

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ ФАКУЛЬТЕТ

Математико-Механический Факультет

Кафедра системного программирования

Землянский Юрий Андреевич

**Использование вариационного подхода в
задаче повышения разрешения
изображения**

Дипломная работа

«Допустить к защите» _____
Заведующий кафедрой
д.ф.-м.н., профессор А.Н. Терехов

Научный руководитель
к.ф.-м.н., доцент А.Т. Вахитов

Рецензент
ст. препод А.А. Пименов

Санкт-Петербург
2012 г.

SAINT-PETERSBURG STATE UNIVERSITY

Mathematics and Mechanics Faculty

System programming department

Zemlyanskiy Yury

Variational approach for Superresolution

Graduate paper

«Approved by» _____
head of department
A.N. Terehov

Supervisor
A.T. Vahitov

Reviewer
A.A. Pimenov

Санкт-Петербург
2012 г.

Содержание

1	Введение	3
2	Модель наблюдений	4
3	Обзор существующих методов	7
3.1	Обучаемые алгоритмы	7
3.2	Алгоритмы, основанные на спектральном представлении изображений	7
3.3	Интерполяционные алгоритмы	8
3.4	Баесовские алгоритмы	8
3.5	Алгоритмы с использованием регуляризации	9
4	BM3D-фильтр	10
4.1	Compressive sensing	10
4.2	BM3D	10
5	Постановка задачи	12
6	Алгоритм	14
6.1	Случай single-frame	14
6.2	Случай multi-frame	15
7	Реализация	16
8	Эксперименты	17
8.1	Single-frame SR	17
8.2	Multi-frame SR	17
9	Итоги	22

1 Введение

На сегодняшний день существуют научные и технологические задачи, для решения которых требуются изображения с высоким разрешением. Высокое разрешение (HR) означает большую плотность пикселей, а следовательно, и более точную детализацию. Например, располагая HR изображением томографии, врач сможет более точно поставить диагноз. Различные объекты проще отличить друг от друга по снимках со спутника, если у них высокое разрешение.

Прямой путь повысить разрешение - чисто технический - использовать камеру со светочувствительной матрицей большего разрешения. В этом случае, однако, во-первых, возрастает стоимость оборудования, а во-вторых, повышается уровень шума, так как на каждый пиксель теперь приходится меньше света. К тому же, не всегда у нас есть возможность поменять используемое оборудование - возможны ситуации, когда все, чем мы располагаем - несколько снимков с недостаточным качеством.

Есть другая возможность - различными алгоритмическими техниками повысить разрешение изображений. Первая публикация алгоритма, который по нескольким изображениям низкого разрешения (LR) восстанавливал HR изображение была в 1989 году - [1]. В настоящее время таким техникам посвящен целый раздел теории обработки сигналов и компьютерного зрения - *superresolution* (SR). В данной работе обзору SR методов посвящена отдельная глава.

Наша работа основывается на серии работ [29],[33]-[36] научной группы из Tampere University of Technology под руководством профессоров К. Егизаряна и В. Катковника. Ими разработан фильтр изображений BM3D, основанный на наличии похожих блоков на изображении и подходе compressive sensing. На данный момент, алгоритмы, основанные на использовании этого фильтра, являются state-of-art для средств фильтрации шума (*denoising*) и компенсации блюра (*deblurring*).

Кроме того, в статье [34] научной группой был предложена вариационная интерпретация задачи denoising и deblurring. В сотрудничестве с ними, мы с Виктором Пасиченко, студентом 542 группы, применили их подход к задаче SR. В итоге, нами был разработан и реализован на Matlab алгоритм SR для изображений на основе вышеупомянутых идей. По результатам экспериментов алгоритм показал хороший результат, в некоторых случаях превосходя предыдущий state-of-art алгоритм SR [29], тоже основанный на BM3D.

2 Модель наблюдений

Для начала мы опишем формальную модель наблюдений, которая связывает восстанавливаемое HR изображение с исходными LR изображениями. Эта модель является общепринятой, и используется в большинстве исследований (статьи [8], [9], [10]).

На рисунке 1, взятым из статьи [9], представлена общая схема получения LR изображения.

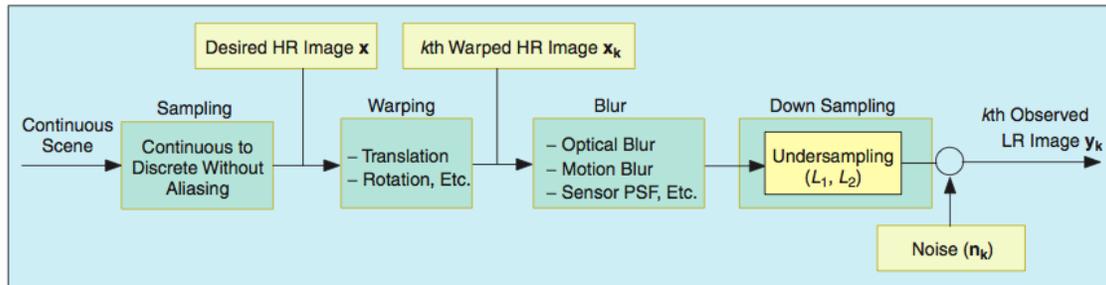


Рис. 1: Схема получения изображений

Прежде всего, мы предполагаем, что сцена, которую мы снимаем, не изменится. И наша задача - восстановить изображение этой сцены в высоком разрешении. Когда мы делаем съемку, камера неизбежно меняет свое положение - это первое, геометрическое искажение (*warping*). Затем на изображения накладывается эффект блюра размытия (*blur*). Изображение проецируется на пиксельную матрицу LR решетки меньшей размерности, чем нам нужно (*down sampling*). К получившемуся изображению добавляется случайный шум.

