

УДК 004.932

СИНТЕЗ И ОПТИМИЗАЦИЯ СТОХАСТИЧЕСКОГО АЛГОРИТМА СОВМЕЩЕНИЯ ИЗОБРАЖЕНИЙ ПО КРИТЕРИЮ МАКСИМУМА ВЗАИМНОЙ ИНФОРМАЦИИ ШЕННОНА

© А. Г. Ташлинский¹, Р. М. Ибрагимов¹, Г. Л. Сафина²

¹Ульяновский государственный технический университет,
432027, г. Ульяновск, ул. Северный Венец, 32

²Московский государственный строительный университет,
129337, Москва, Ярославское шоссе, 26
E-mail: tag@ulstu.ru

На основе математического аппарата стохастической градиентной адаптации синтезирован алгоритм оценивания параметров совмещения изображений по критерию максимума взаимной информации Шеннона. Получены выражения для градиента взаимной информации. При этом производные энтропии изображений по оцениваемым параметрам найдены с применением метода окон Парзена. Проведена оптимизация синтезированного алгоритма, направленная на повышение быстродействия.

Ключевые слова: изображение, оценивание, параметр, алгоритм, совмещение, стохастическая процедура, сходимость, взаимная информация Шеннона.

DOI: 10.15372/AUT20240202

EDN: ATXACK

Введение. Возрастающий поток видеоинформации выдвигает новые требования, направленные на улучшение восприятия, анализа, распознавания и интерпретации изображений для принятия решений. Одной из операций, востребованной при решении прикладных задач обработки видеоинформации, является совмещение изображений. Задача совмещения изображений возникает в медицине [1], обнаружении дефектов элементов промышленных объектов [2], построении траекторий беспилотных летательных и подводных аппаратов [3], мониторинге земной поверхности [4], картировании почвенно-растительного покрова при дистанционном зондировании [5], обнаружении изменений [6], идентификации объектов интереса [7], повышении информативности [8] и многих других областях. Решением является поиск функции или параметров заданной функции, устанавливающей соответствия между сопряжёнными точками двух и более изображений. Это могут быть как кадры одной видеопоследовательности, отражающие динамику сцены, так и изображения, сформированные различными устройствами, полученные в разное время, в разных спектральных диапазонах. Результат совмещения может быть как конечным продуктом, так и этапом при решении другой задачи.

В связи с многообразием задач, вызванным различными ограничениями на исходные данные, методы совмещения изображений уже несколько десятилетий интенсивно развиваются и получают всё более широкое распространение. Предложено множество эффективных подходов, ориентированных на различные прикладные задачи. Условно эти подходы можно разделить на две большие категории: совмещение изображений на основе ключевых признаков [9, 10] и на основе интенсивностей [11]. В ситуациях, когда совмещаемые изображения имеют разную природу (например, сформированы в различных спектральных диапазонах) или структуру (получены с разных пространственных координат и/или в разное время), выделение ключевых признаков часто оказывается затруднено. Совмещение изображений с помощью интенсивностей основано на оценке численного значения некоторой

заданной меры подобия исследуемых изображений. Однако при этом реализация процедуры совмещения требует задания некоторой математической модели [12], описывающей взаимные пространственные деформации совмещаемых изображений.

Совмещение разносектральных и разновременных изображений затруднено дополнительными факторами, связанными с тем, что, кроме различного ракурса и масштаба съёмки, такие изображения часто имеют значительные взаимные нелинейные яркостные искажения и импульсные помехи априорно неизвестного вида. При этом традиционные методы, например SIFT [13] и SURF [14], использующие реперные отметки [15], секвентный анализ [16], в сложных условиях имеют низкую точность и достоверность. Хорошую устойчивость к условиям априорной неопределённости обеспечивают процедуры стохастического безыдентификационного совмещения изображений [17] на основе теоретико-информационных мер [18]. При синтезе таких процедур в качестве целевых функций выбираются теоретико-информационные меры подобия, в частности, взаимная информация (ВИ) изображений.

Постановка задачи. Совмещение изображений, как правило, производится попарно. Поэтому будем рассматривать совмещение двух изображений: опорного \mathbf{Z}^r и деформированного \mathbf{Z}^d в предположении, что сопряжённые точки этих изображений связаны некоторой заданной моделью пространственных деформаций с набором параметров α . Определение оптимальных параметров совмещения при использовании математического аппарата безыдентификационной стохастической адаптации [19] сводится к рекуррентному поиску экстремума многомерной целевой функции $Q(\alpha, \mathbf{Z}^r, \mathbf{Z}^d)$ в пространстве параметров α . Вектор параметров, соответствующий экстремуму, и считается оптимальным.

Стохастические алгоритмы оценивания параметров совмещения изображений основаны на процедуре стохастической градиентной адаптации, которую можно записать в виде

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} \pm \Lambda_t \nabla_Q(\hat{\alpha}_{t-1}, \tilde{\mathbf{Z}}_t^r, \mathbf{Z}^d), \quad (1)$$

где $\nabla_Q(\cdot)$ — стохастический градиент целевой функции $Q(\hat{\alpha}_{t-1}, \tilde{\mathbf{Z}}_t^r, \mathbf{Z}^d)$; Λ_t — матрица усиления, определяющая скорость изменения оценок $\hat{\alpha}_t$; $\tilde{\mathbf{Z}}^r$ — изображение \mathbf{Z}^r , передискретизированное на t -й итерации по оценкам $\hat{\alpha}_{t-1}$, полученным на предыдущей итерации; $t = \bar{1}, T$ — номер итерации. Знак в (1) зависит от того, минимумом или максимумом является экстремум целевой функции, например, у ВИ — максимум, поэтому будет знак «+».

