

Классификация методов повышения разрешения изображения по нескольким кадрам

П.Ю. Сироткина

sirotkinapyu@student.bmstu.ru

МГТУ им. Н.Э. Баумана, Москва, Россия

Аннотация

Статья посвящена обзору методов повышения разрешения изображения (т. н. задаче суперразрешения) по нескольким кадрам одного и того же объекта. Кратко рассмотрена предметная область и основные подходы к решению задачи: методы на основе интерполяции, методы реконструкции с регуляризации, методы на основе теории множеств, примеро-ориентированные методы (выделение патчей), методы оптимизации сверточных нейронных сетей. Проведен критический анализ методов повышения разрешения изображения по нескольким кадрам с определением областей наиболее эффективного использования на основе следующих критериев: вычислительная сложность, качество обработки, необходимость пост- или предобработки. Сделан вывод об области применимости рассмотренных методов. Для достижения наилучшего качества обработки целесообразно применять методы, основанные на сверточных нейронных сетях.

Ключевые слова

цифровое изображение, обработка сигналов, суперразрешение, сверточные нейронные сети, интерполяция, реконструкция, регуляризация, теория множеств, патчи

Введение

Для многих задач, связанных с областью цифровой обработки сигналов, высокое разрешение изображения является ключевым аспектом, позволяющим проводить более качественный анализ и последующую обработку полученной информации. Изображение высокого разрешения содержит больше сведений о деталях, учет которых может быть критически важен во многих областях, таких как медицина, астрономия и многие другие.

Одним из способов решения рассматриваемой проблемы является использование более качественной техники, однако при таком подходе быстро достигается лимит стоимости, веса и размеров оборудования, что делает этот подход нецелесообразным с точки зрения использования ресурсов. Также понижение качества часто может быть связано с физическими ограничениями (например, при регистрации спутников).

В связи с этим существует широкое разнообразие альтернативных методов повышения разрешения изображения, систематизация которых является актуальной задачей для систем фотосъемки.

Модель наблюдения изображения низкого разрешения

В контексте данной работы под термином «кадр» понимается предоставляемое цифровое изображение низкого разрешения (далее НР), используемое для синтеза изображения высокого разрешения (далее ВР). Каждый кадр является изображением одного и того же объекта, незначительно смещенного относительно других кадров из набора.

На рисунке 1 представлена модель наблюдения изображения низкого разрешения.



Рисунок 1 — Модель наблюдения изображения низкого разрешения

Задача повышения разрешения изображения по одному или нескольким кадрам низкого разрешения получила название суперразрешения (англ. super resolution). В общем случае эта задача является некорректной, т. к. в реальных задачах сведения об искажениях изображения заведомо неизвестны.

Представим желаемое изображение ВР размером $N = L_1 N_1 \times L_2 N_2$ как вектор $x = [x_1, x_2, \dots, x_N]^T$. Параметры L_1 и L_2 представляют собой коэффициенты понижения дискретизации в наблюдении для горизонтального и вертикального направлений соответственно. Таким образом, каждое изображение НР имеет размерность $M = N_1 \times N_2$.

Аналогично представим каждое изображение НР в виде вектора $y_k = [y_{k_1}, y_{k_2}, \dots, y_{k_M}]^T$, где $k = \overline{1, p}$, p — размер входного набора.

Для получения изображения высокого разрешения x методом суперразрешения рассматривается следующая система уравнений [1]:

$$A_k x = y_k, \quad (1)$$

где оператор $A_k x$ в общем случае представляется в виде выражения (2):

$$A_k x = D B_k M_l x + n_k, \quad (2)$$

где:

- D — матрица движения фотосистемы (сдвиг, поворот и т. д.);
- B_k — матрица размытия (может быть вызвано оптической системой, относительным движением между фотосистемой и сценой, а также функцией рассеяния в датчике фотосистемы);
- M_l — матрица понижения размерности (генерирует изображение НР путем наложения, т.е. алиасинга);
- n_k — вектор шума.

