamma)$$

$$(b) \quad n = n + 1$$

$$(c) \quad \text{If } (\text{norm}(X^{(n)} - X^{(n-1)}) < \epsilon) \\ \text{then } \gamma^{(n)} = \max\{\gamma_{\text{target}}, k\gamma^{(n-1)}\}$$

$$\text{Until } (\text{norm}(X^{(n)} - X^{(n-1)}) < \epsilon) \text{ and } (\gamma^{(n)} = \gamma_{\text{target}})$$

4. $\hat{x} = X^{(n)}$ где $\alpha, \epsilon, k, \gamma_{\text{target}}$ экспериментально подобранные параметры алгоритма. В нашей реализации мы использовали $\alpha = 0.07, \epsilon = 0.005, k = 0.95, \gamma_{\text{target}} = 10$

$$\text{grad}(x, \gamma) = \frac{1}{\sigma^2} \sum_{r=1}^m W_R^T H_r^T D^T (D H_r W_r x - y_r) + \lambda \cdot G(x, \gamma)$$

где λ экспериментально подобранные регуляризационный параметр и градиент для точки (i, j) задан следующей формулой

$$\begin{aligned} G(i, j) &= 2 [x(i, j) - x(i, j - 1)] \exp (-[x(i, j) - x(i, j - 1)]^2 / \gamma) \\ &+ 2 [x(i, j) - x(i, j + 1)] \exp (-[x(i, j) - x(i, j + 1)]^2 / \gamma) \\ &+ 2 [x(i, j) - x(i + 1, j)] \exp (-[x(i, j) - x(i + 1, j)]^2 / \gamma) \\ &+ 2 [x(i, j) - x(i - 1, j)] \exp (-[x(i, j) - x(i - 1, j)]^2 / \gamma) \end{aligned}$$

6 Результаты

6.1 Обучаемый алгоритм на словарях

графики сравнения PSNR скорость работы примеры изображений

6.2 Алгоритм с использованием регуляризации

графики сравнения PSNR скорость работы примеры изображений
сравнение с билинейной/бикубической интерполяцией

7 Заключение

7.1 какие средства использовались

7.2 что было сделано

7.3 Дальнейшее исследование

7.3.1 исследование других подходов

7.3.2 автоматическое детектирование

Список литературы

- [1] Kang M. G. Park S. C., Park M. K. Super-resolution image reconstruction: a technical overview. *IEEE Signal Processing Magazine*, 20:21–36, 2003.
- [2] Jing Tian and Kai-Kuang Ma. A survey on super-resolution imaging. *Signal, Image and Video Processing*, 5(3):329–342, 2011.
- [3] Rafael C Gonzalez and E Richard. Woods, digital image processing. *ed: Prentice Hall Press, ISBN 0-201-18075-8*, 2002.
- [4] Robert Keys. Cubic convolution interpolation for digital image processing. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 29(6):1153–1160, 1981.
- [5] Jianchao Yang, Zhaowen Wang, Zhe Lin, Scott Cohen, and Thomas Huang. Coupled dictionary training for image super-resolution. *Image Processing, IEEE Transactions on*, 21(8):3467–3478, 2012.
- [6] KV Suresh, G Mahesh Kumar, and AN Rajagopalan. Superresolution of license plates in real traffic videos. *Intelligent Transportation Systems, IEEE Transactions on*, 8(2):321–331, 2007.
- [7] Jie Yuan, Si-dan Du, and Xiang Zhu. Fast super-resolution for license plate image reconstruction. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.
- [8] Mark R Banham and Aggelos K Katsaggelos. Digital image restoration. *Signal Processing Magazine, IEEE*, 14(2):24–41, 1997.

- [9] David Capel and Andrew Zisserman. Computer vision applied to super resolution. *Signal Processing Magazine, IEEE*, 20(3):75–86, 2003.
- [10] Aram Danielyan, Vladimir Katkovnik, and Karen Egiazarian. Bm3d frames and variational image deblurring. *Image Processing, IEEE Transactions on*, 21(4):1715–1728, 2012.
- [11] Aggelos K Katsaggelos, Rafael Molina, and Javier Mateos. Super resolution of images and video. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 1(1):1–134, 2007.
- [12] Юрий Андреевич Землянский. Исследование вариационного подхода в задаче повышения разрешения изображения. Master’s thesis, СПбГУ Математико-Механический факультет, 2012.