Для более удобной записи, изображения мы будем представлять в виде векторов (а не матриц). Пусть HR изображение, которое мы хотим восстановить, размера $l_1 n_1 \times l_2 n_2$, записано по строчкам как вектор $x = [x_1, x_2, \dots, x_n]$, где $n = l_1 n_1 l_2 n_2$. Вообще говоря, x - то изображение, которое бы мы получили, будь у нашего фотоаппарата матрица с большой светочувствительной матрицей. Однако все, чем мы располагаем - это LR изображения, размером $n_1 n_2$. Их мы тоже запишем как вектора $y_k = [y_{k,1}, \dots, y_{k,m}]$, где $m = n_1 n_2$.

Несколько слов про основные стадии модели.

- **Геометрическое искажение**

Отвечает за взаимное расположение снятых сцен в отдельных LR изображениях. Возникает из-за смещения, поворота камеры, хотя возможны и

более сложные преобразования. Операции смещения и или поворота для вектора-изображения x можно представить, как умножение на матрицу W_k . Существует целый класс алгоритмов, занимающихся *motion estimation*, определением параметров геометрического искажения объектов на разных изображениях - [37], [38].

- **Размытие**

“Размытие” изображения. Возникает из-за свойств используемых линз (оптический блюр - *optical blur*), движения камеры (*motion blur*), свойств оптического сенсора (*sensor PSF*).

Преобразование наложения блюра - B - обычно моделируется как свертка изображения с разными ядрами (*blur kernel*) - равномерными, гауссовскими.

Параметры блюра для конкретной камеры можно определить экспериментально. Если же данных о модели камеры нет, существуют алгоритмы, позволяющие определять параметры блюра по самим изображениям - [46]-[48].

- **Уменьшение размерности**

Оператор уменьшения размерности D (*decimation operator*) - просто уменьшает размерность изображения, выбирая только каждую l_1 -ую строчку и l_2 -ый столбец из исходного HR изображения.

- **Шум**

Аддитивный гауссовский белый шум - изображение, где каждый пиксель - нормально распределенная случайная величина, с нулевым математическим ожиданием и конечной дисперсией σ^2 . (Будем обозначать как $\sigma\epsilon$, $\epsilon \sim N^n(0, 1)$, n - количество пикселей в HR изображении.)

Уровень шума зависит от модели камеры и уровня освещенности, и, так же как и блюр, может быть экспериментально определен для конкретной модели камеры. В том случае, когда информации о том, какая аппаратура использовалась для съемки, нет, можно воспользоваться алгоритмами, определяющими уровень шума только по изображениям - [49]

В итоге, согласно нашей модели, верно следующее равенство

$$y_k = DBW_k x + \sigma_k \epsilon$$

где по известным исходным LR изображениям y_1, y_2, \dots , матрицам преобразований D, B, W_k и уровню шума σ требуется восстановить x .

Таким образом задача superresolution - типичная обратная задача.

Первая проблема, которая возникает при ее решении - вычислительная. Дело в том, что размер матриц D, B, W_k очень большой - равен $l_1 l_2 n_1^2 n_2^2$. Допустим, что размер исходного LR изображения 256×256 , и наша задача состоит в том, чтобы увеличить его в 2 раза по обеим координатам. Тогда размерность матриц $2^{16} \times 2^{18}$, т.е. всего $2^{34} \approx 10^{10}$ элементов. Обычные методы решения линейных уравнений здесь просто так не работают.

Вторая проблема в том, что эта задача относится к классу плохо обусловленных (*ill-posed*), т.е. точного решения может и не существовать, или наоборот существовать несколько; решение зависит не непрерывно от начальных условий. Из-за наличия шума в исходных данных, ошибок определения параметров модели получается, что просто найти подходящее HR изображение бывает недостаточно - оно может быть сильно отличаться от реальных изображений. Чтобы получать более качественные решения, разные алгоритмы накладывают на восстанавливаемое изображение дополнительные ограничения - проводят регуляризацию.

В следующей главе мы рассмотрим основные существующие подходы к SR в рамках описанной модели наблюдений.

3 Обзор существующих методов

На данный момент существуют несколько обзорных статей по SR алгоритмам - [9], [10], [11]. Обзор задачи в целом, с теоретической и практической стороны, есть в книге - [8]. В общем случае классические SR алгоритмы могут быть разделены на 5 категорий - обучаемые алгоритмы; алгоритмы, основанные на спектральном представлении изображений; интерполяционные алгоритмы; байесовские алгоритмы; регуляризируемые.

3.1 Обучаемые алгоритмы

Отдельный класс SR алгоритмов, которые для восстановления высокочастотных деталей используют информацию, полученную из заранее подготовленной тестовой базы HR изображений. Хотя такие алгоритмы и не подходят под нашу формулировку задачи, они могут быть применимы на практике, в тех случаях, когда восстанавливаемое изображение однотипные - тогда алгоритм можно будет натренировать для такого класса изображений (например, класс “лица людей”). Такие алгоритмы описаны в статьях [24]-[27].

3.2 Алгоритмы, основанные на спектральном представлении изображений

В исторически первой публикации по проблеме superresolution [1] Tsai и Huang использовали 2D DFT - двумерное дискретное преобразование Фурье, чтобы перевести изображение в спектр, затем объединяли коэффициенты с разных изображений и, применяя обратное преобразование Фурье, получали изображение большего разрешения. Rhee и Kang в статье [2] для получения спектра применили DCT - дискретное косинусное преобразование. В ряде последних статей - [3]-[7] - используется DWT - дискретное вэйвлетное преобразование.