Возможна более короткая форма записи:

$$A_k x = W_k x + n_k = y_k, \quad (3)$$

где матрица W_k размером $(N_1, N_2)^2 \times L_1 N_1 L_2 N_2$ представляет собой вклад пикселей изображения x в изображение y_k посредством размытия, движения и субдискретизации.

Частотные методы

Данная группа методов суперразрешения ориентирована на увеличение разрешения изображения посредством анализа частотных характеристик кадров. Классическим математическим аппаратом для подобных задач является Фурье-анализ.

Этот подход основывается на следующих принципах:

1. Свойство сдвига преобразования Фурье: $f(x - x_0) = e^{-i\omega x_0} \hat{f}(\omega)$.
2. Взаимосвязь между непрерывным преобразованием Фурье исходного изображения ВР и дискретным преобразованием Фурье наблюдаемых изображений НР.
3. Частотный диапазон изображения ограничен.

Пусть $x(t_1, t_2)$ — непрерывное изображение ВР, $X(\omega_1, \omega_2)$ — соответствующее ему непрерывное преобразование Фурье, $y_k[n_1, n_2]$ — k -ое изображение НР, $Y(\Omega_1, \Omega_2)$ — соответствующее ему дискретное преобразование Фурье. Глобальные сдвиги,

которые являются единственным движением, рассматриваемым в подходе частотной группы методов, порождают следующее уравнение для изображений НР:

$$x_k(t_1, t_2) = x(t_1 + \delta_{k_1}, t_2 + \delta_{k_2}), \quad (4)$$

где δ_{k_1} и δ_{k_2} — величины субпиксельных сдвигов между кадрами.

Согласно свойству сдвига преобразования Фурье:

$$X_k(\omega_1, \omega_2) = \exp[i2\pi(\delta_{k_1}\omega_1 + \delta_{k_2}\omega_2)] X(\omega_1, \omega_2). \quad (5)$$

Смещенное изображение $x_k(t_1, t_2)$ дискретизируется с периодами T_1 и T_2 для создания наблюдаемого изображения НР $y_k[n_1, n_2]$ [2].

Таким образом, взаимосвязь между $X_k(\omega_1, \omega_2)$ и $Y(\Omega_1, \Omega_2)$ может быть описана (6):

$$Y(\Omega_1, \Omega_2) = \frac{1}{T_1 T_2} \sum_{n_1=0}^{L_1-1} \sum_{n_2=0}^{L_2-1} X_k \left[\frac{2\pi}{T_1} \left(\frac{\Omega_1}{N_1} + n_1 \right), \frac{2\pi}{T_2} \left(\frac{\Omega_2}{N_2} + n_2 \right) \right]. \quad (6)$$

Учитывая, что функция X финитна, матрично-векторная форма уравнения может быть записана (7):

$$Y = \Phi X. \quad (7)$$

Пространственные методы

Регистрация – интерполяция – восстановление

Линейный подход к решению задачи суперразрешения на основе интерполяции является наиболее тривиальным, однако в связи с этим имеет ряд существенных недостатков: возникновение алиасинга (эффект «лесенки»), размытия и эффекта Гиббса. На рисунке 2 представлен алгоритм суперразрешения на основе рассматриваемого подхода [3].