Быстродействие (1) определяется в первую очередь вычислительными затратами на нахождение стохастического градиента $\nabla_Q(\cdot)$. Для их уменьшения целесообразно на каждой итерации использовать как можно меньше отсчётов $\tilde{z}_{\mathbf{j}_t}^r \in Z_t^r \subset \tilde{\mathbf{Z}}_t^r$ и $z_{\mathbf{j}_t}^d \in Z_t^d \subset \mathbf{Z}^d$ изображений $\tilde{\mathbf{Z}}^r$ и \mathbf{Z}^d . Данное множество отсчётов будем называть локальной выборкой $Z_t = Z_t^r \cup Z_t^d$, где $\mathbf{j}_t \in \Omega_Z$ — координаты отсчётов, совокупность которых представляет план локальной выборки. При этом координаты могут выбираться как случайным образом, так и по некоторым правилам.

Устойчивость формируемых оценок параметров совмещения к условиям априорной неопределённости изображений и яркостных помех обеспечивает релейная стохастическая процедура, в которой варибельность оценок ограничивает знаковая функция:

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} \pm \Lambda_t \text{sign} \nabla_Q(\hat{\alpha}_{t-1}, Z_t). \quad (2)$$

Задачей исследования является синтез и оптимизация релейного стохастического алгоритма оценивания параметров совмещения изображений при выборе целевой функцией

качества оценивания ВИ Шеннона. Решение этой задачи важно, поскольку оно даёт основу для анализа эффективности других, более упрощённых подходов, в частности, использования метода гистограмм [20] или оценки градиента целевой функции конечными разностями.

Отметим также, что при использовании в качестве меры подобия изображений ВИ необходимо исследование ещё одного аспекта, связанного с оптимизацией плана локальной выборки на итерациях оценивания. Если некий критерий оптимальности применять ко всей локальной выборке, то это может вступить в противоречие с возможностью определения элементов выборки со всей площади изображения. Связано это с тем, что при расчёте ВИ используется как одиночная, так и взаимная энтропии изображений. Одиночная энтропия рассчитывается через разности яркостей элементов выборки. Однако, если координаты этих элементов отстоят друг от друга достаточно далеко, то они слабо коррелированы и не вносят существенного вклада во ВИ. Поэтому требуется также определиться: проводить ли оптимизацию плана всей локальной выборки или выборку разбить на подвыборки и оптимизировать план каждой из подвыборок с последующей интеграцией результатов. Если использовать второй вариант, то подвыборки по заданному правилу могут формироваться по всему изображению и вышеуказанное противоречие снимается.

Синтез алгоритма совмещения изображений на основе взаимной информации Шеннона.

Стохастический градиент ВИ исследуемых изображений. Численные значения ВИ Шеннона [21] на t -й итерации можно найти по известному выражению с использованием энтропии Шеннона:

$$Q_t = \hat{H}(\tilde{\mathbf{Z}}_t^r) + \hat{H}(\mathbf{Z}^d) - \hat{H}(\tilde{\mathbf{Z}}_t^r, \mathbf{Z}^d), \quad (3)$$

где $\hat{H}(\tilde{\mathbf{Z}}_t^r)$ и $\hat{H}(\mathbf{Z}^d)$ — оценки энтропии Шеннона изображений $\tilde{\mathbf{Z}}_t^r$ и \mathbf{Z}^d , $\hat{H}(\tilde{\mathbf{Z}}_t^r, \mathbf{Z}^d)$ — оценка совместной энтропии Шеннона этих изображений.

В свою очередь, оценки энтропий по локальной выборке могут быть найдены как

$$\hat{H}(\tilde{\mathbf{Z}}_t^*) = - \sum_{\mathbf{j} \in \Omega_Z} p(z_{\mathbf{j}_t}^*) \log_2 p(z_{\mathbf{j}_t}^*),$$

$$\hat{H}(\tilde{\mathbf{Z}}_t^r, \mathbf{Z}^d) = - \sum_{\mathbf{j}_i, \mathbf{j}_k \in \Omega_Z} p(\tilde{z}_{\mathbf{j}_i}^r) p(z_{\mathbf{j}_k}^d) \log_2 p(\tilde{z}_{\mathbf{j}_i}^r) p(z_{\mathbf{j}_k}^d),$$

где $p(z_{\mathbf{j}_t}^d)$, $p(\tilde{z}_{\mathbf{j}_t}^r)$ и $p(\tilde{z}_{\mathbf{j}_{ti}}^r, z_{\mathbf{j}_{tk}}^d)$ — значения плотности распределения вероятностей (ПРВ) и совместной ПРВ при яркостях отсчётов $\tilde{z}_{\mathbf{j}_{ti}}^r, z_{\mathbf{j}_{tk}}^d \in Z_t$.