Преимущества этого подхода в том, что мелкие детали в изображении, которые нужно восстановить, как раз соответствуют высокочастотным компонентам спектра - их удастся экстраполировать за счет информации с нескольких LR изображений.

Кроме того, преобразование в спектр может быть эффективно реализовано как на программном, так и на аппаратном уровнях. Поэтому такие алгоритмы выигрывают в скорости работы по сравнению с остальными.

Однако, этот подход предполагает, что геометрическое искажение между исходными изображениями низкого разрешения - только глобальный сдвиг, что в

реальных данных совсем не обязательно.

3.3 Интерполяционные алгоритмы

Подобные алгоритмы проецируют исходные изображения на HR сетку - что требует определения взаимного расположения изображений. Затем, HR картинка достраивается разными алгоритмами интерполяции. Так например, самая наивная интерполяция - *nearest neighbor* - присваивает неизвестному пикселю значение, равное значению ближайшего известного. Bose и Ahuja в статье [12] используют более сложную полиномиальную интерполяцию.

Irani и Peleg [13] предложили *Iterative Back Projection* (IBP) алгоритм, где HR изображение восстанавливается последовательным проецированием разницы между исходными LR изображениями и смоделированными LR изображениями (т.е. которые могли бы получиться при текущем HR изображении).

Как уже указывалось ранее, задача восстановления HR изображения при нашей модели - плохо обусловленная, и решение, во-первых, не единственное, а во-вторых, некоторые возможные решения не очень соответствуют реальности. Различные интерполяционные алгоритмы находят лишь одно, какое-то HR изображение, которое соответствует исходным LR. Поэтому зачастую делается еще один шаг - регуляризация полученного изображением с помощью разных методов удаление смазывания (*deblurring*) и шума (*denoising*).

Метод *Projection onto Convex Sets* (POSC), предложенный в статье [14] (см. также [15], [16]; существует реализация алгоритма на Matlab'e - [17], [18]), использует теоретико-множественный подход к SR задаче. При известной модели, для каждого LR изображения есть набор HR изображений, применив нашу модель к которым, мы получим LR изображение, отличающееся от нашего не больше, чем на некоторую фиксированную величину - ошибку. Алгоритм состоит в эффективном проецировании множества возможных HR изображений на такие множества, при постепенном уменьшении ошибки. Кроме того, регуляризацию HR изображения можно провести, если определить, какие изображения мы считываем хорошим, и делать проекцию в том числе и на это множество. Недостаток этого метода - большая вычислительная сложность.

3.4 Баесовские алгоритмы

Подход, моделирующий вероятностное распределение возможных HR изображений, при известных исходных LR. Два наиболее популярных подхода - *maximum likelihood* (ML) и *maximum a posterior* (MAP).

Алгоритм типа ML был впервые предложен Tom and Katsaggelos в статье [19]. Искомое HR изображение x^{ML} , при исходных LR изображениях y_1, y_2, \dots, y_p находится, как

$$x^{ML} = \arg \max_x p(y_1, y_2, \dots, y_p | x)$$

MAP-алгоритм описан в статьях - [20], [21]. Здесь ищется решение x^{MAP} такое, что

$$x^{MAP} = \arg \max_x p(x | y_1, y_2, \dots, y_p) = \arg \max_x p(y_1, y_2, \dots, y_p | x) p(x).$$

Т.е. необходимо определить априорную вероятность $p(x)$. Для этого используются модель *Markov Random Field* (MRF), подробнее описанная в [22], [23]. В статье [31] описан алгоритм, являющийся state-of-art на данный момент.

3.5 Алгоритмы с использованием регуляризации

Подходы, основанные на регуляризации, избавляются от неоднозначности в решениях, используя априорные знания о природе изображений. Например, делают предположения о вероятностном распределении HR изображений (как в случае с MAP).

Несколько последних статей ([28], [32]) используют для регуляризации подход *compressive sensing*, о котором будет написано далее. Кроме того, на этом же подходе основан *BM3D* фильтр ([33], [34]), алгоритм с использованием которого ([29], [35], [36]) на сегодняшний день является одним из state-of-art. Реализованный нами метод является продолжением применения *BM3D* фильтра для задачи SR. Поэтому мы коснемся темы *compressive sensing* и *BM3D* в следующей главе.

4 ВМ3D-фильтр

4.1 Compressive sensing

Compressive sensing - это парадигма, используемая в теории сигналов [32], основанная на том, что интересующий нас сигнал (в данном случае изображение) - это не случайный шум, а нечто, имеющее структуру и поэтому, в случае изображения, представление его в виде набора пикселей является избыточным. Т.е. изображения из реального мира это всего лишь некоторое, достаточно малое подпространство всех возможных изображений. Поэтому в некотором базисе оно должно иметь много нулевых коэффициентов. Обычно используют спектральные представления в виде коэффициентов ряда Фурье, синусов, косинусов или в виде коэффициентов некоторого вейвлетного разложения.

4.2 ВМ3D

ВМ3D (block matching 3d) техника, представленная в цикле статей [29],[33]-[36], позволяющая получить разреженное представление изображения в спектральном базисе, адаптирующемся под данное изображение. Основные идеи ВМ3D заключаются в том, что реальные изображения имеют повторяющиеся блоки, а так же в применении совместного 3D преобразования к группе блоков, которое дает более разреженное представление, чем 2D преобразования к каждому блоку по отдельности.

