МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ
НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
Механико-математический факультет

Всероссийская молодежная научная конференция студентов, аспирантов и молодых ученых «Все грани математики и механики»

(23–27 апреля 2019 г.)

Сборник статей

Под редакцией
д-ра физ.-мат. наук, профессора А.В. Старченко

Томск
Издательский Дом Томского государственного университета
2019

Совмещение изображений на основании Shearlet-декомпозиции

Потоцкая А.А., Старченко А.В., Захарова А.А.

Томский государственный университет, Томск

Институт прикладных наук Руана, Руан

e-mail: bubuzyonok@yandex.ru

Аннотация. В работе рассматривается задача совмещения двух изображений. Целью совмещения изображений является нахождение преобразования между двумя изображениями одной сцены или одного объекта, снятых в разное время, с разных точек зрения или с разных датчиков. Данная задача была исследована с помощью различных методов, которые часто зависят от конкретного применения. В данной работе рассмотрен алгоритм совмещения изображений на основании Shearlet-декомпозиции. Отображение между изображениями ищется на основании отображения между ключевыми признаками изображений, которые выделяются с помощью Shearlet-декомпозиции.

Ключевые слова: Совмещение изображений, теория фреймов, Shearlet-декомпозиция.

Совмещение изображений - это процесс преобразования различных наборов данных в одну систему координат. Данные могут представлять собой несколько фотографий, данные с разных датчиков, глубин, точек обзора или полученные в разные моменты времени, Совмещение необходимо для того, чтобы иметь возможность сравнивать или интегрировать данные, полученные в процессе разных измерений. Алгоритмы совмещения изображений обычно делят на прямые (на основании пикселей) и на алгоритмы, основанные на деталях изображения. [1]

Сформулируем математическую постановку задачи совмещения изображений . Пусть I_1и I_2 – изображения в пространстве $L^2 = \{f: \mathbb{R}^2 \to \mathbb{R} : \int_{-\infty}^{\infty} |f(x)|^2 dx < \infty\}$ со скалярным произведением $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)dx$ и нормой $\|f\|_2 = \sqrt{\int_{-\infty}^{\infty} |f(x)|^2 dx}.$ Необходимо найти оптимальное преобразование между изображениями I_1и I_2. Таким образом, сформулируем задачу совмещения изображений [1]:

$$(P'): \text{Найти } \eta_0' = \mathrm{argmin}_{\eta \in \mathrm{T}} \|U(\eta)I_1 - I_2\|_2,$$

где Т – группа преобразований между двумя изображениями, $\eta \in \mathrm{T}$ - преобразование, представленное в виде вектора из пространства $\mathbb{R}^P$, содержащего параметры преобразования, $U(\eta) \in L^2$ – альтернативное унитарное представление преобразования. То есть, для любого $\eta \in \mathrm{T}$, $U(\eta)$ - это функция, которая отображает изображение f в его преобразованное изображение $U(\eta)f \in L^2$. Более того, так как $U(\eta)$ – унитарный оператор, то $\|U(\eta)f\|_2 = \|f\|_2$. Обозначим через $d(I_1, I_2) = \|U(\eta_0')I_1 - I_2\|_2$ инвариантное расстояние относительно преобразования η_0' между изображениями I_1и I_2.К сожалению, поиск преобразования η_0' и соответствующего инвариантного расстояния является сложной задачей, так как целевая функция обычно не выпукла и имеет множество локальных минимумов.

Чтобы обойти эту проблему, предположим, что изображения хорошо аппроксимированы их разреженным разложением в ряд геометрических функций. Пусть D – параметрический словарь геометрических особенностей, построенный преобразованием порождающей функции $\varphi \in L^2$ следующим образом:

$$D = \{\varphi_\gamma : \gamma \in T_d\} \subset L^2 \ , \tag{1}$$

где $T_d \subset T$ – конечная дискретизация группы преобразований T и $\varphi_\gamma = U(\gamma)\varphi$ обозначает преобразование порождающей функции φ с помощью γ. Обозначим через p и q разреженные аппроксимации функций (изображений) I_1и I_2 соответственно в словаре D:

$$\begin{aligned} p&=\sum_{i=1}^{K} c_i \varphi_{\gamma_i}, \\ q&=\sum_{i=1}^{K} d_i \varphi_{\delta_i}. \end{aligned} \tag{2}$$

Так как словарь D содержит особенности, которые показывают потенциальные части изображения, то предположим, что коэффициенты c_i и d_i – все неотрицательные, то есть различные особенности не погашают друг друга.