Тогда в соответствии с (3) стохастический градиент ВИ Шеннона определяется выражением:

$$\nabla_{Q_t} = \frac{\partial \hat{H}(\tilde{\mathbf{Z}}_t^r)}{\partial \alpha} - \frac{\partial \hat{H}(\tilde{\mathbf{Z}}_t^r, \mathbf{Z}^d)}{\partial \alpha}. \quad (4)$$

Таким образом, нахождение стохастического градиента ВИ Шеннона сводится к определению на каждой итерации оценок ПРВ и совместной ПРВ изображений $\tilde{\mathbf{Z}}_t^r$ и \mathbf{Z}^d , а также производных энтропий по оцениваемым параметрам.

Оценка ПРВ опорного и деформированного изображений. Для нахождения ВИ на каждой итерации процедуры (2) требуется оценка одиночной и совместной ПРВ изображений $\tilde{\mathbf{Z}}_t^r$ и \mathbf{Z}^d по локальной выборке Z_t , имеющей небольшой объём (единицы, десятки отсчётов),

что предъявляет к качеству аппроксимации ПРВ повышенные требования. Оценки ПРВ могут быть найдены либо с использованием гистограммного метода [20], либо по методу окон Парзена [22]. Исследования показывают, что более информативным для рассматриваемой задачи является метод окон Парзена, предполагающий к тому же несложную программную реализацию. Идея этого метода состоит в том, что оценка ПРВ $p_t(z)$ находится как нормированная суперпозиция элементарных симметричных распределений, центрированных на интенсивностях отсчётов $\tilde{z}_{\mathbf{j}t}^d$, $\mathbf{j}t \in \Omega_Z$, попавших в локальную выборку. В частности, если в качестве элементарного распределения выбрано гауссово, то

$$p_t(z) = \frac{1}{\mu} \sum_{\mathbf{j}t \in \Omega_Z} \frac{1}{\sqrt{2\pi} \sigma_G} \exp\left(-\frac{(z - \tilde{z}_{\mathbf{j}t}^d)^2}{2\sigma_G^2}\right),$$

где μ — объём локальной выборки, соответствующий мощности множества Ω_Z ; σ_G^2 — дисперсия элементарного распределения, которая выбирается по некоторому критерию и существенно влияет на качество оценок энтропии. Вместо гауссова распределения можно использовать и любую другую подходящую дифференцируемую функцию, например ПРВ Коши. Однако особенности гауссовой функции упрощают последующий анализ, поэтому в настоящем исследовании выбрана она.

Производные энтропии ВИ по оцениваемым параметрам. Для расчёта стохастического градиента ВИ (4) необходимо найти частные по оцениваемым параметрам совмещения одиночной энтропии изображения $\tilde{\mathbf{Z}}_t^r$ и совместной энтропии изображений $\tilde{\mathbf{Z}}_t^r$ и \mathbf{Z}^d . Это можно сделать либо численными методами с использованием конечных разностей, либо, как в данной работе, получить аналитические выражения с учётом принятой модели взаимных пространственных деформаций изображений.

Известен подход к нахождению оценок энтропии Шеннона, основанный на том, что она может быть выражена как математическое ожидание отрицательного логарифма ПРВ [23]. При этом оценки ПРВ находятся по одной локальной выборке, а среднего значения логарифма ПРВ — по другой. Применение двух выборок ведёт к неэффективному использованию содержащейся в них информации. В данной работе для аналитического нахождения производных энтропии применяется другой подход, при котором используется только одна выборка из изображений $\tilde{\mathbf{Z}}_t^r$ и \mathbf{Z}^d . Упуская промежуточные выкладки, приведём полученные выражения.

При известной Z_t частные производные для энтропии Шеннона можно найти как

$$\frac{\partial \hat{H}_t(\tilde{\mathbf{Z}}^r)}{\partial \alpha} = \frac{1}{\mu \sigma_G^2} \sum_{z_k^r \in Z_t} \left(\left(\lg \left(\sum_{z_i^r \in Z_t} \frac{\nabla_{ik}^r}{\mu} \right) + 1 \right) \sum_{z_i^r \in Z_t} \nabla_{ik}^r \Delta_{ik}^r \frac{\partial \Delta_{ik}^r}{\partial \alpha} \right), \quad (5)$$

где $\nabla_{ik}^r = p_t^r(z_i^r - z_k^r)$; $\Delta_{ik}^r = z_i^r - z_k^r$; z_i^r и z_k^r — яркости отсчётов $\tilde{z}_{\mathbf{j}t}^r \in Z_t$.

Аналогичные соотношения можно записать и для $\partial \hat{H}_{St}(\mathbf{Z}^d, \tilde{\mathbf{Z}}_t^r) / \partial \alpha$:

$$\frac{\partial \hat{H}_t(\mathbf{Z}^d, \tilde{\mathbf{Z}}_t^r)}{\partial \alpha} = \frac{1}{\mu \sigma_G^2} \sum_{z_k^r, z_k^d \in Z_t} \left(\ln \left(\frac{1}{\mu} \sum_{z_i^r, z_i^d \in Z_t} \nabla_{ik}^r \nabla_{ik}^d \right) + 1 \right) \left(\sum_{z_i^r, z_i^d \in Z_t} \nabla_{ik}^d \nabla_{ik}^r \Delta_{ik}^r \frac{\partial \Delta_{ik}^r}{\partial \alpha} \right), \quad (6)$$

где $\nabla_{ik}^d = w_t(z_i^d - z_k^d)$; $\Delta_{ik}^d = z_i^d - z_k^d$; z_i^d и z_k^d — яркости отсчётов $z_{\mathbf{j}t}^d \in Z_t$.

Отметим, что при суммировании в выражениях (5) и (6) $i \neq k$.

Полученные формулы для стохастического градиента ВИ Шеннона позволяют в соответствии с процедурой (2) разрабатывать релейные алгоритмы оценивания параметров совмещения изображений.

Оптимизация плана локальной выборки. Учтём, что в выражениях (5) и (6) производные $\partial\Delta_{ik}^r/\partial\alpha$ зависят от принятой модели пространственных деформаций. Представим $\partial\Delta_{ik}^r/\partial\alpha$ в виде

$$\frac{\partial\Delta_{ik}^r}{\partial\alpha} = \frac{\partial\Delta_{ik}^r}{\partial\mathbf{j}} \frac{\partial\mathbf{j}}{\partial\alpha} = \frac{\partial(z_i^r - z_k^r)}{\partial j_1} \frac{\partial j_1}{\partial\alpha} + \frac{\partial(z_i^r - z_k^r)}{\partial j_2} \frac{\partial j_2}{\partial\alpha}. \quad (7)$$

При этом, поскольку зависимость яркостей z_i^r отсчётов $\tilde{z}_{\mathbf{j}_t}^r \in Z_t^r$ от базовых координат \mathbf{j} изображения \mathbf{Z}^r априорно не известна, производную $\partial(\tilde{z}_{\mathbf{j}_t}^r - z_k^r)/\partial\mathbf{j}$ можно оценить только через конечные разности. Например, в простейшем случае, полагая изображение двумерным $\mathbf{j} = (j_1, j_2)^\top$ и используя интерполяцию второго порядка, производную $\partial z_{\mathbf{j}_t}^r/\partial j_1$ в узле $(j_1 = n, j_2 = m)$ сетки отсчётов можно оценить как

$$\left. \frac{\partial z_{\mathbf{j}_t}^d}{\partial j_1} \right|_{j_1=n, j_2=m} \cong \frac{z_{n+\Delta_{j_1}, m} - z_{n-\Delta_{j_1}, m}}{2\Delta_{j_1}},$$

где Δ_{j_1} — некоторое приращение значения координаты j_1 отсчёта $z_{\mathbf{j}_t}^d$.

Рассмотрим выражение (7) в предположении, что изображения $\tilde{\mathbf{Z}}_t^r$ и \mathbf{Z}^d имеют гауссово распределение яркостей с нулевым средним и в качестве модели исследуемых изображений принята аддитивная смесь полезного изображения \mathbf{S} и помех Θ :

$$\tilde{z}_{\mathbf{j}_t}^r = \tilde{s}_{\mathbf{j}_t}^r + \theta_{\mathbf{j}_t}^r, \quad z_{\mathbf{j}_t}^d = s_{\mathbf{j}_t}^d + \theta_{\mathbf{j}_t}^d,$$

где $s_{\mathbf{j}_t}^d, \tilde{s}_{\mathbf{j}_t}^r \in \mathbf{S}$ — отсчёты изображения \mathbf{S} с нулевым средним и дисперсией σ_s^2 ; $\theta_{\mathbf{j}_t}^d, \theta_{\mathbf{j}_t}^r \in \Theta$ — независимые центрированные гауссовы помехи с дисперсией σ_θ^2 .

С учётом принятых ограничений и, введя для краткости обозначения: $\nu_{1i} = \partial j_1/\partial\alpha|_{z_i^r}$, $\nu_{1k} = -\partial j_1/\partial\alpha|_{z_k^r}$, $\nu_{2i} = \partial j_2/\partial\alpha|_{z_i^r}$, $\nu_{2k} = -\partial j_2/\partial\alpha|_{z_k^r}$, получаем:

$$\begin{aligned} \frac{\partial\Delta_{ik}^r}{\partial\alpha} &= (z_i^r - z_k^r) \left[(z_i^r(\Delta_{j_1}) - z_i^r(-\Delta_{j_1}))\nu_{1i} + \right. \\ &\left. + (z_i^r(\Delta_{j_2}) - z_i^r(-\Delta_{j_2}))\nu_{2i} - (z_k^r(\Delta_{j_1}) - z_k^r(-\Delta_{j_1}))\nu_{1k} + (z_k^r(\Delta_{j_2}) - z_k^r(-\Delta_{j_2}))\nu_{2k} \right], \quad (8) \end{aligned}$$

где $z_i^r(\Delta_{j_1})$ — яркость изображения $\tilde{\mathbf{Z}}_t^r$ в точке с координатами $(j_1 + \Delta_{j_1}, j_2)$; (j_1, j_2) — координаты точки z_i^r , аналогично и для Δ_{j_2} ; $\partial j_*/\partial\alpha|_{z_{i,k}^r}$ — значение производной по соответствующей координате в точке отсчёта $z_{i,k}^r$.

Опуская громоздкие промежуточные преобразования, запишем выражения для математического ожидания и дисперсии $\partial\Delta_{ik}^r/\partial\alpha$. Математическое ожидание:

$$M \left[\frac{\partial\Delta_{ik}^r}{\partial\alpha} \right] = \sigma_s^2 \left[(\nu_{1i} - \nu_{1k}) (R(-\Delta_{j_1}) - R(\Delta_{j_1})) + (\nu_{2i} - \nu_{2k}) (R(-\Delta_{j_2}) - R(\Delta_{j_2})) \right], \quad (9)$$

где $R(\Delta_{j_1}), R(-\Delta_{j_1})$ — значения автокорреляционной функции (АКФ) изображения \mathbf{S} для расстояния между координатами точек z_*^d и $z_*^r(\Delta_{j_1})$; аналогично и для Δ_{j_2} .

Соответственно, дисперсия:

$$D \left[\frac{\partial\Delta_{ik}^r}{\partial\alpha} \right] = 4\sigma_s^2 g \left(\nu_{1i}^2 \nu_{1k}^2 \nu_{2i}^2 \nu_{2k}^2 \right) \left[(1 - R(2)) (1 - R(0)) + g^{-1} (2 - R(0) - R(2\Delta_j) + g) + \right.$$

$$\begin{aligned}
 & + 0,5 \left(\left(\nu_{1i} - \nu_{1k} \right) \times \left(R(-\Delta_{j_1}) - R(\Delta_{j_1}) \right) + \left(\nu_{2i} - \nu_{2k} \right) \left(R(-\Delta_{j_2}) - R(\Delta_{j_2}) \right) \right)^2 + \\
 & + \nu_{1i}\nu_{1k} \left(R(-\Delta_{j_1}) - R(\Delta_{j_1}) \right)^2 - \left(R(2\Delta_{j_1}) - R(2\Delta_{j_1}) - 2R(0) \right) \left(1 + g^{-1} \right) + \\
 & + \nu_{2i}\nu_{2k} \left(R(-\Delta_{j_2}) - R(\Delta_{j_2}) \right)^2 - \left(R(2\Delta_{j_2}) - R(2\Delta_{j_2}) - 2R(0) \right) \left(1 + g^{-1} \right) + \\
 & + \left(\nu_{1i}\nu_{1k} + \nu_{2i}\nu_{2k} \right) \left(R(\pm\Delta_j) - R(\mp\Delta_j) + R(\pm\Delta_j)R(\mp\Delta_j) \right), \tag{10}
 \end{aligned}$$

где $R(\pm\Delta_j)$, $R(\mp\Delta_j)$ — значения АКФ изображения \mathbf{S} для расстояния между координатами точек z_*^d и $z_*^r(\Delta_{j_1}, -\Delta_{j_2})$; $g = \sigma_s^2/\sigma_\theta^2$ — отношение сигнал/шум.

Максимум полезной информации из выражения (8) будет извлекаться при максимуме модуля отношения математического ожидания $\partial\Delta_{ik}^r/\partial\alpha$ градиента к его СКО:

$$\max \left| \frac{M[\partial\Delta_{ik}^r/\partial\alpha]}{\sqrt{D[\partial\Delta_{ik}^r/\partial\alpha]}} \right|. \tag{11}$$

Условие (11) и найденные выражения (9) и (10) позволяют оптимизировать план локальной выборки.

Пример оптимизации плана локальной выборки. Производная $\partial j/\partial\alpha$ может быть найдена аналитически по принятой модели взаимных пространственных деформаций изображений \mathbf{Z}^d и \mathbf{Z}^r . Например, при использовании частного случая общей аффинной модели — модели подобия [13]:

$$j_1 = \kappa \left(h_{j_1}^* \cos \varphi - h_{j_2}^* \sin \varphi \right) + h_{j_1} + h_{j_1 0}, \quad j_2 = \kappa \left(h_{j_1}^* \sin \varphi + h_{j_2}^* \cos \varphi \right) + h_{j_2} + h_{j_2 0},$$

параметрами $\alpha = \left(h_{j_1}, h_{j_2}, \varphi, \kappa \right)^\top$ которой являются сдвиги h_{j_1} и h_{j_2} по базовым осям изображений, угол поворота φ и коэффициент масштаба κ :

$$\frac{\partial j_1}{\partial j_1} = \frac{\partial j_2}{\partial j_2} = 1, \quad \frac{\partial j_1}{\partial \varphi} = \kappa \left(-h_{j_1}^* \sin \varphi - h_{j_2}^* \cos \varphi \right), \quad \frac{\partial j_2}{\partial \varphi} = \kappa \left(-h_{j_1}^* \cos \varphi - h_{j_2}^* \sin \varphi \right), \tag{12}$$

$$\frac{\partial j_1}{\partial j_2} = \frac{\partial j_2}{\partial j_1} = 0, \quad \frac{\partial j_1}{\partial \kappa} = \left(h_{j_1}^* \cos \varphi - h_{j_2}^* \sin \varphi \right), \quad \frac{\partial j_2}{\partial \kappa} = \left(h_{j_1}^* \sin \varphi + h_{j_2}^* \cos \varphi \right),$$

где $h_{j_1}^* = \left(h_{j_1} - h_{j_1 0} \right)$; $h_{j_2}^* = \left(h_{j_2} - h_{j_2 0} \right)$; $\left(h_{j_1 0}, h_{j_2 0} \right)$ — координаты центра поворота.

В качестве примера рассмотрим оптимизацию плана при минимальном объёме локальной подвыборки, равном трём. Для равных условий отсчёты $z_{jt}^d \in Z_t^d$ подвыборки будем выбирать из условия их расположения примерно на окружности радиуса ρ . Тогда при оптимальном расстоянии между элементами подвыборки все разности яркостей будут давать в среднем максимум информации для выполнения итерации стохастической процедуры в нужном направлении. Если, например, возможными деформациями являются

только сдвиги h_{j_1} и h_{j_2} , то в соответствии с (12) $\partial j_1/\partial\varphi = \partial j_2/\partial\varphi = \partial j_1/\partial\kappa = \partial j_2/\partial\kappa = 0$. Несложно показать, что в этом случае выражения (9) и (10) принимают вид

$$M \left[\frac{\partial \Delta_{ik}^r}{\partial \alpha} \right] = \frac{2\sigma_s^2}{\Delta_j} \left(R(-\Delta_j) - R(\Delta_j) \right), \quad (13)$$

$$D \left[\frac{\partial \Delta_{ik}^r}{\partial \alpha} \right] = 4\sigma_s^2 \left(2 - R(\sqrt{3}\rho) - R(2\Delta_j) + 2g^{-1} \right) + \left(1 - R(\sqrt{3}\rho) \right) \left(1 - R(2\Delta_j) \right) - \left(1 - R(\sqrt{3}\rho) + g^{-1} \right) \left(R(\mp\Delta_j) + R(\pm\Delta_j) - 2R(\sqrt{3}\rho) \right) + \left(R(-\Delta_j) - R(\Delta_j) \right)^2. \quad (14)$$

Как видно из (14), оптимальный план подвыборки зависит от АКФ изображений и отношения сигнал/шум.

Используя (13) и (14) и условие (11), можно найти оптимальный план выбора отсчётов подвыборки (оптимальное значение ρ_{opt}), обеспечивающий в условиях аддитивных шумов извлечение максимума информации из производной $\partial \Delta_{ik}^r/\partial h$.

Экспериментальные результаты. Экспериментальные исследования проводились с использованием релейного стохастического алгоритма совмещения изображений, синтезированного на основе выражений (5) и (6). Для апробации алгоритма использовались имитированные и реальные изображения оптического диапазона. При этом имитированные изображения формировались с помощью волновой модели [24], достоинством которой при экспериментальной проверке аналитических результатов является изотропность получаемых изображений, а также близость ПРВ яркостей и АКФ к гауссовым функциям, поскольку именно такие допущения часто используются при получении аналитических результатов.

При проведении эксперимента в (2) использовалась положительно определённая диагональная матрица усиления с постоянными коэффициентами обучения. Для нахождения дисперсии σ_G^2 элементарного гауссова распределения использовался известный подход [23], при котором σ_G^2 определяется через оценку дисперсии локальной выборки Z_t . Число итераций алгоритмов $T = 250$. Локальная выборка Z_t на каждой итерации состояла из подвыборок по 3 отсчётам. План подвыборок оптимизировался в соответствии с критерием (11). Подвыборки формировались по изображению случайным образом. Координаты отсчётов подвыборок определялись с точностью до 0,5 шага сетки отсчётов от окружности радиуса ρ_{opt} . Быстродействие алгоритмов оценивалось числом итераций до достижения сходимости оценок параметров совмещения (числом итераций сходимости). Гипотеза о достижении сходимости к оптимальным значениям считалась подтверждённой, если дисперсия и среднее значение погрешности оценки параметра в скользящем окне не превысили заданных пороговых значений.

Пример имитированного изображения приведён на рис. 1, *a*, а его АКФ — на рис. 1, *b*. Радиус корреляции изображения составляет примерно 37 шагов сетки отсчётов. При отсутствии шумов ($g = \infty$) получаем $\rho_{\text{opt}} \approx 58$, при $g = 10$ — $\rho_{\text{opt}} \approx 51$.

На рис. 2, *a* показаны экспериментальные зависимости количества итераций для сходимости разработанного алгоритма от параметра ρ (кривая 1 — для $g = \infty$, кривая 2 — для $g = 10$). Для ситуации без шума использовалась одна подвыборка, для ситуации с шумом — 9 подвыборок. Результаты усреднены по 30 реализациям. Видно, что зависимости имеют явно выраженные минимумы, соответствующие примерно ρ_{opt} . При этом отклонение от ρ_{opt} на 80 % почти вдвое уменьшает скорость сходимости алгоритмов.

Выше приведены результаты эксперимента при использовании имитированных изображений в условиях гауссовых шумов. Но как уже отмечалось, оптимальное значение

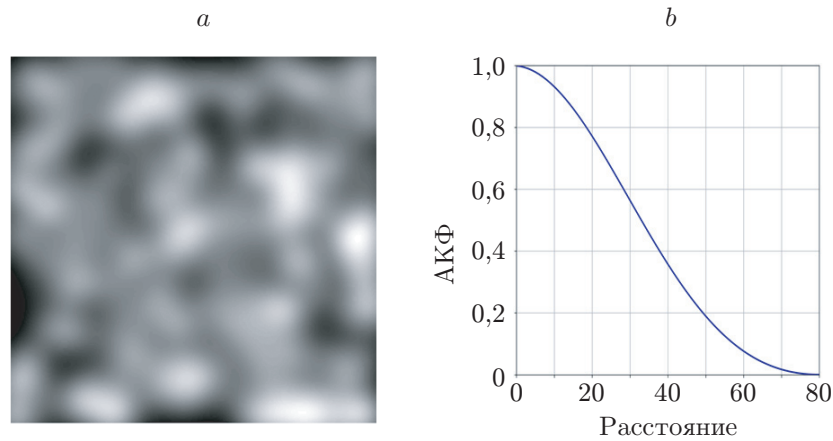


Рис. 1. Имитированное изображение (a) и его АКФ (b)

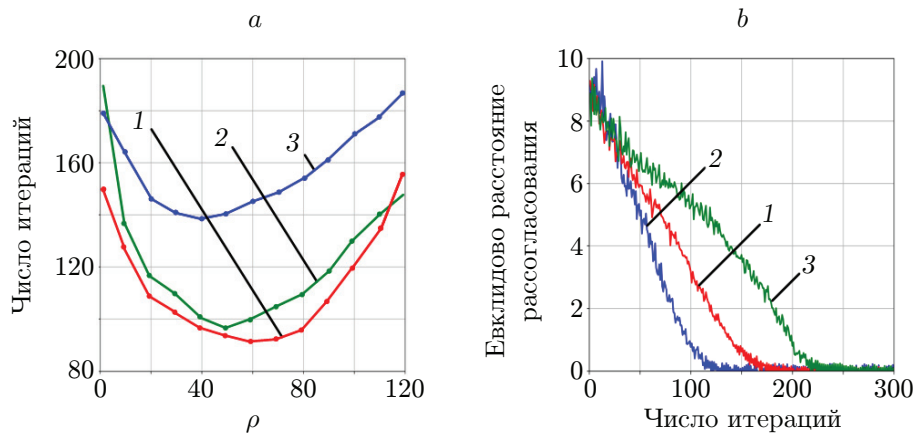


Рис. 2. Графики зависимостей: числа итераций сходимости от радиуса окружности (a) и евклидова расстояния рассогласования от числа итераций (b)

определяется АКФ и отношением сигнал/шум и мало зависит от вида ПРВ и природы формирования изображений. Так, кривая 3 на рис. 2, a построена при тех же параметрах деформаций для реального изображения инфракрасного диапазона радиуса корреляции, равного 24, для которого $\rho_{\text{opt}} \approx 40$.

На рис. 2, b приведены зависимости от числа итераций евклидова расстояния рассогласования [25], которое интегрально характеризует вектор оценок параметров совмещения. Оценивался вектор параметров: $h_{j_1} = 3,7$, $h_{j_2} = -4,1$, $\varphi = 5,5$, $\kappa = 1,06$ для ситуации без оптимизации (кривая 1) и с оптимизацией (кривая 2) при равных объёмах локальной выборки. Объём локальной выборки составлял 12 отсчётов (четыре подвыборки по три отсчёта). Из графиков видно, что при оптимизации скорость сходимости алгоритма увеличивается примерно в 1,7 раза. На этом же графике приведены результаты (кривая 3) для ситуации нахождения энтропии изображений алгоритмом ЕММА [23]. Использование алгоритма ЕММА проигрывает результату с оптимизацией более, чем в 2 раза.

Таким образом, экспериментальные результаты подтвердили предположение об эффективности оптимизации плана локальной выборки и адекватность полученных аналитических выводов.

Заключение. Релейный стохастический алгоритм совмещения изображений, разработанный на основе полученных аналитических выражений для стохастического градиен-

та ВИ Шеннона, показал высокую эффективность. В частности, при размерах исследуемых изображений 400×400 общее число отсчётов, использованных на всех итерациях оценивания, не превышало 1000, что составляет 0,7 % от их общего числа. При увеличении объёма локальной выборки возрастают точность и устойчивость оценок параметров совмещения, но также растёт и вычислительная трудоёмкость.

Показано, что оптимизация плана локальной выборки, по которой на каждой итерации оценивания определяется стохастический градиент ВИ Шеннона, позволяет значительно (в проведённых экспериментах до двух раз) улучшить быстродействие алгоритма. При этом при оптимизации плана локальной выборки её целесообразно разделить на подвыборки небольшого объёма, провести оптимизацию плана подвыборок, затем при нахождении энтропии для расчёта стохастического градиента ВИ данные подвыборок интегрировать. При этом подвыборки могут формироваться произвольно или по какому-либо правилу по всему изображению.

Финансирование. Работа выполнена при поддержке Российского научного фонда в рамках научного проекта № 22-21-00513, <https://rscf.ru/en/project/22-21-00513/>.