ВМ3D преобразование получается следующим образом:

1. Рассмотрим произвольный блок фиксированного размера (блок это квадратная часть изображения $k \times k$)
2. Для каждого такого блока находятся блоки, похожие на него. Похожесть определяется как l_2 норма разницы блоков.
3. Таким образом, мы получаем группы блоков. Для каждой группы блоков применяется 3D преобразование. Сначала для каждого блока применяется 2D преобразование (2D -DFT, -DCT, -DWT), а затем применяется межблочное 1D преобразование (например, преобразование Хаара). Получившиеся коэффициенты образуют спектр данного блока.
4. Объединение спектров всех групп и является результатом ВМ3D преобразование, т.е. ВМ3D спектром изображения.

Таким образом, VM3D преобразование задается набором групп. При известном наборе групп, можно применить обратное VM3D преобразование. Сначала нужно обратить 3D преобразование для каждой группы, а затем вернуть полученные блоки на их изначальные места на изображении. Блоки могут пересекаться, в этом случае значение пикселя определяется как взвешенная сумма значений этого пикселя на пересекающихся блоках. Вес можно определить по-разному. Например, в статье [34] предпочтение отдается блокам, спектр группы которых наиболее разрежен(*sparse*). Т.е. чем меньше ненулевых коэффициентов в спектре группы, из которой взят этот блок, тем больше его вес. VM3D фильтрация изображения заключается в нахождении его VM3D спектра, фильтрации коэффициентов спектра некоторым внутренним фильтром и преобразованием обратно из спектра в изображение. В качестве внутреннего фильтра может применяться *hard-thresholding*, который обнуляет коэффициенты, меньшие определенной границы. Более подробно сам VM3D и его применение в различных областях обработки изображений можно найти в статьях [29],[33]-[36].

5 Постановка задачи

Теперь перейдем к постановке задачи для нашего алгоритма. Наша модель наблюдений:

- для случая одного LR изображения

$$y = DBX + n = Ax + \sigma\epsilon,$$

где $\epsilon \sim N^n(0, 1)$, n - количество пикселей.

- для нескольких LR изображений

$$y_k = DBW_k x + \sigma_k \epsilon = A_k x + \sigma_k \epsilon$$

Как уже указывалось ранее, точного решения у этой системы может и не существовать. Поэтому часто задачу найти HR изображение, соответствующее данному набору LR изображений формулируют в виде наименьших квадратов

$$E_f(x) = \sum_k \|y_k - A_k x\|_2^2 \rightarrow \min$$

Для того, чтобы получить в результате изображение, больше соответствующее реальности, вводят дополнительный регуляризирующий функционал $E_r(x)$. Тогда задача восстановления HR изображения сводится к минимизации функционала

$$E_f(x) + E_r(x) \rightarrow \min$$

Руководствуясь парадигмой compressive sensing, член регуляризации должен выглядеть, как l_0 норма (т.е. количество ненулевых элементов) некоторого спектрального представления x . Пусть Φ - преобразование изображения в спектр. Тогда $E_r(x) = \|\Phi x\|_{l_0}$, и итоговая формулировка SR задачи выглядела бы так:

$$\sum_k \|y_k - A_k x\|_2^2 + \lambda \|\Phi x\|_{l_0} \rightarrow \min$$

В статье [34] для задачи удаления смазывания был предложен метод разделения минимизируемого функционала на два отдельных, причем качество работы алгоритма в этом случае было выше, чем когда минимизировался один функционал. В нашем случае первый будет отвечать за соответствие HR изображения x нашим наблюдениям y_k , а второй - за l_0 норму Φx .

Обозначим обратное преобразование из спектра за Ψ (в том случае, когда пространство спектра имеет большую размерность, то действует, как проекция на ближайшее изображение). Мы будем одновременно восстанавливать не только HR изображение X , но еще и его спектральное представление w . Причем, для большей устойчивости метода будем допускать неточное соответствие их друг другу.

В случае, когда LR изображение у нас всего одно, функционалы имеют вид

$$\begin{cases} L_{im}(x, w) = \frac{1}{2\sigma^2} \|y - Ax\|^2 + \frac{1}{\gamma} \|x - \Psi w\|^2 \\ L_{sp}(x, w) = \tau \|w\|_{l_0} + \|\Phi x - w\|^2 \end{cases},$$

где γ и τ - регуляризирующие параметры. Задача формулируется в виде

$$\begin{cases} x^* = \arg \min_x L_{im}(x, w^*) \\ w^* = \arg \min_w L_{sp}(x^*, w) \end{cases} \quad (1)$$

Когда восстановление HR изображения идет по нескольким LR изображениям, функционалы записываются в виде

$$\begin{cases} L_{im}(x, w) = \sum_k \frac{1}{2\sigma_k^2} \|y_k - A_k x\|^2 + \frac{1}{\gamma} \|x - \Psi w\|^2 \\ L_{sp}(x, w) = \tau \|w\|_{l_0} + \|\Phi x - w\|^2 \end{cases}$$

При использовании в качестве спектрального преобразования ВМЗД, разделение функционала на два и вариационный подход позволяют улучшить результат, что объясняется теоретически ([34] стр. 21 VIII Discussion) и подтверждается экспериментально.