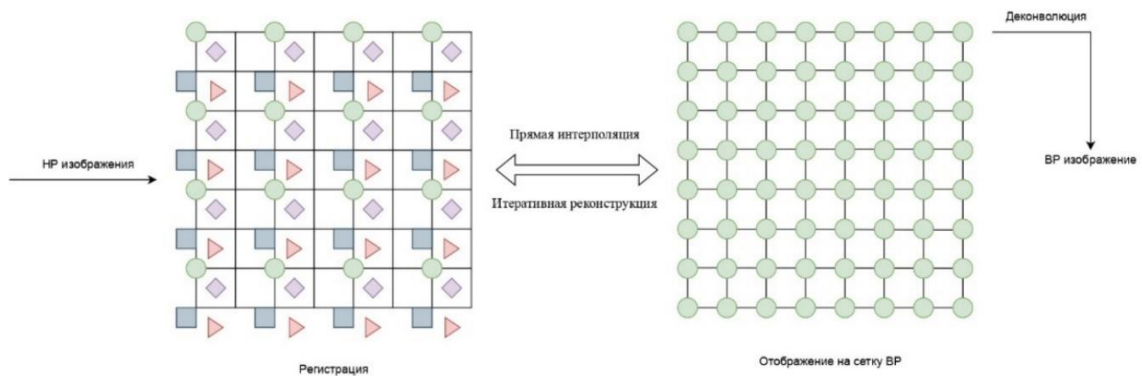


Рисунок 2 — Алгоритм суперразрешения на основе подхода регистрация - интерполяция - восстановление

Примеро-ориентированные методы

Данная группа методов реализовывает подход, основанный на глубоком обучении [4]. Создается набор изображений (изображения ВР подвергаются деградации (см. рисунок 1) и формируется пара (ВР;НР)), который затем используется для анализа деталей, соответствующих отдельным областям цифрового изображения. Выявленные взаимосвязи используются для прогнозирования мелких деталей на других изображениях.

Алгоритм суперразрешения осуществляется путем разбиения кадра на патчи и составления изображения ВР путем перебора соответствующих патчей в базе данных. На рисунке 3 представлен пример сравнения патчей.

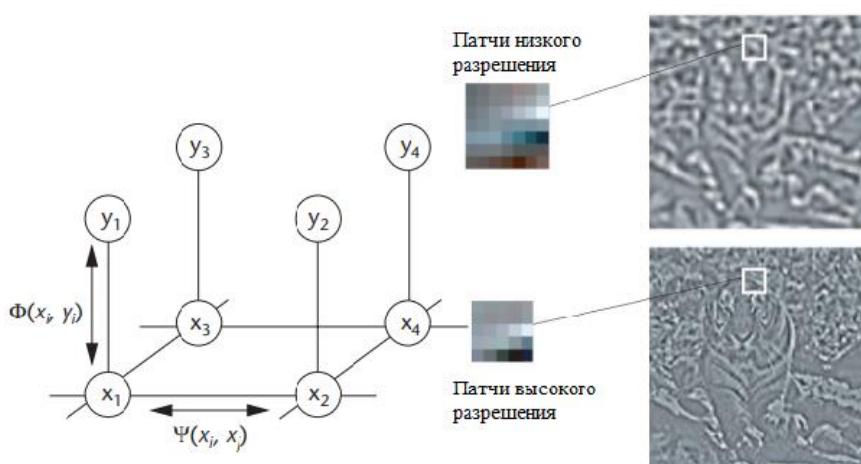


Рисунок 3 — Пример сравнения патчей

На рисунке 4 представлен пример базы данных патчей, предложенный в работе «Example-Based Super-Resolution» [5].

