

Реконструкция изображения как задача квадратичного программирования

В.В. Мясников^{1,2}, А.В. Веричев²

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

²Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация. В рамках классического (не нейросетевого) подхода к решению задачи реконструкции (англ.: inpainting) цифрового изображения в работе предлагается метод ее решения с использованием квадратичного программирования. Рассматриваются постановки задачи реконструкции без обучения и с обучением, анализируются результаты решения с использованием различных модификаций предлагаемого метода. Представлены примеры решения на известных данных. Обсуждаются достоинства и недостатки предложенного метода, направления его развития.

1. Введение

Развитие современных методов и средств обработки цифровой информации привели к развитию значительного числа новых направлений. Одним из таких направлений, активно развивающихся в последнее десятилетие, является направление по реконструкции (англ.: inpainting) цифровых изображений [1-3]. Под реконструкцией изображения в данном случае понимается «доопределение» цифрового изображения на части отсчетов изображения (областях, или т.н. *лакунах*), где значения яркости полагаются неизвестными (см. рисунок 1). Особый интерес представляют методы, которые позволяют реконструировать изображение для областей достаточно большого размера (относительно размеров самого изображения или изображенных на нем объектов). Актуальность методов и алгоритмов решения задачи реконструкции цифровых изображений объясняется целым рядом объективных потребностей, наиболее типичными из которых являются [1]:

- необходимость реконструкции старых и поврежденных изображений, в том числе, при их оцифровке;

- необходимость улучшения потребительских свойств цифрового изображения за счет устранения и замены некоторых нежелательных его фрагментов, например, «удаление» со снимка изображения небольшого технического объекта для получения фотографии живой природы;

- при синтезе изображений со сменным ракурсом съемки необходимость реконструкции областей на изображении, «затененных» объектами при исходной позиции камеры.

Хотя сами по себе методы и алгоритмы реконструкции цифровых изображений, как неспециализированные математико-технические решения, не несут какого-либо позитивного или отрицательного контекста, цели их использования могут быть совершенно различными. В

связи с этим, наряду с самими методами реконструкции развиваются и средства выявления «следов» такой реконструкции. И этому вопросу также в последнее десятилетие посвящено значительное число работ [4-6].

Настоящая работа построена следующим образом. Во втором разделе представлен краткий обзор существующих направлений ее решения. В третьем разделе приводится предлагаемая математически корректная постановка задачи реконструкции, указаны все рассматриваемые ограничения и используемые соотношения. В этом же разделе представлены два алгоритма реконструкции цифрового изображения: без обучения и с обучением, то есть в случае наличия прототипов изображений, а также результаты экспериментов по реконструкции цифровых изображений.

2. Краткий обзор существующих направлений и подходов

Формально, постановка задачи реконструкции цифрового изображения в наиболее простой постановке может быть представлена следующим образом. Пусть функция яркости изображения

$$f : \Omega \rightarrow V \quad (1)$$

задана на множестве отсчетов $\Omega \equiv Z_{N_1} \times Z_{N_2}$, где пара N_1, N_2 определяет вертикальный и горизонтальный размеры изображения, $\omega \equiv (n_1, n_2) \in \Omega$, а множество $Z_N \equiv \{0, 1, \dots, N-1\}$ - множество целых чисел до N . Значениями функции яркости выступают величины из множества V , в качестве которого на практике обычно выступают множества $Z_{2^B-1}^K$ или $R_{[0,1]}^K$, где K - число «каналов» изображения (обычное цветное изображение имеет три спектральных канала, соответствующих различным длинам волн), а B - число требуемых для представления яркости отсчета битов. Область задания изображения - множество Ω - в задаче реконструкции изображения (inpainting) предполагается состоящей из двух непересекающихся областей:

$$\Omega = \Omega_0 \cup \Omega_1,$$

где Ω_1 определяет множество отсчетов изображения с известными значениями функции яркости (1), а Ω_0 , напротив, с неизвестными. Таким образом, в простейшем варианте постановки задачи реконструкции без обучения необходимо доопределить значения функции (1) на множестве Ω_0 (f_ω - известные величины функции яркости):

$$\begin{cases} f(\omega) = ?, & \omega \in \Omega_0, \\ f(\omega) = f_\omega, & \omega \in \Omega_1. \end{cases} \quad (2)$$

В случае реконструкции с обучением, наряду в самим реконструируемым изображением в постановке задачи могут фигурировать другие изображения того же типа (например, изображения лиц или изображения зданий), заданные обычно на всем множестве Ω . При этом в наиболее общей постановке для различных изображений множества Ω задания функции яркости могут быть различными.