6 Алгоритм

6.1 Случай single-frame

Решение системы может быть получено итеративным методом поочередной минимизации каждого из функционалов ([34])

$$\begin{cases} x_{t+1} = \arg \min_x L_{im}(x, w_t) \\ w_{t+1} = \arg \min_w L_{sp}(x_{t+1}, w) \end{cases}$$

Минимизация L_{im}

$$\begin{aligned} \frac{\partial L_{im}}{\partial X} &= 0 \\ -\frac{1}{\sigma^2} A^T (y - Ax) + \frac{1}{\gamma} (x - \Psi w) &= 0 \\ \left(\frac{\gamma}{\sigma^2} A^T A + I\right) x &= \Psi w + \frac{\gamma}{\sigma^2} A^T y \end{aligned}$$

В итоге, нахождение x_{t+1} свелось к численному решению линейного уравнения, чтобы можно эффективно реализовать, например, методом градиентного спуска.

Минимизация L_{sp}

$$w_{t+1} = Th_\tau\{\Phi x_{t+1}\},$$

где Th_τ - оператор hard-thresholding с параметром τ .

Таким образом итеративный алгоритм выглядит так

1. $\left(\frac{\gamma}{\sigma^2} A^T A + I\right) x_{t+1} = \Psi w_t + \frac{\gamma}{\sigma^2} A^T y$
2. $w_{t+1} = Th_\tau\{\Phi x_{t+1}\}$

Так как преобразование из спектра Ψw_t , операция $Th_\tau\{\Phi x_{t+1}\}$ делается внутри ВМЗД фильтра, то алгоритм можно переписать в виде

1. $\left(\frac{\gamma}{\sigma^2} A^T A + I\right) x_{t+1} = v_t + \frac{\gamma}{\sigma^2} A^T y$
2. $v_{t+1} = VM3D\{\Phi x_{t+1}\}$

6.2 Случай multi-frame

В случае нескольких исходных LR изображений алгоритм выглядит следующим образом

1. $(\sum_k \frac{\gamma}{\sigma_k^2} A_k^T A_k + I)x_{t+1} = \Psi w_t + \sum_k \frac{\gamma}{\sigma_k^2} A_k^T y_k$
2. $w_{t+1} = Th_\tau\{\Phi x_{t+1}\}$.

Если использовать BM3D фильтр

1. $(\sum_k \frac{\gamma}{\sigma_k^2} A_k^T A_k + I)x_{t+1} = v_t + \sum_k \frac{\gamma}{\sigma_k^2} A_k^T y_k$
2. $v_{t+1} = BM3D\{\Phi x_{t+1}\}$

7 Реализация

Мы реализовали предложенный алгоритм на Matlab. Исходный код находится по адресу [45]. В. Катковник предоставил для пользования авторскую реализацию VM3D фильтра.

В ходе реализации выяснилось, что напрямую задавать изображение как вектор, а преобразования хранить в виде матриц – невозможно, поскольку в этом случае размер матрицы будет очень большим. В самом деле, для изображения с разрешением $N \times M$, размер матрицы, задающей линейное преобразование для этого изображения, будет $NM \times NM$. Поэтому изображения мы хранили в виде матриц, а линейные преобразования задавали в виде функций – причем, в этом случае мы отдельно реализовывали функцию применения транспонированного преобразования. Это было возможным, поскольку мы рассматривали лишь простейшие случаи *subsampling* операторов.

В нашем случае *subsampling* оператор усреднял пиксели в блоке размера $k \times k$, а затем брал каждый k -ый пиксель.

В качестве модели *motion estimation* – геометрического преобразования между двумя изображениями – мы рассматривали простой сдвиг (по обеим осям). Для оценки параметров сдвига мы пробовали несколько существующих алгоритмов (библиотека FAIR - [38]) Лучшие результаты показал простой алгоритм на основе FFT (например, этот алгоритм реализован в [40]).

Параметры нашего алгоритма – коэффициент $k = \frac{\gamma}{\sigma^2}$ и уровень шума σ – мы подбирали для набора изображений. В наших примерах $k = 10$, а σ мы оценивали по уровню шума в входных изображениях, в случае реальных тестовых данных, или присваивали известному уровню шума, в случае синтетических.

8 Эксперименты

Для сравнения результатов экспериментов, в тех случаях, когда нам было доступно исходное HR изображения, результат работы алгоритма мы сравнивали по метрике *Peak signal-to-noise ratio* (PSNR) - [51]. Этой метрикой пользуются для сравнения изображений практически во всех работах по SR. Для двух изображений $Im1$ и $Im2$ метрика PSNR будет равна

$$PSNR(Im1, Im2) = 10 \log_{10} \sum_{(i,j)} (Im1(i, j) - Im2(i, j))^2$$

8.1 Single-frame SR

Здесь мы продемонстрируем возможности предложенного алгоритма по увеличению одного изображения. В качестве первого теста мы взяли известное изображение Cameraman 256x256, уменьшенное в 4 раза с добавлением шума с дисперсией 8. Результат применения представлен на рисунке 2.

8.2 Multi-frame SR

Для оценки возможностей нашего алгоритма первые тесты для нескольких изображений мы провели на синтетически полученных данных. Для этого мы взяли изображение “Text”, сгенерировали для него 9 циклических сдвигов (на 0,1 и 2 пикселя по строкам и столбцам). Затем уменьшили изображение в 3 раза по обоим осям равномерно усредняя пиксели в квадратах 3x3, и добавили случайный шум с дисперсией 2. Алгоритм для multi-frame SR научной группы из Tampere University of Technology [29], а так же те алгоритмы, с которыми они проводили сравнение - [41], тестировались на этих же данных. По количественной метрике PSNR и просто по визуальному сравнению получившихся изображений наш алгоритм превосходит методы, описанные в статьях [41], [42], [43], [50]. Лучший из них получает изображение с PSNR = 18.5dB, в то время, как наш результат – 22dB (см таблицу (1)). Фрагменты полученных изображений показаны на рисунке 3.