Считаем, что любой элемент φ_γ из D (1) является геометрической особенностью изображения. Предположим сначала, что функция φ неотрицательна. В дополнение, предположим для простоты, что $\gamma \mapsto \varphi_\gamma$ представляет собой взаимно-однозначное соответствие. Наконец, предположим, не теряя общности, что порождающая функция φ нормирована.

Теперь мы можем определить задачу совмещения, как задачу нахождения оптимального отображения между разреженными образцами. Строгая постановка [1]:

$$(P)\text{: Найти } \eta_0 = argmin_{\eta \in T} \|U(\eta)p - q\|_2.$$

Наименьшее расстояние $d(p,q) = \|U(\eta_0)p - q\|_2$ является преобразованным инвариантным расстоянием между разреженными аппроксимациями изображений p и q. В сравнении с основной постановкой, изображения I_1 и I_2 заменены их соответствующими разреженными представлениями p и q. Это дает некоторые потенциальные преимущества в приложениях. Необходимо ответить, что если исходные изображения не удачно аппроксимированы соответствующими разреженными представлениями, то решение задачи (P) может значительно отличаться от решения основной задачи (P').

Рассмотрим далее алгоритм решения задачи совмещения, предложенный в [1] для изображений, представленных их разреженными представлениями. Ключевая идея данного алгоритма основывается на свойстве ковариации словаря D (1): глобальное преобразование, примененное к изображению, вызывает эквивалентное преобразование соответствующих признаков. Благодаря этому свойству ковариации можно вывести глобальное преобразование между изображениями простым вычислением относительных преобразований между особенностями обоих изображений.

В частности, пусть $T_a^{p,q}$ - множество относительных преобразований между парами признаков, взятых соответственно в p и q: $T_a^{p,q} = \{\delta_i \circ \gamma_j^{-1}: 1 \leq i, j \leq K\}$. Таким образом, можно оценить преобразование между изображениями, решив следующую более слабую задачу:

$$(\hat{P})\text{: Найти } \hat{\eta} = \operatorname{argmin}_{\eta \in T_a^{p,q}} \|U(\eta)p - q\|_2.$$

Минимум целевой функции $d_a(p,q) = \|U(\hat{\eta})p - q\|_2$ определяется как приближенное расстояние, инвариантное к преобразованию между I_1 и I_2.

Хотя задачи (P) и $(\hat{P})$ имеют некоторые общие черты, они различаются в важном аспекте, а именно в пространстве поиска. Оно сокращается от T до конечного множества $T_a^{p,q}$. Это заставляет оцениваемое преобразование быть равным преобразованию, которое точно отображает две особенности, взятые соответственно из p и q. Предположение, что T можно заменить на $T_a^{p,q}$, вытекает из наблюдения, что особенности ковариантны глобальному преобразованию, примененному к исходному изображению. Даже если это предположение не обязательно верно для всех признаков, когда между изображениями существует отображение (отличное от искомого глобального преобразования), ожидается, что существует, по крайней мере, один признак, преобразование которого согласуется с оптимальным преобразованием η_0. Преимущество замены T на $T_a^{p,q}$, тем не менее, является непосредственным: неразрешимая проблема сводится к проблеме, пространство поиска которой имеет мощность не более K^2. Поскольку K обычно выбирается достаточно малым, задача $(\hat{P})$ может быть эффективно решена путем полного поиска по всем элементам $T_a^{p,q}$.