СПИСОК ЛИТЕРАТУРЫ

1. **Fida A. D., Gaidel A. V., Demin N. S. et al.** Automated combination of optical coherence tomography images and fundus images // *Comp. Opt.* 2021. **45**, N 5. P. 721–727. DOI: 10.18287/2412-6179-CO-892.
2. **Dementiev V. E., Magdeev R. G., Tashlinskii A. G.** Detecting anomalies in temporal image sequences based on object identification by the stochastic gradient adaptation // *Proc. of the IEEE Xplore VII Int. Conf. Information Technology and Nanotechnology (ITNT)*. Samara, Russia, 20–24 Sept., 2021. DOI: 10.1109/ITNT52450.2021.9649175.
3. **Kamaev A. N., Karmanov D. A.** Visual navigation of an autonomous underwater vehicle based on the global search of image correspondences // *Comp. Opt.* 2018. **42**, N 3. P. 457–467. DOI: 10.18287/2412-6179-2018-42-3-457-467.
4. **Фролов В. Н., Тупиков В. А., Павлова В. Ф., Александров В. А.** Методы информационного совмещения изображений в многоканальных оптико-электронных системах // *Изв. Тульского гос. ун-та*. 2016. Вып. 11, № 3. С. 95–104.
5. **Zhang H., Xu R.** Exploring the optimal integration levels between SAR and optical data for better urban land cover mapping in the Pearl River Delta // *Int. Journ. Appl. Earth Obs. Geoinf.* 2018. **64**. P. 87–95. DOI: 10.1016/j.jag.2017.08.013.
6. **Wan L., Xiang Y., You H.** A post-classification comparison method for SAR and optical images change detection // *IEEE Geosci. Remote Sensing Lett. (GRSL)*. 2019. **16**, Iss. 7. P. 1026–1030. DOI: 10.1109/LGRS.2019.2892432.
7. **Magdeev R. G., Tashlinskii A. G.** Efficiency of object identification for binary images // *Comp. Opt.* 2019. **43**, N 2. P. 277–281. DOI: 10.18287/2412-6179-2019-43-2-277-281.
8. **Marcos D., Hamid R., Tuia D.** Geospatial correspondences for multimodal registration // *Proc. of the IEEE Conf. Computer Vision and Pattern Recognition*. Las Vegas, USA, 27–30 June, 2016. P. 5091–5100.
9. **Park H., Bland P. H., Brock K. K., Meyer C. R.** Adaptive registration using local information measures // *Medical Image Analysis*. 2004. **8**, Iss. 4. P. 465–473. DOI: 10.1016/j.media.2004.03.001.
10. **Can A., Stewart C., Roysam B., Tanenbaum H.** A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina // *IEEE Trans. Pattern Analysis Machine Intell.* 2002. **24**, Iss. 3. P. 347–363. DOI: 10.1109/34.990136.

11. **Mani V. R. S., Arivazhagan S.** Survey of medical image registration // Journ. Biomed. Eng. Technol. 2013. **1**, Iss. 2. P. 8–25. DOI: 10.12691/jbet-1-2-1.
12. **Gonzalez R. C., Woods E.** Digital Image Processing. Chapter 2: Digital Image Fundamentals. London: Pearson Publisher, 2018. P. 34–74.
13. **Wu J., Cui Zh., Sheng V. S. et al.** A comparative study of SIFT and its variants // Measurement Sci. Rev. 2013. **13**, Iss. 3. P. 122–131. DOI: 10.2478/msr-2013-0021.
14. **Bay H., Ess A., Tuytelaars T., Gool L. V.** SURF: Speeded-up robust features // Computer Vision and Image Understanding. 2008. **110**, N 3. P. 346–359. DOI: 10.1016/j.cviu.2007.09. 014.
15. **Алексанин А. И., Морозов М. А., Фомин Е. В.** Проблемы совмещения изображений с пиксельной точностью // Современные проблемы дистанционного зондирования Земли из космоса. 2019. **16**, № 1. С. 9–16. DOI: 10.21046/2070-7401-2019-16-1-9-16.
16. **Злобин В. К., Колесенков А. Н., Костров Б. В.** Корреляционно-экстремальные методы совмещения аэрокосмических изображений // Вестн. Рязанского гос. радиотехнич. ун-та. 2011. Вып. 37, № 3. С. 12–17.
17. **Tashlinskii A. G., Safina G. L., Kovalenko R. O., Ibragimov R. M.** Usage of mutual information as similarity measures for stochastic binding images // Proc. of the IEEE Xplore VIII Int. Conf. Information Technology and Nanotechnology (ITNT). Samara, Russia, 20–24 Sept., 2021. DOI: 10.1109/ITNT52450.2021.9649386.
18. **Tashlinskii A., Ibragimov R., Safina G.** Application of Renyi Mutual Information in Stochastic Referencing of Multispectral and Multi-temporal Images // Proc. of the IEEE Xplore VIII Int. Conf. Information Technology and Nanotechnology (ITNT). Samara, Russia, 23–27 May, 2022. DOI: 10.1109/ITNT5410.2022.9848648.
19. **Цыпкин Я. З.** Информационная теория идентификации. М.: Физматлит, 1995. 336 с.
20. **Kovalenko R. O., Tashlinskii A. G.** Optimization of the Histogram Intervals Number which Approximate Brightness Probability Distributions in Stochastic Image Alignment Based on Information Similarity Measures // Proc. of the IEEE Xplore 24th Int. Conf. Digital Signal Processing and its Applications (DSPA). Moscow, Russia, 30 March – 1 April, 2022. DOI: 10.1109/DSPA53304.2022.9805456.
21. **Shannon C. E., Weaver W.** The Mathematical Theory of Communication. Urbana: University of Illinois Press, 1998. 144 p.
22. **Mussa H. Y., Mitchell J. B. O., Afzal A. M.** The Parzen Window method: In terms of two vectors and one matrix // Patt. Recogn. Lett. 2015. **63**. P. 30–35. DOI: 10.1016/j.patrec.2015.06. 002.
23. **Viola P., Wells III W. M.** Alignment by maximization of mutual information // Int. Journ. Comp. Vis. **24**, Iss. 2. 1997. P. 137–154.
24. **Крашенинников В. Р.** Основы теории обработки изображений. Ульяновск: УлГТУ, 2003. 152 с.
25. **Ташлинский А. Г., Сафина Г. Л., Коваленко Р. О.** Оптимальное евклидово расстояние рассогласования оценок при псевдоградиентном оценивании параметров межкадровых геометрических деформаций изображений // Информационно-измерительные и управляющие системы. 2018. **16**, № 11. С. 33–39.

Поступила в редакцию 12.07.2023

После доработки 16.08.2023

Принята к публикации 24.08.2023