Совершенно очевидно, что сформулированная задача реконструкции является некорректной по Адамару, поскольку имеет произвольное множество решений. Дополнительной проблемой также является невозможность математически обоснованного введения критерия качества реконструкции [1,12], что также затрудняет использование формальных подходов к построению методов и алгоритмов ее решения.

Как и во многих задачах машинного обучения и зрения, в настоящее время можно выделить две принципиально различные тенденции решения задачи реконструкции изображений. Исторически первой - *классической* - является использование определенных математических моделей для построения алгоритмов реконструкции. Введение моделей позволяет перейти от неформальной и некорректной в математическом плане постановки (2) к корректным математическим постановкам. Обычно [1] выделяют три принципиально различных подхода в рамках этого направления.

Диффузные методы (англ.: diffusion methods) [1,7] используют в качестве ключевого момента понятие гладкости: они распространяют локальную информацию, заложенной в функцию яркости, сохраняя гладкость ее определенных характеристик. Используемый для этого аппарат – дифференциальные уравнения в частных производных, где входные данные (значения функции яркости в отсчетах области Ω_0) выступают в качестве начальных условий. В процессе решения возникающей таким образом задачи Коши локальная «структура» функции яркости «распространяется» от границы области Ω_0 внутрь ее, сохраняя присущие для выбранной дифференциальной модели свойства. Сохранение границ для диффузных методов обычно осуществляется путем дополнительного учета градиента существующей функции яркости. Диффузные метод обычно дают хороший результат при реконструкции на изображений «узких» линейных дефектов или областей достаточно малого размера.

Прецедентные методы (англ.: exemplar-based methods) основаны на простой идеи поиска в качестве прецедентов для заполнения Ω_0 похожих фрагментов на том же самом изображении [1,8-9]. В качестве «шаблона» поиска выступают отсчеты из Ω_1 , находящиеся вблизи границы. Обычно итерационный метод инкрементным образом заполняет все более «сужающуюся» область Ω_0 теми фрагментами из Ω_1 , которые оказались наиболее «близкими» по некоторому критерию к соответствующему «шаблону». Одним из наиболее известных методов этого типа является метод, предложенный в работах [8-9]. На классах изображений, содержащих регулярные в детерминированном или статистическом смысле текстуры (фотографии зданий, водной поверхности, городской инфраструктуры и т.п.) он демонстрирует результаты, близкие к превосходным с точки зрения человеческого восприятия. Одним из основных недостатков прецедентных методов является их высокая вычислительная сложность, связанная с объективной сложностью решения задач поиска «дубликатов» на изображениях [1,8-9]. Для ее снижения разрабатываются специализированные быстрые методы [10].

Разреженные представления (англ.: sparse representation) – еще один классический подход к решению задачи реконструкции цифрового изображения [1,11]. Суть его заключается в представлении изображения и/или его фрагмента с использованием линейной комбинации небольшого числа специализированных базисных функций, определяемых по самому анализируемому изображению. Идея, как очевидно, пересекается с известным методом главных компонент (англ.: PCA – principle component analysis), реализация которого для представления нормализованных по геометрическому положению лиц – eigenfaces – показывает возможность достаточно качественного описания изображения лица с использованием одной-двух сотен базисных функций. Как следствие, также возможно эффективное решение задачи реконструкции подобных изображений. Ключевой проблемой методов этого класса является выбор подходящего базиса, который собственно и определяет потенциально возможный класс изображений. Поиск его осуществляется различными способами, например, путем разделения области определения изображения на подобласти, соответствующие различным объектам и, как следствие, имеющим различное текстурное содержание. На каждой из областей в этом случае производится построение «своего» базиса. Но все же вопрос сложности определения подходящего базиса – одна из основных проблем этого подхода.

В классическом подходе наряду с отдельными методами может использоваться и их комбинация [1].

С определенным временем в задачах компьютерного зрения значительного успеха достигли искусственные нейронные сети (ИНС), в частности, сети свертки и глубокие ИНС. Их использование в задаче реконструкции цифровых изображений не только позволило получить выдающиеся результаты в случае больших областей Ω_0 [2], но также перейти к решению гораздо более сложных задач синтеза целых новых сцен и человеческих образов [13]. Несмотря на это, использование классического подхода остается все еще актуальным, особенно когда вычислительные ресурсы ограничены. Подробный обзор методов на основе ИНС выходит за рамки настоящей работы.

3. Предлагаемый метод реконструкции

Представленное ниже решение приводится для случая однокомпонентного изображения, то есть при $K=1$. Для случая многокомпонентных изображений указанные ниже задачи решаются независимо K раз.