Вычислительная сложность рассмотренного алгоритма напрямую зависит от выбора константы K: большое значение K приводит к большой мощности пространства поиска $T_a^{p,q}$. Кроме того, от значения K также обычно зависит погрешность аппроксимации исходных изображений их разреженными разложениями p и q. Наконец, нужно отметить, что для простоты необходимо предположить, что оба изображения I_1и I_2 аппроксимируются одинаковым количеством признаков. Тем не менее, легко видеть, что можно обобщить это на случай, когда количество признаков различно. В этом случае имеем $\left|T_a^{p,q}\right| = K_1 K_2$ вместо K^2, где K_1 и K_2 - количество признаков в I_1и I_2 соответственно.

Рассмотрим подробнее переход от исходной задачи в $(\hat{P})$ с точки зрения эффективности совмещения. Для начала определим метрику для измерения точности совмещения изображений. Так как мы хотим получить представление о производительности описанного алгоритма в отношении оптимального совмещения изображений, полученного путем решения (P'), очевидно определить метрику через разницу между расстоянием, инвариантным для преобразования, и его приближенной версией, то есть $E'(p\,, q,\ I_1, I_2) = |\, d_a(p, q) - d\,(I_1, I_2)\,|$. Однако, как говорилось ранее, предполагается, что изображения задаются их разреженными разложениями.

Поэтому используется альтернативная метрика, заданная $E\ (p\ ,q) = d_a(p,q) - d\ (p,q)$, где используются приближения разреженного изображения p и q вместо оригинальных изображений. Отметим, что $E\ (p\ ,q) \geq 0$, так как $T_a^{p,q} \subset T$.

Следующая теорема связывает две метрики $E\ (p\ ,q)$ и $E'(p\ ,q,\ I_1,I_2)$ с ошибками разреженной аппроксимации $\|I_1 - p\|_2$ и $\|I_2 - q\|_2$.

Теорема 1. $E'(p\ ,q,\ I_1,I_2) \leq E\ (p\ ,q) + \|I_1 - p\|_2 + \|I_2 - q\|_2$.

Доказательство. Доказательство данного утверждения можно найтив [1].

Рассмотрим частный случай, когда $d(p,q) = 0$. Это означает, что существует преобразование $\eta_0 \in T$, для которого $q = U(\eta_0)p$, т. е. отображение между разреженными аппроксимациями изображений может быть найдено точно. В данном случае описанный алгоритм совмещения способен восстановить точное глобальное преобразование между p и q, если любое подмножество размера $2K$ в D линейно независимо. [1] Линейная независимость любого подмножества размера $2K$ в словаре D фактически гарантирует более сильный результат: это гарантирует, что у любого разреженного сигнала есть уникальное разложение в D [2]. Другими словами, это гарантирует, что когда два разреженных сигнала равны, все их признаки равны.

Предыдущий результат применим только к идеальному сценарию, поскольку на практике условие $d(p, q) = 0$ редко выполнимо. В общем же случае, между изображениями обычно присутствуют небольшие отличия (помимо преобразования из T), которые приводят к ненулевому расстоянию $d(p, q)$. Кроме того, даже когда исходные изображения точно связаны глобальным преобразованием (т.е. $d(I_1, I_2) = 0$), нет гарантии, что разреженные приближения могут быть идеально выровнены (т.е. $d(p, q) = 0$) из-за дискретизации словаря.

В общем случае, когда разреженные приближения p и q имеют различия, которые не могут учитываться в глобальном геометрическом преобразовании из T. То есть, предположим что существует число $\varepsilon > 0$ такое, что $d(p, q) < \varepsilon\sqrt{\|c\|_2^2 + \|d\|_2^2}$, где c и d обозначают соответственные векторы коэффициентов для приближений p и q из (2). Следовательно, величина ε измеряет нормализованную разницу между p и q.

В работе рассматривается задача совмещения изображений путем поиска относительных преобразований между отличительными особенностями изображений, представленных в виде их разреженных аппроксимаций (2).

Методы, основанные на вейвлетах, были эффективно применены к проблеме совмещения многовременных изображений. Разлагая изображение в соответствии с дискретным вейвлет-алгоритмом, изображение уменьшается до наболее важных характеристик, которые легче сопоставить. Хорошо известно, что вейвлеты достаточно хорошо фиксируют некоторые особенности изображения, такие как текстуры. Однако известно, что вейвлеты являются изотропными, то есть они не эффективно представляют изображения, имеющие характерные особенности. В частности, изображения с четкими краевидными элементами, включая реки, дороги и горы, неоптимально представлены с помощью вейвлет-подобных алгоритмов.

Существует много анизотропных обобщений вейвлетов. В частности, shearlets [3], дают быстрое, оптимизированное и зависящее от направления разложение изображений, которое дает надежный алгоритм регистрации изображений.

Рассмотрим полученные результаты. В качестве тестового изображения рассмотрим стандартное изображение Barbara, представленное на рисунке 1. В качестве второго изображения для совмещения возьмем то же изображение, масштабированное с параметром scale и повернутое на угол alfa (Рис.2).

Рис. 1. Тестовое изображение Barbara

Рис. 2. Масштабированное и повернутое изображение (scale=0.7, alfa=30)

На рисунке 3 представлены результаты поиска особенностей изображений и соотношения между ними. Кружками обозначены особенности на оригинальном изображении, а крестиками – на измененном, линии указывают соответствие между особенностями.

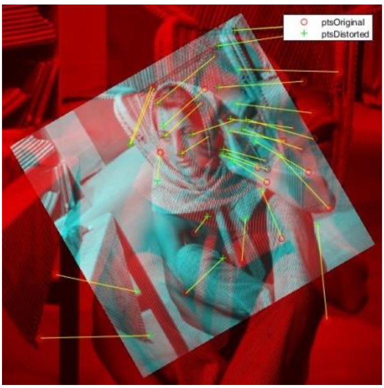

Рис. 3. Соответствие обнаруженных особенностей

Для проверки эффективности использованного алгоритма, проведем обратное преобразование (восстановление) измененного изображения. Для критерия сравнения оригинального и восстановленного изображений будем использовать отношение пикового сигнала к шуму (PSNR). На рисунке 4a приведено оригинальное изображение, а на рисунке 4b – восстановленное. Значение PSNR, в данном случае, равно 26,5.

Рис. 4. Исходное (a) и восстановленное (b) изображения

Рассмотрим влияние масштабирующего параметра и угла поворота на эффективность метода. Для этого рассмотрим результаты численных экспериментов, приведенных в таблице 1. Здесь представлена зависимость «отличия» исходного и восстановленного изображений от масштабирующего параметра и угла поворота.

Таблица 1. Зависимость PSNR от масштабирующего параметра и угла поворота

Масштаб	Поворот	PSNR
scale = 1	alfa = 30	28.5201
scale = 1	alfa = 45	29.4196
scale = 1	alfa = 60	29.8522
scale = 1	alfa = 90	Inf
scale = 0.9	alfa = 45	25.9420
scale = 0.7	alfa = 45	24.0888
scale = 0.5	alfa = 45	24.3584
scale = 0.3	alfa = 45	17.9779

Исходя из представленных результатов, можно сделать вывод, что данный метод эффективно совмещает повернутые изображения. Однако даже небольшое масштабирование значительно ухудшает полученные результаты.

Список литературы

[1] Fawzi A., Frossard P. Image registration with sparse approximations in parametric dictionaries // SIAM Journal on Imaging Sciences. – 2013. – Т. 6. – №. 4. – С.2370—2403.

[2] Donoho D. L., Elad M. Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ1 minimization //Proceedings of the National Academy of Sciences. – 2003. – Т. 100. – №. 5. – С.2197—2202.

[3] Потоцкая А. А. и др. Обработка изображений с помощью Shearlets//Сборник статей девятой сибирской конференции по параллельным и высокопроизводительным вычислениям. – 2017. - С.55—61