Стоит отметить, что алгоритмы, с которыми мы сравнивали, не используют *motion estimation*, то есть решают несколько другую задачу, чем рассматриваемая здесь, в которой геометрическое преобразование W_k неизвестно. Предложенный нами метод нуждается в оценке глобального геометрического преобразования и довольно чувствителен к ошибкам в измерении. Второй набор синтетических тестовых данных, описанных в работе [29] был получен похожим образом



а) Исходное изображение



б) Уменьшенное в 4 раза



с) Восстановленное из исходного изображения

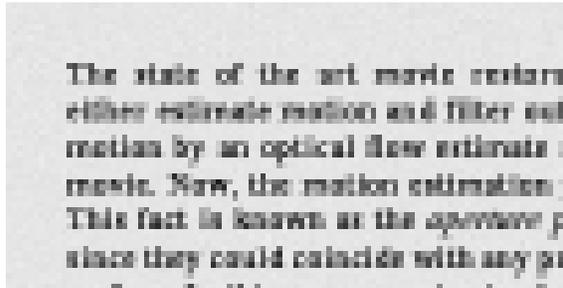


с) Восстановленное из 4 LR изображений

Рис. 2: Cameraman

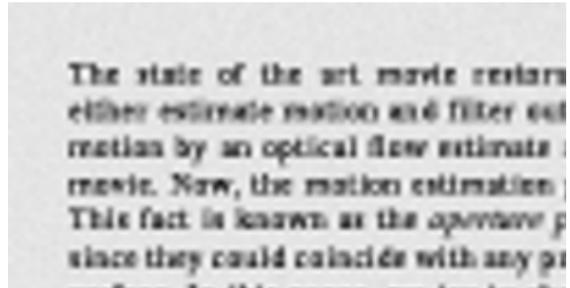
Алгоритм	PSNR (dB)
Nearest-neighbour интерполяция [41]	13.5 dB
Lanczos интерполяция [41]	13.8 bB
Shift-and-add алгоритм [50]	18.4 dB
Алгоритм [41]	18.4 dB
Наш метод	22.1 dB

Таблица 1: Сравнение результатов работы алгоритмов для набора "Text"



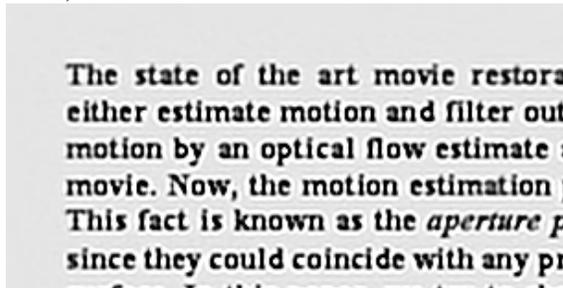
The state of the art movie restoration either estimate motion and filter out motion by an optical flow estimate movie. Now, the motion estimation This fact is known as the *aperture p* since they could coincide with any p

a) Одно из 9-ти LR изображений



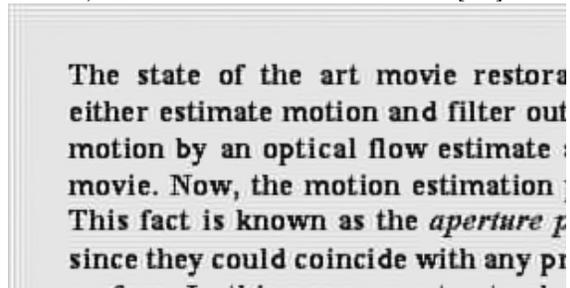
The state of the art movie restoration either estimate motion and filter out motion by an optical flow estimate movie. Now, the motion estimation This fact is known as the *aperture p* since they could coincide with any p

b) Lanczos интерполяция - [41]



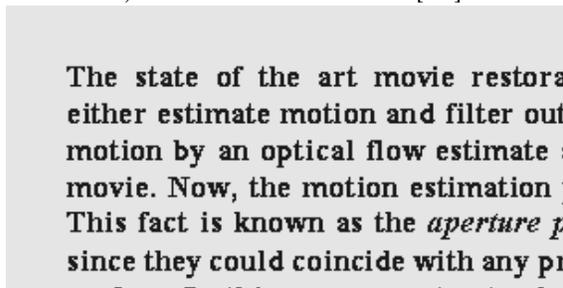
The state of the art movie restoration either estimate motion and filter out motion by an optical flow estimate movie. Now, the motion estimation This fact is known as the *aperture p* since they could coincide with any p

d) Non-Local Means - [41]



The state of the art movie restoration either estimate motion and filter out motion by an optical flow estimate movie. Now, the motion estimation This fact is known as the *aperture p* since they could coincide with any p

e) Наш алгоритм

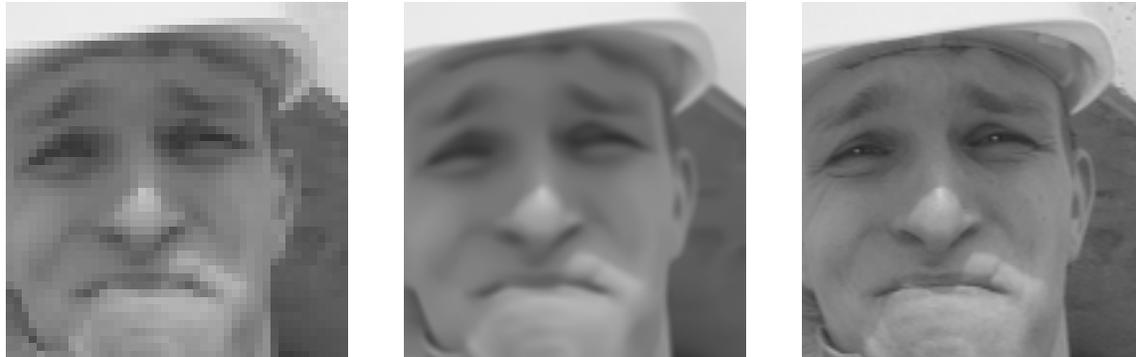


The state of the art movie restoration either estimate motion and filter out motion by an optical flow estimate movie. Now, the motion estimation This fact is known as the *aperture p* since they could coincide with any p

f) Исходное изображение

Рис. 3: Text

из другого исходного изображения – первого кадра видеозаписи “Foreman” (см. рисунок 4)