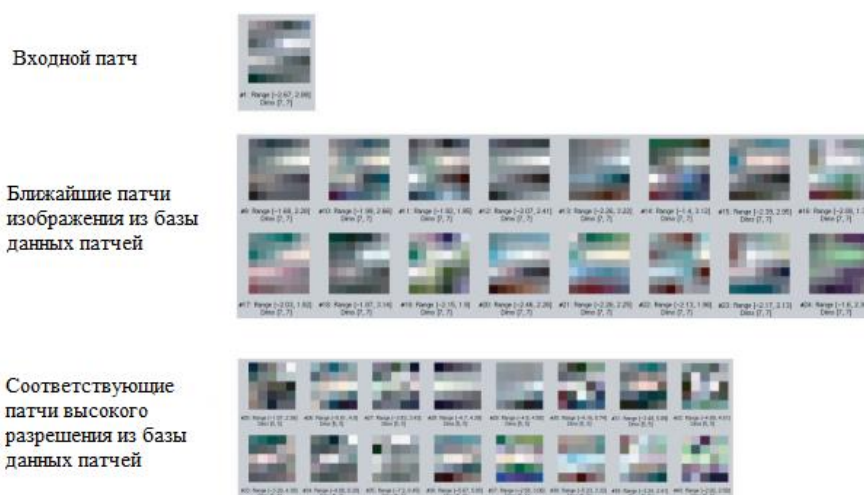


Рисунок 4 — Пример базы данных патчей

Реконструкция с регуляризацией

Задача реконструкции (решение обратной задачи) изображения ВР является некорректной в виду недостаточного количества информации о деталях и отсутствия априорной информации о функции размытия в общем случае. Методы регуляризации сводят некорректно поставленную задачу к корректной посредством введения дополнительных ограничений.

В данном разделе представлены детерминированные и стохастические (статистические) методы реконструкции.

Детерминированные регуляризационные методы

Оценив параметры регистрации, можно полностью уточнить модель наблюдения. Рассматриваемый подход решает обратную задачу, используя априорную информацию о решении, которая может быть использована для корректной постановки задачи.

Рассмотрим формулировку метода наименьших квадратов [6], выбрав x для минимизации Лангранжиана:

$$\sum_{k=1}^p \|y_k - W_k\|^2 + \alpha \|Cx\|^2, \quad (8)$$

где оператор C — фильтр верхних частот, $\|\cdot\|$ — l_2 -норма, α — параметр регуляризации (множитель Лангранжа), контролирующий компромисс между точностью данных. Чем больше значение α , тем более гладкое решение.

Основные методы детерминированной итеративной регуляризации решают следующую задачу:

$$\left[\sum_{k=1}^p W_k^T W_k + \alpha C^T C \right] \hat{x} = \sum_{k=1}^p W_k^T y_k, \quad (9)$$

что приводит к формулировке следующей рекуррентной формулы для \hat{x} :

$$\hat{x}^{n+1} = \hat{x}^n + \beta \left[\sum_{k=1}^p W_k^T (y_k - W_k \hat{x}^n) - \alpha C^T C \hat{x}^n \right], \quad (10)$$

где β — параметр сходимости.

Стохастические регуляризационные методы

Данная группа методов основана на байесовской статистике и используется, когда можно установить апостериорную функцию плотности исходного изображения. Оценка x с помощью метода максимального правдоподобия максимизирует апостериорную функцию плотности $P(x|y_k)$ относительно x :

$$x = \operatorname{argmax} P(x | y_1, y_2, \dots, y_p). \quad (11)$$

Запишем задачу максимизации, применив логарифмирование, а также теорему Байеса к условной вероятности:

$$x = \operatorname{argmax} \{ \ln P(y_1, y_2, \dots, y_p | x) + \ln P(x) \}. \quad (12)$$

Здесь и априорная модель изображения $P(x)$, и условная плотность $P(y_1, y_2, \dots, y_p | x)$ будут определяться априорной информацией об изображении ВР и статической информацией о шуме. Поскольку этот метод учитывает априорные ограничения, то предоставляется стабильная оценка изображения ВР.

Главным преимуществом байесовской модели является использование предварительной модели изображения с сохранением краев изображения.

Методы резолуции на основе теории множеств

Согласно методу POCS (англ. Projection Onto Convex Sets) [7, 8], учет априорных знаний в решение можно интерпретировать как наложение ограничения на то, чтобы решение было членом замкнутого выпуклого множества C_i , которое определяется как набор векторов, удовлетворяющих определенному свойству. Ограничение представляет собой предположительные сведения о гладкости, структуре, текстуре и других характеристиках изображения.

На рисунке 5 представлена иллюстрация метода суперрезолуции на основе теории множеств.

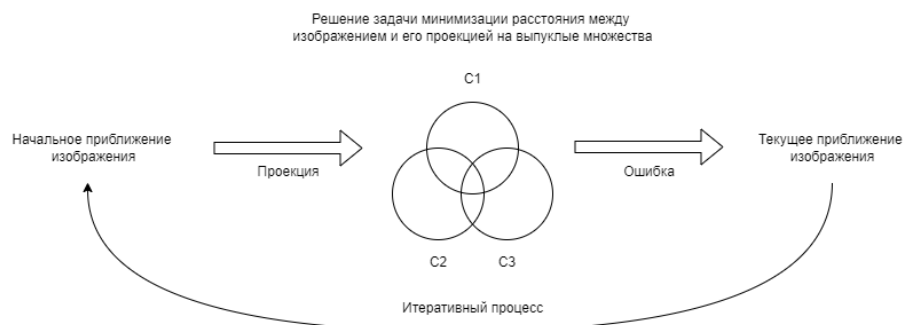


Рисунок 5 — Метод суперрезолуции на основе теории множеств

Методы на основе сверточных нейронных сетей

Сверточные нейронные сети в области распознавания образов на изображениях и их классификации нашли широкое применение, что объясняется достаточно высоким качеством решения подобных задач. Архитектура сети получила название в виду использования операции свертки.

На рисунке 6 представлена типовая архитектура сверточной нейронной сети.

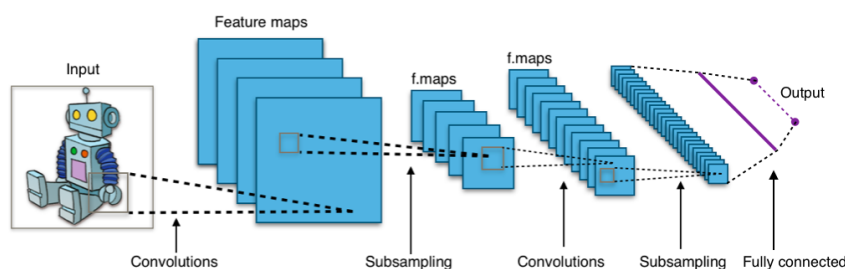


Рисунок 6 — Типовая архитектура сверточной нейронной сети

Как можно увидеть, СНС состоит из 3 типов слоев [9, 10, 11]:

- слой свертки — основной блок СНС, производящий свертку входной матрицы с ядром свертки. Количество ядер равно количеству карт признаков, выделяемых на изображении;
- слой пуллинга, или субдескрипторизация — данный слой принимает результат свертки предыдущего слоя в вид матрицы и сжимает данную матрицу с целью выделения низкоуровневых признаков;
- полносвязный слой — на данный слой подается одномерный вектор от стоящего перед ним слоя, причем данный вектор получен из матрицы путем записи ее элементов построчно в одну строку.

Оптимизация глубоких нейронных сетей

Одна из ключевых задач в тренировке глубоких нейронных сетей — оптимизация их параметров. Вследствие большого совокупного числа параметров на всех слоях нейронной сети применение методов второго порядка (таких как BFGS, SR1 и других квазиньютоновских методов) крайне затруднительно, а методы нулевого порядка в большинстве случаев не позволяют найти качественное решение за приемлемое время.

По этой причине для оптимизации глубоких и в том числе сверточных нейронных сетей, наиболее распространены методы первого порядка, требующие лишь знания градиента сети как функции от ее параметров.

Проекция — аппроксимация — восстановление

Исходя из приведенных выше соображений, были разработаны методы, основанные на подходе проекции-аппроксимации-восстановления [12, 13].

Под подходом проекции — аппроксимации — восстановления понимается следующая процедура (далее обозначаемая акронимом ПАВ):

1. Проекция: проекция нескольких последних точек из оригинального пространства в пространство меньшей размерности посредством умножения на специальным образом построенную прямоугольную матрицу с ортонормальными строками.
2. Аппроксимация: построение квадратичного полинома, аппроксимирующего полученные проекции точек и соответствующие им значения целевой функции.
3. Восстановление: аппроксимация параметров целевой функции (например, Гессиана, точки минимума) в оригинальном пространстве на основе полученного полинома в пространстве меньшей размерности и прямоугольной матрицы из первого пункта.