3.1. Формальная постановка задачи реконструкции

Мы предлагаем решение, сочетающее в себе свойства всех трех классических подходов. Как и в случае разреженных представлений, мы ищем решение в виде изображения, представимого в виде разложения по некоторому заранее выбранному базису:

$$f(\omega) = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} \alpha_{m_1, m_2} \varphi_{m_1, m_2}(\omega), \quad \omega \in \Omega \quad (3)$$

В данной работе мы ограничиваемся рассмотрением косинусного базиса типа IV, для которого существуют алгоритмы быстрого дискретного разложения. Для этого базиса вид базисных функций:

$$\varphi_{m_1, m_2}(n_1, n_2) = \varphi_{m_1}(n_1) \varphi_{m_2}(n_2),$$

где

$$\varphi_{m_i}(n_i) = \cos\left((m_i + 0.5)(n_i + 0.5) \frac{\pi}{N_i}\right), \quad n_i, m_i = \overline{0, N_i - 1}$$

Однако, в отличие от подхода на основе разреженных представлений, мы не накладываем ограничений на число базисных функций в (3), то есть в общем случае $M = |\Omega|$. Дополнительно, для изображения целиком и/или его фрагмента мы отталкиваемся от ограничений, накладываемых условиями-равенствами (2). Поскольку указанных ограничений-равенств, очевидно, недостаточно для решения задачи представления изображения в виде (3), мы добавляем к ним дополнительные условия для гарантии единственности решения:

- условия на значения функции яркости и ее производных как ограничения-неравенства,
- условия на значения коэффициентов в виде выпуклого функционала. Вид функционала определяется типом задачи (с обучением или без), а также конфигурацией и размером Ω_0 и Ω_1

В результате неформально сформулированная задача (2) представляется в виде задачи квадратичного программирования [14]:

$$\begin{cases} \mathbf{a}^T P \mathbf{a} + q^T \mathbf{a} \rightarrow \min_{\mathbf{a}} \\ G \mathbf{a} \leq h, \\ A \mathbf{a} = b. \end{cases} \quad (4)$$

Здесь вектор неизвестных \mathbf{a} состоит из линейных коэффициентов разложения (3):

$$\mathbf{a} = (\alpha_{0,0}, \alpha_{0,1}, \dots, \alpha_{0, N_2-1}, \alpha_{1,0}, \alpha_{1,1}, \dots, \alpha_{1, N_2-1}, \dots, \alpha_{N_1-1,0}, \alpha_{N_1-1,1}, \dots, \alpha_{N_1-1, N_2-1})^T$$

Матрица A и вектор b соответствуют условиям-равенствам в (2). Матрица A размера $|\Theta_1| \times N_1 N_2$ состоит из $|\Theta_1|$ строк вида:

$$[\varphi_0(n_1)\varphi_0(n_2) \quad \dots \quad \varphi_0(n_1)\varphi_{N_2-1}(n_2) \quad \varphi_1(n_1)\varphi_0(n_2) \quad \dots \quad \varphi_{N_1-1}(n_1)\varphi_{N_2-1}(n_2)], \quad (n_1, n_2) \in \Theta_1, \quad (5)$$

а вектор b состоит из $|\Theta_1|$ значений функции яркости, соответствующих строкам матрицы A с совпадающими (n_1, n_2) :

$$f_{(n_1, n_2)}, \quad (n_1, n_2) \in \Theta_1 \quad (6)$$

Матрицы P , G и вектора q , h определяются в зависимости от типа задачи по разному.

3.2. Задача реконструкции без обучения

В случае задачи без обучения матрицы P , G и вектора q , h могут быть определены исключительно по данным самого реконструируемого изображения.

Определим по анализируемому изображению максимальное и минимальное значение функции яркости:

$$f^{\max} = \max_{(n_1, n_2) \in \Theta_1} f_{(n_1, n_2)}, \quad f^{\min} = \min_{(n_1, n_2) \in \Theta_1} f_{(n_1, n_2)}, \quad (7)$$

А также аналогичные величины для производных функции яркости:

$$f_{r,k}^{\max} = \max_{(n_1, n_2) \in \Theta_1^k} \frac{\partial f_{(n_1, n_2)}^k}{\partial n_r^k}, \quad f_{r,k}^{\min} = \min_{(n_1, n_2) \in \Theta_1^k} \frac{\partial f_{(n_1, n_2)}^k}{\partial n_r^k}, \quad r=1,2; \quad k=1,2,\dots \quad (8)$$