а) Одно из 9-ти LR изображений

б) После обработки нашим алгоритмом

с) Исходное изображение

Рис. 4: Foreman

К сожалению, результаты восстановления изображений в статье [29] не открыты, поэтому визуально сравнить их нельзя. Однако, авторы [29] сравнивают получившееся изображение с исходным по метрике PSNR. Максимальное значение среди алгоритмов в статье [29] – 34,9dB. В то время как значение этой метрики для нашего изображения – 35,8dB.

Как уже отмечалось ранее, в случае, когда геометрические отношения между LR изображениями известны и глобальны (т.е., например, простой сдвиг) – как в случае синтетических данных – наш алгоритм показывает высокое качество работы, сопоставимое с существующими алгоритмами. Однако на реальных данных, результат работы не настолько качественен. Для примера реальной последовательности, мы взяли кадры “Library” из Оксфордской тестовой базы для SR алгоритмов [44] (на данный момент, эта база – один из немногих существующих benchmark’ов по проблеме superresolution). На рисунке 5 показаны фрагменты изображения с надписью

Одно из преимуществ нашего алгоритма – устойчивость к шуму на исходных изображениях. На рисунке 6 приведен пример работы нашего алгоритма на тестовых данных “Foreman”, но с уровнем шума 20 (8% от диапазона значений интенсивности 0..255)

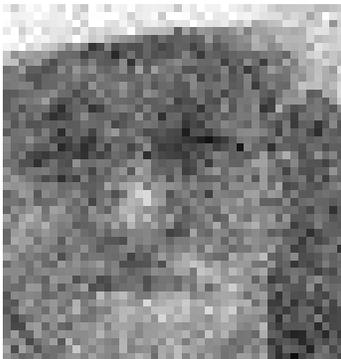


а) Одно из исходных изображений



б) После обработки нашим алгоритмом

Рис. 5: Library



а) Одно из 9-ти LR изображений



б) После обработки нашим алгоритмом



с) Исходное изображение

Рис. 6: Foreman с сильно зашумленный

9 Итоги

В ходе написания дипломной работы были изучены существующие подходы к superresolution. Был разработан и реализован алгоритм superresolution на основе вариационного подхода с использованием VM3D фильтра. Был проведен ряд экспериментов для сравнения предложенного подхода с существующими и в некоторых экспериментах он показал результаты лучше, чем методы, представленные для сравнения, что указывает на перспективность этого подхода.

В будущем планируется продолжить работу над оптимизацией алгоритма. В частности, есть идея использовать методы стохастической оптимизации. Кроме того, можно попробовать адаптировать алгоритм для видео - восстанавливать не одно изображение, а повышать разрешение сразу всех изображений.

Список литературы

- [1] Tsai, R.Y., Huang, T.S. "Multiframe image restoration and registration". In: Huang, T.S. (ed.) *Advances in Computer Vision and Image Processing*. JAI Press Inc., London (1984)
- [2] Rhee, S., Kang, M.G. "Discrete cosine transform based regularized high-resolution image reconstruction algorithm". *Opt. Eng.* 38, 1348–1356 (1999)
- [3] Nguyen, N., Milanfar, P. "A wavelet-based interpolation-restoration method for superresolution". *Circuits Syst. and Signal Process.* 19, 321–338 (2000)
- [4] Ei-Khamy, S.E., Hadhoud, M.M., Dessouky, M.I., Salam, B.M., Ei-Samie, F.E. "Regularized super-resolution reconstruction of images using wavelet fusion". *Opt. Eng.* 44, 097001.1–097001.10 (2005)
- [5] Ei-Khamy, S.E., Hadhoud, M.M., Dessouky, M.I., Salam, B.M., Ei-Samie, F.E. "Wavelet fusion: a tool to break the limits on LMMSE image super-resolution". *Int. J. Wavel. Multiresolut. Inf. Process.* 4, 105–118 (2006)
- [6] Ji, H., Fermuller, C. "Wavelet-based super-resolution reconstruction: theory and algorithm". In: *Proceedings of the European Conference on Computer Vision*, pp. 295–307. Graz, Austria (2006)
- [7] Ji, H., Fermuller, C. "Robust wavelet-based super-resolution reconstruction: theory and algorithm". *IEEE Trans. Pattern Anal. Mach. Intell.* 31, 649–660 (2009)
- [8] Katsaggelos A.K., Molina R., Mateos J. "Super Resolution of Images and Video (2007)
- [9] Park S. C., Park M. K., Kang M. G. "Super-resolution image reconstruction: a technical overview *IEEE Signal Processing Magazine*, 20(3):21–36, (2003).
- [10] Tian J., Ma K.-K. "A survey on super-resolution imaging. *Signal, Image and Video Processing* (2011)
- [11] Babu S., Murthy S., "A Survey on the methods of Super-resolution Image Reconstruction", *International Journal of Computer Applications*, vol. 15, issue 2, pp. 1-6