Классификация и сравнительный анализ рассмотренных методов

На рисунке 7 представлены критерии, согласно которым можно классифицировать рассматриваемые методы повышения разрешения изображения.

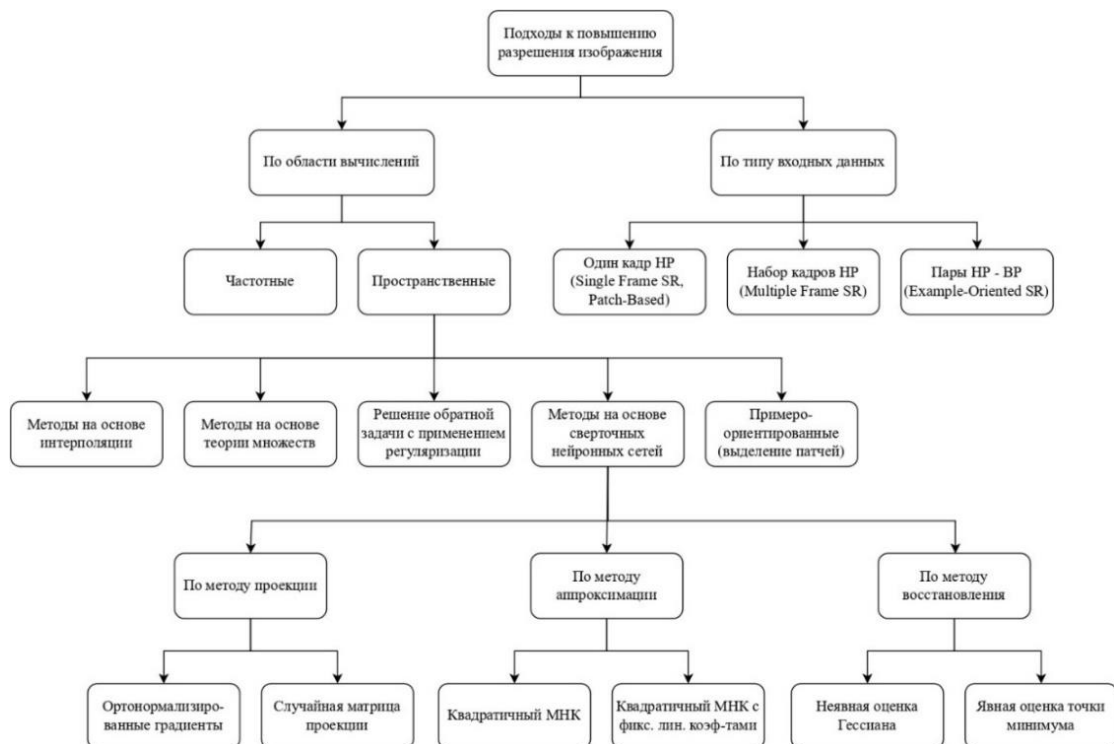


Рисунок 7 — Классификация подходов повышения разрешения изображения

При сравнительном анализе методов повышения разрешения изображения были выделены следующие критерии: вычислительная сложность метода, качество обработки и необходимость пост- или предобработки.

В таблице 1 представлен сравнительный анализ рассмотренных методов.

Таблица 1 — Сравнительный анализ рассмотренных методов

Метод	Критерий		
	Вычислит. сложность	Качество обработки	Необходимость доп. обработки
Частотные методы	Низкая	Низкое	Нет
Интерполяционные методы	Низкая	Очень низкое	Нет
Методы теории множеств	Средняя	Низкое	Да
Регуляризация	Средняя	Среднее	Нет
Выделение патчей	Очень высокая	Высокое	Да
Сверточные нейронные сети	Очень высокая	Очень высокое	Да

Заключение

Выбор метода суперразрешения зависит от конкретной задачи и имеющихся вычислительных и временных ресурсов. Методы на основе глубокого обучения часто дают наилучшие результаты, но они требуют большого объема данных и мощных вычислительных ресурсов для обучения сетей. Аналитические методы более вычислительно эффективны, но могут не давать такие высококачественные результаты, а также не учитывать структурные особенности изображений.

На основе проведенной классификации и сравнительного анализа для дальнейшей разработки было выбрано направление оптимизации сверточных нейронных сетей для повышения разрешения изображения по нескольким кадрам, т.к. в рамках поставленной задачи ключевым фактором является качество получаемого результата, другие же факторы являются второстепенными.

Список литературы

[1] Насонов А. В., Крылов А. С. Быстрое суперразрешение изображений с использованием взвешенной медианной фильтрации //Труды. – 2010. – С. 101-104.

- [2] Park S. C., Park M. K., Kang M. G. Super-resolution image reconstruction: a technical overview //IEEE signal processing magazine. – 2003. – Т. 20. – №. 3. – С. 21-36.
- [3] Кокошкин А. В. и др. Оценка ошибок синтеза изображений с суперразрешением на основе использования нескольких кадров //Компьютерная оптика. – 2017. – Т. 41. – №. 5. – С. 701-711.
- [4] Wang Q., Tang X., Shum H. Patch based blind image super resolution //Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1. – IEEE, 2005. – Т. 1. – С. 709-716.
- [5] Freeman W. T., Jones T. R., Pasztor E. C. Example-based super-resolution //IEEE Computer graphics and Applications. – 2002. – Т. 22. – №. 2. – С. 56-65.
- [6] Katsaggelos A. K. Digital image restoration. – Springer Publishing Company, Incorporated, 2012.
- [7] Fan C. et al. POCS Super-resolution sequence image reconstruction based on improvement approach of Keren registration method //Sixth International Conference on Intelligent Systems Design and Applications. – IEEE, 2006. – Т. 2. – С. 333-337.
- [8] Stark H., Oskoui P. High-resolution image recovery from image-plane arrays, using convex projections //JOSA A. – 1989. – Т. 6. – №. 11. – С. 1715-1726.
- [9] Бредихин А. И. Алгоритмы обучения сверточных нейронных сетей //Вестник Югорского государственного университета. – 2019. – №. 1 (52). – С. 41-54.
- [10] Umehara K., Ota J., Ishida T. Application of super-resolution convolutional neural network for enhancing image resolution in chest CT //Journal of digital imaging. – 2018. – Т. 31. – С. 441-450.
- [11] Dong C. et al. Learning a deep convolutional network for image super-resolution //Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part IV 13. – Springer International Publishing, 2014. – С. 184-199.
- [12] Senov A. Projective approximation based quasi-Newton methods// In: Proc. of International Workshop on Machine Learning, Optimization, and Big Data. 2017. P. 29–40.
- [13] Сенов А. А. Глубокое обучение в задаче реконструкции суперразрешения изображений //Стохастическая оптимизация в информатике. – 2017. – Т. 13. – №. 2. – С. 38-57.

Сироткина Полина Юрьевна – студент, МГТУ им. Н.Э. Баумана, кафедра «Программное обеспечение ЭВМ и информационные технологии».

+7-904-959-88-21

sirotkinapyu@student.bmstu.ru

Classification of methods for increasing image resolution over several frames

Annotation

The article is devoted to a review of methods for increasing image resolution (the so-called super-resolution problem) using several frames of the same object. The subject area and the main approaches to solving the problem are briefly reviewed: methods based on interpolation, methods of reconstruction with regularization, methods based on set theory, example-based methods (patch selection), methods for optimizing convolutional neural networks. A critical analysis of methods for increasing image resolution over several frames was carried out, identifying areas of the most effective use based on the following criteria: computational complexity, processing quality, and the need for post- or preprocessing. A conclusion is made about the scope of applicability of the considered methods. To achieve the best processing quality, it is advisable to use methods based on convolutional neural networks.

Keywords

digital image, digital signal processing, super-resolution, convolutional neural networks, interpolation, reconstruction, regularization, set theory, patches