В представленных выражениях множество Θ_1^k определяет внутренность множества Θ_1^{k-1} , где $\Theta_1^0 \equiv \Theta_1$ и введено с учетом с учетом невозможности рассмотрения значений численных производных на границах множества Θ_1 . Аналогичные обозначения считаем справедливыми и для множеств Θ_0^k, Θ_0 . В качестве производных используются численные значения производных ($f_{(n_1, n_2)}^0 \equiv f_{(n_1, n_2)}$):

$$\frac{\partial f_{(n_1, n_2)}^k}{\partial n_1^k} = f_{(n_1, n_2)}^{k-1} - f_{(n_1-1, n_2)}^{k-1}, \quad \frac{\partial f_{(n_1, n_2)}^k}{\partial n_2^k} = f_{(n_1, n_2)}^{k-1} - f_{(n_1, n_2-1)}^{k-1} \quad r=1,2; \quad k=1,2,\dots \quad (9)$$

С учетом введенных обозначений задача (4) в части ограничений-неравенств доопределяется следующим образом:

- строки матрицы G и соответствующие значения вектора h , определяющие ограничения по значению функции яркости для каждого отсчета в области Θ_0 :

$$\begin{aligned} & \left[\varphi_0(n_1)\varphi_0(n_2) \quad \dots \quad \varphi_0(n_1)\varphi_{N_2}(n_2) \quad \varphi_1(n_1)\varphi_0(n_2) \quad \dots \quad \varphi_{N_1-1}(n_1)\varphi_{N_2-1}(n_2) \right] \square f^{\max}; \\ & \left[-\varphi_0(n_1)\varphi_0(n_2) \quad \dots \quad -\varphi_0(n_1)\varphi_{N_2}(n_2) \quad -\varphi_1(n_1)\varphi_0(n_2) \quad \dots \quad -\varphi_{N_1-1}(n_1)\varphi_{N_2-1}(n_2) \right] \square -f^{\min}; \\ & (n_1, n_2) \in \Theta_0; \end{aligned} \quad (10)$$

- строки матрицы G и соответствующие значения вектора h , определяющие ограничения по значению k -ой производной функции яркости для каждого отсчета в области Θ_0 . Число контролируемых производных по каждому из направлений является параметром для предлагаемого метода ($1 \leq k \leq k^{\max}$, $r=1,2$):

$$\begin{aligned} & \left[n_r^k \varphi_0(n_1)\varphi_0(n_2) \quad \dots \quad n_r^k \varphi_0(n_1)\varphi_{N_2}(n_2) \quad n_r^k \varphi_1(n_1)\varphi_0(n_2) \quad \dots \quad n_r^k \varphi_{N_1-1}(n_1)\varphi_{N_2-1}(n_2) \right] \square f_{r,k}^{\max}; \\ & \left[-n_r^k \varphi_0(n_1)\varphi_0(n_2) \quad \dots \quad -n_r^k \varphi_0(n_1)\varphi_{N_2}(n_2) \quad -n_r^k \varphi_1(n_1)\varphi_0(n_2) \quad \dots \quad -n_r^k \varphi_{N_1-1}(n_1)\varphi_{N_2-1}(n_2) \right] \square -f_{r,k}^{\min}; \\ & (n_1, n_2) \in \Theta_0. \end{aligned} \quad (11)$$

Указанные ограничения вытекают из известных соотношений, связывающих значения производных функции, представленной рядом Фурье и ее коэффициентов спектрального разложения. А именно (для упрощения изложения соотношения представлены в одномерном непрерывном случае):

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega \Rightarrow \frac{\partial f(x)}{\partial x} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} i\omega F(\omega) e^{i\omega x} d\omega$$

Наконец, вектор q полагается нулевым, а матрица P подбираются таким образом, чтобы текстурные свойства результирующего изображения на множествах Θ_0 и Θ_1 были максимально близки. Причем матрица P , характеризующая отношения между парами отсчетов $(n_1, n_2), (m_1, m_2)$ оказывается зависимой от двух параметров: бинарного типа t и

вещественного порядка s . В общем виде элементы матрицы (формально записываемой как $P_{t,s}$) имеют следующий вид:

$$p_{t,s}((n_1, n_2), (m_1, m_2)) = \begin{cases} 0, & (n_1, n_2) \neq (m_1, m_2); \\ \text{else} & \begin{cases} (N_1 - n_1 - 1)^{-s} + (N_2 - n_2 - 1)^{-s}, & s < 0, t = 1; \\ (n_1)^s + (n_2)^s, & s > 0, t = 1; \\ (N_1 - n_1 - 1 + N_2 - n_2 - 1)^{-s}, & s < 0, t = 0; \\ (n_1 + n_2)^{-s}, & s < 0, t = 0. \end{cases} \end{cases} \quad (12)$$

Под текстурными свойствами изображения на множествах Θ_r ($r \in \{0,1\}$) в данной работе понимаются первые T отсчетов эмпирической корреляционной функции (по каждому из направлений): ρ_{t_1, t_2}^r ($0 \leq t_1, t_2 \leq T$, $r \in \{0,1\}$). В результате окончательное решение задачи реконструкции цифрового изображения получается в результате подбора таких параметров t и s , которые обеспечивали бы минимальное рассогласование наборов ρ_{t_1, t_2}^0 и ρ_{t_1, t_2}^1 по некоторой наперед заданной норме: $\|(\rho_{t_1, t_2}^0) - (\rho_{t_1, t_2}^1)\|$. Учитывая, что аналитическое решение такой задачи не представляется возможным, для получения окончательного результата используется перебор по потенциально возможным прецедентам t и s , что может быть представлено в виде алгоритма-псевдокода:

Алгоритм (параметры k_r^{\max} , λ)

For t in $\{0,1\}$

For s in $\{-6, -5.5, -5, \dots, 5, 5.5, 6\}$

// инициализация

матрицу A и вектор b задать по формулам (5) и (6)

матрицу G и вектор h задать по формулам (10) и (11) с учетом параметра k_r^{\max} по каждой из координат $r=1,2$;

Сформировать $P_{t,s}$ по (12)

$q = 0$

// решение

Решение задачи (4), получение набора $\alpha^{t,s}$;

Оценка матриц статистических свойств для Θ_r ($r \in \{0,1\}$): $(\rho_{t_1, t_2}^0) (\rho_{t_1, t_2}^1)$;

Оценка рассогласования статистических свойств $\varepsilon_{t,s} = \|(\rho_{t_1, t_2}^0) - (\rho_{t_1, t_2}^1)\|$

$t^*, s^* = \arg \min_{t,s} \varepsilon_{t,s}$

Результат: изображение по формуле (3) с коэффициентами α^{t^*, s^*} .

Примеры реконструированных с использованием указанного алгоритма изображений известного оперного театра г. Сиднея (Австралия) и фасада дома приведены на рисунках 1-2.

3.3. Задач реконструкции с обучением

Принципиальным отличием задачи реконструкции с обучением является наличие дополнительной обучающей информации, которая может быть использована для (до)определения матриц P , G и векторов q , h несколько иным от стандартного способа методом. При этом следует различать варианты постановок с геометрически согласованными изображениями одного типа (например, для случая анализа изображений лиц, где их расположение нормализуется по положению зрачков глаз) и просто изображения одного типа (например, изображения домов, природных ландшафтов, моря и т.п.). Хотя нахождение матрицы P и вектора q не зависит от факта геометрической согласованности, выбор матрицы G

и вектора h производится по разному. Далее в подразделе представлены 3 секции, первая из которых определяет способ задания матрицы P и вектора q , а две последующие - способы определения матрицы G и вектора h для различных постановок.

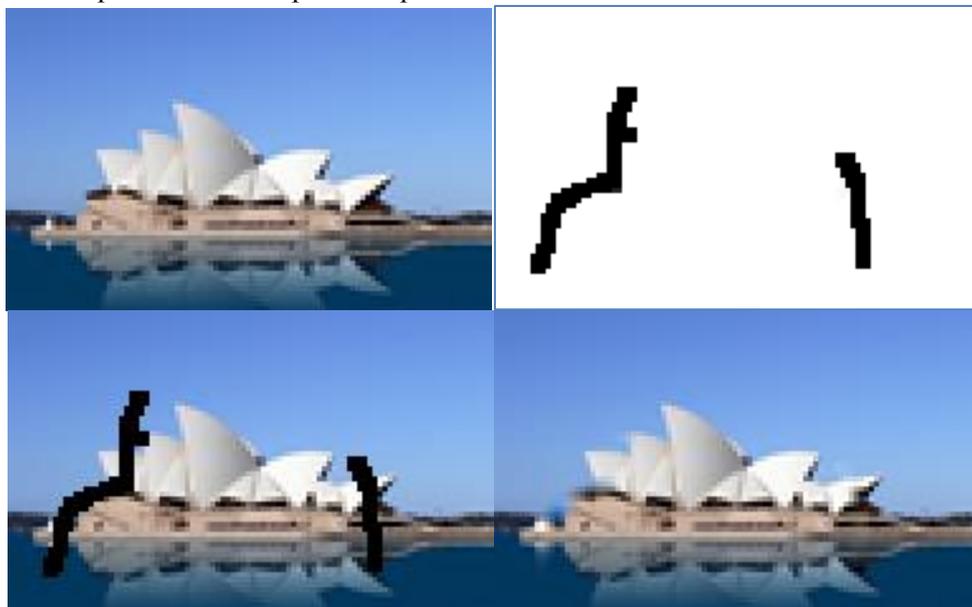


Рисунок 1. Иллюстрация к предложенному методу (сверху-вниз) для изображения оперного театра г. Сиднея (Австралия): - исходное изображение 96x60; - области Θ_0, Θ_1 ; - области Θ_0, Θ_1 на исходном изображении; - реконструированное изображение.

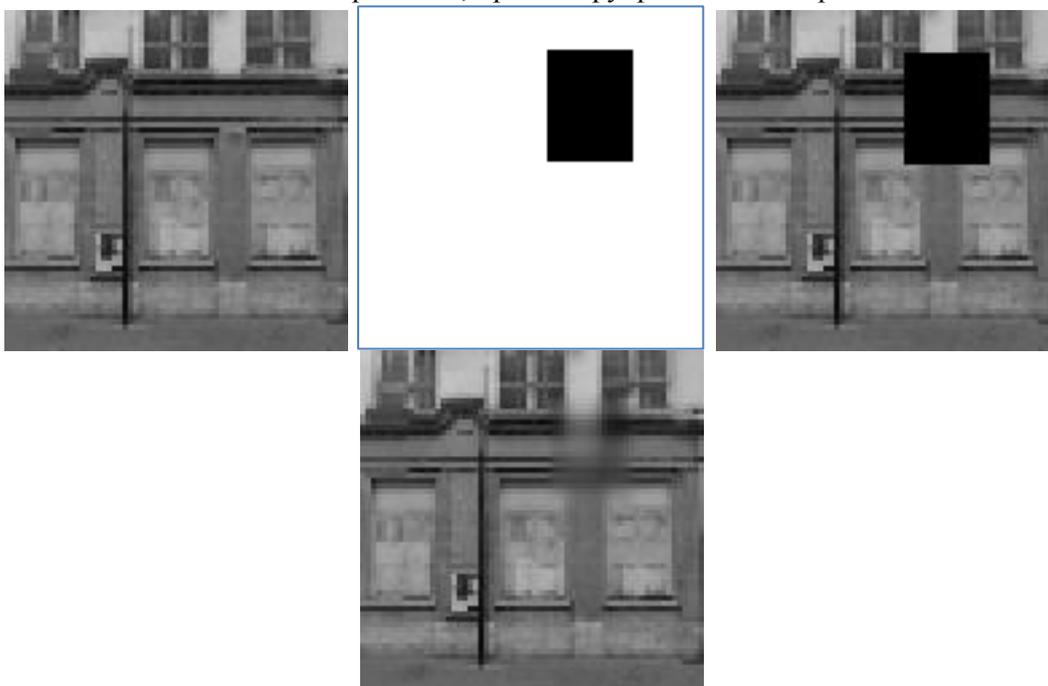


Рисунок 2. Иллюстрация к предложенному методу (сверху-вниз) для изображения здания: - исходное изображение 60x64; - области Θ_0, Θ_1 ; - реконструированное изображение.

Для определенности ниже считаем, что в качестве обучающего множества задано множество полностью известных изображений $\{f_{(n_1, n_2)}^j, (n_1, n_2) \in \Theta^j\}_{j=0}^{J-1}$ (изображения считаем нормированными по среднему и дисперсии).

3.3.1. Задание матрицы P и вектора q в задаче реконструкции с обучением

Пусть для всех изображений из обучающей выборке известно их представление в виде разложения (3) по косинусному базису, что вычислительно эффективно может быть сделано с использованием быстрого алгоритма [Блейхут]: $\{F_{(m_1, m_2)}^j, (m_1, m_2) \in \Theta^j\}_{j=0}^{J-1}$. Величины $F_{(n_1, n_2)}^j$ играют ту же роль, что коэффициенты α_{m_1, m_2} . Учитывая, что коэффициенты разложения - суть сумма большого числа случайных величин, их закон распределения может приближенно считаться нормальным. В этом случае параметры такого нормального закона распределения имеют вид:

$$\text{cov}\{F_{(m_1, m_2)}, F_{(n_1, n_2)}\}, E\{F_{(m_1, m_2)}\}, (m_1, m_2) \in \bigcup_{j=0}^{J-2} \Theta^j$$

где оценки ковариационной матрицы и среднего по выборке имеют вид:

$$E\{F_{(m_1, m_2)}\} \square \frac{\sum_{j=0}^{J-1} F_{(m_1, m_2)}^j \cdot I((m_1, m_2) \in \Theta^j)}{\sum_{j=0}^{J-1} I((m_1, m_2) \in \Theta^j)}, (m_1, m_2) \in \bigcup_{j=0}^{J-2} \Theta^j;$$

$$\text{cov}\{F_{(m_1, m_2)}, F_{(n_1, n_2)}\} \square \frac{\sum_{j=0}^{J-1} (F_{(m_1, m_2)}^j - E\{F_{(m_1, m_2)}\}) (F_{(n_1, n_2)}^j - E\{F_{(n_1, n_2)}\}) \cdot I((n_1, n_2) \in \Theta^j \wedge (m_1, m_2) \in \Theta^j)}{\sum_{j=0}^{J-1} I((n_1, n_2) \in \Theta^j \wedge (m_1, m_2) \in \Theta^j)},$$

$$(m_1, m_2) \in \bigcup_{j=0}^{J-2} \Theta^j, (n_1, n_2) \in \bigcup_{j=0}^{J-2} \Theta^j.$$

Векторно-матричные аналоги приведенных скалярных величин обозначим: $E\{\mathbf{F}\}, \text{cov}\{\mathbf{F}, \mathbf{F}\}$. Максимизация вероятности принадлежности конкретного спектра $F_{(n_1, n_2)}$ к нормальному закону с указанными параметрами, как известно, эквивалентно минимизации расстояния Махаланобиса до среднего $E\{F_{(m_1, m_2)}\}$, рассчитываемого с учетом приведенной ковариационной матрицы. Тогда, то есть в качестве оптимизируемой функции критерия естественно взять следующую:

$$(\mathbf{a} - E\{\mathbf{F}\})^T (\text{cov}\{\mathbf{F}, \mathbf{F}\})^{-1} (\mathbf{a} - E\{\mathbf{F}\})$$

или

$$\mathbf{a}^T (\text{cov}\{\mathbf{F}, \mathbf{F}\})^{-1} \mathbf{a} - 2E\{\mathbf{F}\}^T (\text{cov}\{\mathbf{F}, \mathbf{F}\})^{-1} \mathbf{a} + E\{\mathbf{F}\}^T E\{\mathbf{F}\} \rightarrow \min_{\mathbf{a}}$$

Сопоставляя полученное выражение с критерием задачи (4) и игнорируя константное последнее слагаемое, имеем:

$$P = (\text{cov}\{\mathbf{F}, \mathbf{F}\})^{-1}, \quad q = 2E\{\mathbf{F}\}^T (\text{cov}\{\mathbf{F}, \mathbf{F}\})^{-1}$$

3.3.2. Задание матрицы G и вектора h в задаче реконструкции с обучением: случай геометрически несогласованных изображений одного типа

Задание производится точно тем же методом, что и в случае задачи реконструкции без обучения, то есть с использованием выражений (7)-(11).

3.3.3. Задание матрицы G и вектора h в задаче реконструкции с обучением: случай геометрически согласованных изображений одного типа

В отличие от остальных случаев диапазон значений функции яркости или ее производной определяется для каждой позиции отдельно. А именно:

$$f_{(n_1, n_2)}^{\max} = \max_{j=0, J-1} f_{(n_1, n_2)}^j, \quad f_{(n_1, n_2)}^{\min} = \min_{j=0, J-1} f_{(n_1, n_2)}^j, \\ f_{r, k}^{\max} = \max_{j=0, J-1} \frac{\partial (f_{(n_1, n_2)}^j)^k}{\partial n_r^k}, \quad f_{r, k}^{\min} = \min_{j=0, J-1} \frac{\partial (f_{(n_1, n_2)}^j)^k}{\partial n_r^k}, \quad r=1, 2; \quad k=1, 2, \dots \quad (8)$$

В дальнейшем именно эти величины выступают в качестве компонент вектора h . При этом составляющие матрицы G остаются теми же, что и в выражениях (10)-(11).

4. Выводы и результаты

В работе предложен метод решения задачи реконструкции цифрового изображения с использованием квадратичного программирования. Для двух постановок задачи - без обучения и с обучением - представлены модификации предлагаемого метода. Показана их работоспособность на примерах. Наряду с очевидным достоинством метода (простота постановки задачи, произвольная форма реконструируемых областей) метод обладает и недостатками:

- высокой вычислительной сложностью, следующей из необходимости решения задачи квадратичного программирования;
- проблемами при его использовании на больших изображениях, также связанной с необходимостью решения задачи квадратичного программирования;
- чрезмерная (по сравнению с лучшими прецедентными методами) склонность к сглаживанию, свойственная также диффузным методам и методам на основе разреженных представлений.

В расширенной версии работы и последующих публикациях предлагается преодолеть некоторые из указанных недостатков.

5. Благодарности

Работа выполнена при поддержке Министерства науки и высшего образования в рамках выполнения работ по Государственному заданию ФНИЦ «Кристаллография и фотоника» РАН в части разделов (1) «Введение» и (2) «Краткий обзор существующих направлений и подходов», и грантов РФФИ (№ 18-01-00748, 17-29-03190-офи, 20-37-70053-Стабильность) в части раздела (3) «Предлагаемый метод реконструкции».

6. Литература

- [1] Guillemot, C. Overview and Recent Advances / C. Guillemot, O. Meur // IEEE Signal Processing Magazine. – 2014. – Vol. 31(1). – P. 127-144.
- [2] Liu, G. Image inpainting for irregular holes using partial convolutions / G. Liu, F.A. Reda, K.J. Shih, T.C. Wang, A. Tao, B. Catanzaro // Proceedings of the European Conference on Computer Vision (ECCV). – 2018. – P. 85-100.
- [3] Кокошкин, А.В. Ретуширование и восстановление отсутствующих частей изображений с помощью итерационного вычисления их спектров / А.В. Кокошкин, В.А. Коротков, К.В. Коротков, Е.П. Новичихин // Компьютерная оптика. – 2019. – Т. 43, № 6. – С. 1030-1040. DOI: 10.18287/2412-6179-2019-43-6-1030-1040.
- [4] S. Teerakanok, T. Uehara. Copy-Move Forgery Detection: A State-of-the-Art Technical Review and Analysis // IEEE ACCESS. – 2019. – Vol. 7. – P. 40550-40568.
- [5] Кузнецов, А.В. Алгоритм обнаружения искажённых дубликатов на цифровых изображениях с использованием бинарных градиентных контуров / А.В. Кузнецов, В.В. Мясников // Компьютерная оптика. – 2016. – Т. 40, № 2. – С. 284-293. DOI: 10.18287/2412-6179-2016-40-2-284-293.
- [6] Мясников, В.В. Локальное порядковое преобразование цифровых изображений / В.В. Мясников // Компьютерная оптика. – 2015. – Т. 39, № 3. – С. 397-405. DOI: 10.18287/0134-2452-2015-39-3-397-405.

- [7] Li, H.W. Localization of Diffusion-Based Inpainting in Digital Images / H. Li, W. Luo, J. Huang // IEEE Transactions on Information Forensics and Security. – 2017. – Vol. 12(12). – P. 3050-3064.
- [8] Barnes, C. PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing / C. Barnes, E. Shechtman, A. Finkelstein, D.B. Goldman // ACM Transactions on Graphics (Proc. SIGGRAPH). – 2009. – Vol. 28(3).
- [9] Barnes, C. The Generalized PatchMatch Correspondence Algorithm / C. Barnes, E. Shechtman, D.B. Goldman, A. Finkelstein // European Conference on Computer Vision, 2010.
- [10] Kumar, L.N. Nayar What is a good nearest neighbors algorithm for finding similar patches in images / L.N. Kumar, S.K. Zhang // Proc. European Conf. Computer Vision (ECCV). – 2008. – Vol. II. – P. 364-378.
- [11] Elad, M. Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing – Springer Science+Business Media, LLC 2010. – 376 p.
- [12] Qureshi, M.A. A critical survey of state-of-the-art image inpainting quality assessment metrics / M.A. Qureshi, M.A. Deriche, A. Amin // Journal of Visual Communication and Image Representation. – 2017. – Vol. 49. – P. 177-191.
- [13] Soiffer, R. Generating Images with Neural Networks / R. Soiffer, J. Zhang // Neural networks Image processing. – 2018. – Vol. 1(1). – P. 1-5.
- [14] Кюнци, П. Нелинейное программирование / П. Кюнци, В. Крелле – Москва: «Советское радио», 1965.

Image Inpainting as a Quadratic Programming Task

V.V. Myasnikov^{1,2}, A.V. Verichev²

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

Abstract. In the framework of the classical (non-neural network) approach to solving the digital image inpainting problem, a method for solving this problem using quadratic programming is proposed. The statements of the reconstruction problem without training and with training are considered, the results of the solution are analyzed using various modifications of the proposed method. Examples of solutions based on known data are presented. The advantages and disadvantages of the proposed method, the directions of its development are discussed.