- [12] Bose N.K., Ahuja N.A. "Superresolution and noise filtering using moving least squares IEEE Trans. Image Process. 15, 2239–2248, (2006)
- [13] Irani, M., Peleg, S. "Improving resolution by image registration". CVGIP Graph. Models Image Process. 53, 231–239, (1991)
- [14] Katsaggelos A.K. "A Multiple Input Image Restoration Approach". Journal of Visual Communication and Image Representation, 1:93–103, (1990)
- [15] Patti A.J., Tekalp A.M. "Super resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time". IEEE Trans. Image Process. 6, 1446–1451, (1997)
- [16] Elad M. and Feuer A., "Restoration of a Single Superresolution Image from Several Blurred, Noisy, and Undersampled Measured Images
- [17] <http://www.codeforge.com/read/209555/pocs.html>
- [18] <http://www.mathworks.com/matlabcentral/fileexchange/30488-superresolution-demo>
- [19] Tom B.C., Katsaggelos A.K. "Reconstruction of a high-resolution image by simultaneous registration, restoration, and interpolation of low-resolution images". In: Proceedings of the IEEE International Conference on Image Processing, pp. 539–542. Washington, (1995)
- [20] Hardie R.C., Barnard K.J., Armstrong E.E. "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images". IEEE Trans. Image Process. 6, 1621–1633, (1997)
- [21] Tian, J., Ma, K.-K. "Stochastic super-resolution image reconstruction". J. Vis. Commun. Image Represent. 21, 232–244, (2010)
- [22] Li, S.Z. "Markov Random Field Modeling in Computer Vision". Springer, New York, (1995)
- [23] Rue H. "Gaussian Markov Random Fields: Theory and Applications". Chapman & Hall, Boca Raton, (2005)
- [24] Hertzmann A., Jacobs C.E., Oliver N., Curless B., Salesin D.H. "Image analogies". In: Proceedings of the SIGGRAPH, pp. 327–340. Los Angeles, (2001)

- [25] Chang H., Yeung D.Y., Xiong Y. "Super-resolution through neighbor embedding". In: Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pp. 275–282. Washington, D.C., (2004)
- [26] Freeman W.T., Pasztor E.C., Carmichael O.T. "Learning low-level vision". *Int. J. Comput. Vis.* 40, 25–47, (2000)
- [27] Pickup L.C., Roberts S.J., Zisserman A. "A sampled texture prior for image super-resolution". In: Thrun, S., Saul, L., Scholkopf, B. (eds.) *Advances in Neural Information Processing Systems*, MIT Press, Cambridge, (2003)
- [28] Edeler T., Ohliger K., Hussmann S., Mertins A. "Multi image super resolution using compressed sensing"
- [29] Danielyan A., Foi A., Katkovnik V., Egiazarian K. "Image and video super-resolution via spatially adaptive block-matching"
- [30] Lyndsey C. "Machine Learning in Multi-frame Image Super-resolution Robotics Research Group Department of Engineering Science University of Oxford, Michaelmas Term, (2007)
- [31] Liu C., Sun D. "A Bayesian Approach to Adaptive Video Super Resolution". *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [32] Divekar A., Ersoy O. "Theory and Applications of Compressive Sensing"
- [33] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering" *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095,(2007)
- [34] Danielyan A., Katkovnik V., Egiazarian K. "BM3D frames and variational image deblurring"
- [35] Danielyan A., Foi A., Katkovnik V., Egiazarian K. "Image upsampling via spatially adaptive block-matching filtering"
- [36] Danielyan A., Foi A., Katkovnik V., Egiazarian K. "Spatially adaptive filtering as regularization in inverse imaging: compressive sensing, super-resolution and upsampling"
- [37] http://en.wikipedia.org/wiki/Motion_estimation

- [38] Modersitzki J. "FAIR: Flexible Algorithms for Image Registration <http://www.ec-securehost.com/SIAM/FA06.html>
- [39] M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing". SpringerPress, (2010)
- [40] <http://www.mathworks.com/matlabcentral/fileexchange/30488-superresolution-demo>
- [41] Protter M., Elad M., Takeda H., Milanfar P. "Generalizing the Non-Local-Means to Super-Resolution Reconstruction", to appear in IEEE Trans. Image Process., (2008). Test sequences available online at <http://www.cs.technion.ac.il/~matanpr/NLM-SR>
- [42] Wolberg G., "Digital Image Warping" , IEEE Computer Society Press, Washington, USA, (1990)
- [43] Turkowski K., "Filters for common resampling tasks", Graphics Gems, Academic Press, (1990)
- [44] http://www.robots.ox.ac.uk/~vgg/data/super_resolution/sr_sequences.tgz
- [45] <https://github.com/vpas/IDSR-BM3D>
- [46] Liu J., Zhang T.J., "Recognition of the blurred image by complex moment invariants Pattern Recognition Letters, (2005)
- [47] Aizenberg I., Aizenberg N., Bregin T., Butakov C., Farberov E., Merzlyakov N., Milyukova O., "Blur Recognition on the Neural Network based on Multi-Valued Neurons
- [48] Premaratne P. "Image blur recognition using under-sampled discrete Fourier transform"
- [49] <http://people.csail.mit.edu/celiu/denoise/estnoise/>
- [50] Farsiu S., Robinson D., Elad M., Milanfar P., "Robust shift and add approach to superresolution", Proc. SPIE Conf. Applications of Digital Signal and Image Processing, pp. 121- 130, (2003)
- [51] http://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio