

ВЫБОР МОДЕЛИ ПРОСТРАНСТВЕННОГО ПРЕОБРАЗОВАНИЯ ИЗОБРАЖЕНИЙ ПО КРИТЕРИЮ МИНИМАЛЬНОЙ ДЛИНЫ ОПИСАНИЯ

© 2007 г. А. С. Потапов, канд. техн. наук

НПК «Государственный оптический институт им. С.И. Вавилова, Санкт-Петербург

Рассмотрена задача определения лучшего пространственного преобразования по набору опорных точек, найденных в процессе сопоставления двух изображений. Показано, что использование критерия среднеквадратичного отклонения допустимо только при выборе среди преобразований с одинаковым числом параметров. В противном случае предпочтение отдается наиболее сложному преобразованию, которое, однако, не обеспечивает наилучшей точности совмещения изображений. Предложен критерий, построенный на основе принципа минимальной длины описания, для выбора среди различных классов преобразований (движения, подобия, аффинного и др.). Экспериментально установлено, что выбор пространственного преобразования на основе этого критерия позволяет увеличить точность совмещения, особенно при небольшом числе опорных точек или в случае, когда положения точек заданы с большими ошибками, что характерно для задач структурного совмещения изображений, подверженных изменениям.

Коды OCIS: 150.0150.

Поступила в редакцию 13.12.2006.

1. Введение

Выбор модели, наилучшим образом описывающей некоторый набор данных, является общей задачей, встречающейся во всех областях научной деятельности. Существует множество методов решения этой задачи для каждого конкретного случая. Все эти методы условно можно разделить на параметрические и непараметрические. В случае построения параметрических моделей число параметров обычно считается фиксированным, что позволяет сравнивать модели лишь на основе той ошибки, с которой они описывают имеющиеся данные (например, среднеквадратичного отклонения (СКО)).

Однако достаточно часто может возникать потребность сравнивать модели с разным числом параметров. Оказывается, что минимизация ошибки, с которой данные описываются моделью, в результате дает модель с большим числом параметров, очень хорошо описывающую имеющиеся данные, но плохо предсказывающую новые данные. В работах по искусственным нейронным сетям этот эффект известен, как низкая обобщающая способность в результате переобучения. Таким образом, при сравнении моделей с разным числом параметров критерии наподобие СКО недостаточны для оценки качества моделей.

Классическим способом решения данной проблемы является разделение имеющихся данных на обучающую и тестовую выборки, что позволяет непосредственно проверять прогностическую силу модели. При этом, однако, теряется значительная часть информации, что может оказаться критичным при малом объеме имеющихся данных.

Другим подходом является включение сложности модели (например, числа ее параметров) в критерий качества модели совместно с ошибкой описания. В байесовских методах это соответствует заданию априорных вероятностей моделей на основе их сложности. Однако как именно следует учитывать сложность модели, остается не совсем ясным.

Формализация понятия сложности модели выполняется в методах, основанных на теории информации. Здесь сложность модели рассматривается как число бит, необходимых, чтобы закодировать эту модель. Это обычно выражается в принципе минимальной длины описания (МДО) для выбора модели [1]: следует выбирать ту модель, которая минимизирует сумму длин описания модели и данных, закодированных с помощью этой модели.

В настоящей статье рассматривается практическое применение принципа МДО для решения задачи нахождения пространственного преобразования между двумя наборами точек. Данная задача возникает при совмещении изображений, т. е. при установлении пространственного преобразования, отображающего одно изображение в систему координат другого изображения таким образом, чтобы соответствующие точки на двух изображениях совпадали. Практически всегда полагается, что тип пространственного преобразования известен априори [2, 3], но для разработки достаточно универсальных алгоритмов совмещения изображений от этого условия необходимо отказаться. Именно задача выбора между преобразованиями с различным числом параметров и является основной темой данной работы. При этом мы воспользуемся алгоритмами

поиска опорных точек и отождествления точек на двух изображениях, которые нами были описаны ранее [4–6].

Большое внимание здесь уделяется практической эффективности разрабатываемых алгоритмов и демонстрации того, как именно может применяться принцип МДО для решения практических проблем.

2. Постановка задачи

Пусть даны два набора отождествленных опорных точек $\{(x_i, y_i)\}_{i=1}^N$ и $\{(x'_i, y'_i)\}_{i=1}^N$, где x_i, y_i – координаты i -й опорной точки на первом изображении, x'_i, y'_i – координаты соответствующей ей опорной точки на втором изображении, а N – число пар опорных точек. И пусть задано множество пространственных преобразований $\{g_k(x, y, \mathbf{a}_k): R^2 \rightarrow R^2 | \mathbf{a}_k \in R^{d_k}\}_{k=1}^K$, где K – количество классов преобразований, $\mathbf{a}_k$ – d_k -компонентный вектор параметров k -го класса. Здесь мы включаем в это множество все возможные преобразования параллельного переноса, группы движения, группы подобия, аффинные преобразования, билинейные и полиномиальные (точнее, квадратичные) преобразования.

Выбор допустимых классов преобразований имеет значение только с точки зрения приложения, но не самого рассматриваемого подхода. Перечисленные преобразования характерны для задачи совмещения аэрокосмических изображений.

В рассматриваемой задаче считается, что класс, к которому принадлежит пространственное преобразование, которое требуется найти, априори неизвестен, а опорные точки заданы с некоторой погрешностью. Это приводит к следующей широко известной трудности. Если данные наборы опорных точек связаны преобразованием более простого вида, чем используемая модель, то “лишние” коэффициенты модели будут отличны от нуля. Такая излишне усложненная модель будет с меньшей ошибкой описывать данный набор точек, но с большей ошибкой отображать точки, не вошедшие в выборку опорных точек.

Таким образом, более сложной модели будет соответствовать меньшее значение СКО и в то же время она может оказаться хуже, чем более простая модель. Как сложность, так и точность модели адекватно учитываются в критерии МДО, что позволяет решить проблему выбора между пространственными преобразованиями разной сложности.

3. Оценка длины описания

Для произвольного пространственного преобразования $g_k(x, y, \mathbf{a}_k)$ запишем связь между опорными точками двух изображений:

$$\begin{cases} x'_i = g_{x,k}(x_i, y_i, \mathbf{a}_k) + e_{x,i}, \\ y'_i = g_{y,k}(x_i, y_i, \mathbf{a}_k) + e_{y,i}, \end{cases} \quad (1)$$

где $\mathbf{e}_i = (e_{x,i}, e_{y,i})$ – величины невязок. Тогда координаты точек второго множества $\{(x'_i, y'_i)\}_{i=1}^N$ могут быть восстановлены по координатам точек первого множества $\{(x_i, y_i)\}_{i=1}^N$, если известны номер k класса выбранной модели, вектор значений параметров этой модели $\mathbf{a}_k$ и величины невязок $\{(e_{x,i}, e_{y,i})\}_{i=1}^N$.

Согласно принципу МДО лучшей будет та модель, которая позволяет минимизировать суммарную длину описания, состоящего из перечисленных компонентов. Длину описания точек первого множества $\{(x_i, y_i)\}_{i=1}^N$ можно не учитывать, так как это постоянное слагаемое, не зависящее от рассматриваемой модели. Также можно условиться, что номера классов представляются кодовыми словами с одинаковым числом бит, и не учитывать это значение. Это соответствует заданию одинаковых априорных вероятностей для разных классов моделей. Если эти априорные вероятности известны (или могут быть определены в процессе многократного применения системы), то внесение дополнительного слагаемого в оценку длины описания не вызывает затруднений.

Таким образом, остаются два компонента, длины описания которых требуется оценить, – это значения параметров $\mathbf{a}_k$ модели данного класса k (регулярная составляющая) и невязки $\{(e_{x,i}, e_{y,i})\}_{i=1}^N$ (случайная составляющая). Рассмотрим сначала второй компонент описания.

3.1. Оценка длины описания невязок

Предположим сначала, что параметры $\mathbf{a}_k$ модели пространственного преобразования известны. Тогда значения невязок $\{(e_{x,i}, e_{y,i})\}_{i=1}^N$ также могут быть вычислены. Удобно считать векторы $(e_{x,i}, e_{y,i})$ независимыми и одинаково распределенными, т. е. считать эти значения отсчетами некоторого случайного вектора (e_x, e_y) . Тогда длина описания невязок может быть выражена через энтропию этого случайного вектора, умноженную на N . Для получения вычислительно эффективных процедур обычно предполагается, что невязки распределены по нормальному закону, тогда энтропия оказывается [7, с. 129] пропорциональной логарифму среднеквадратичного отклонения, а длина описания невязок может быть оценена как

$$L_e = N \left(\log_2 \sigma(e_x) + \log_2 \sigma(e_y) + 2c \right), \quad (2)$$

где $\sigma(e) = \frac{1}{N} \left(\sum_{i=1}^N e_i^2 \right)^{0,5}$ – среднеквадратичное значение невязок, а c – некоторая константа (не зависящая от модели преобразования и потому не участвующая

при сравнении моделей). Отметим, что в более общем случае [7, с. 130] вместо суммы логарифмов величин $\sigma(e_x)$ и $\sigma(e_y)$ в формуле (2) нужно использовать логарифм определителя ковариационной матрицы пары случайных величин (e_x, e_y) , однако для практических нужд формула (2) обычно бывает приемлемой.

В действительности, при вычислении энтропии непрерывной случайной величины необходимо учитывать опорную плотность (которая в простейшем случае отвечает точности задания значений этой величины), поэтому оценка (2) не вполне корректна [7, с. 30–33]. Это особенно хорошо видно при $\sigma \rightarrow 0$ (тогда $L_e \rightarrow -\infty$, что, конечно, лишено смысла). Для нужд совмещения изображений эту трудность легко обойти, если считать точность задания координат опорных точек известной (например, 1 пиксел) и полагать длину описания $L_e = 0$, когда невязки оказываются меньше этой точности (более элегантно эта трудность разрешается при использовании алгоритмической теории информации А.Н. Колмогорова, рассмотрение которой, однако, выходит за рамки статьи).

Если бы задача ставилась как поиск параметров пространственного преобразования, минимизирующих величину L_e , то мы бы получили классический метод наименьших квадратов (МНК). Однако согласно принципу МДО требуется также учесть длину описания параметров преобразования.

3.2. Оценка длины описания параметров преобразования

Каждый параметр преобразования должен описываться определенным числом бит информации. Как правило, это число считают одинаковым для всех параметров [8]. Здесь же мы будем считать, что на описание разных параметров преобразования может приходиться разное количество информации, причем это количество может зависеть от расположения опорных точек. К примеру, если все опорные точки расположены на одной прямой, то изменение масштаба в направлении, перпендикулярном этой прямой, оценено быть не может (или может, но с очень низкой точностью), т. е. в исходных данных просто нет информации для оценки этого параметра. Можно было бы закодировать этот параметр преобразования большим числом бит, но эта информация была бы нерелевантна содержанию исходных данных, более того, она бы никак не влияла на величины невязок.

Мы применили следующую простую идею оценки числа бит, которое необходимо выделить на описание того или иного параметра. Сначала параметр оценивается с максимальной точностью, а затем он

округляется до некоторого двоичного разряда. Число разрядов уменьшается до тех пор, пока увеличение длины описания невязок (которое возникает из-за понижения точности описания параметра) компенсируется уменьшением длины описания самого параметра.

Итак, длина описания модели вычисляется как

$$L_m = \sum_{i=1}^d L(a_i), \quad (3)$$

где $L(a_i)$ – число бит, оптимальное для описания i -го параметра.

4. Алгоритм выбора модели

Итак, результирующее выражение для длины описания будет:

$$DL = L_e + L_m = N(\log_2 \sigma(e_x) + \log_2 \sigma(e_y) + 2c) + \sum_{i=1}^d L(a_i). \quad (4)$$

Задача заключается в выборе такой модели преобразования, при которой бы значение DL было бы минимально. Величины L_e и L_m можно минимизировать последовательно. При фиксированном классе моделей $L_m = \text{const}$, т. е. параметры оптимальной модели можно находить путем минимизации L_e , для чего вполне подходит МНК. Таким образом, алгоритм выбора моделей состоит из следующих шагов:

1. Взять один из допустимых классов преобразований.
2. С помощью МНК определить оптимальные параметры преобразования данного класса.
3. Оценить число бит, оптимальное для описания каждого из параметров преобразования.
4. Посчитать общую длину описания $DL = L_e + L_m$ и сравнить ее с длинами описания моделей ранее рассмотренных классов.

Заметим, что этим алгоритмом в дополнение к определению класса пространственного преобразования и оценке его параметров также определяется и точность, с которой эти параметры оценены. Это свойство может быть весьма полезным на практике.

5. Искключение выбросов

При решении задачи совмещения изображений часто оказывается, что некоторые опорные точки на двух изображениях отождествлены неверно. В результате такие пары точек искажают общее пространственное преобразование, снижая точность совмещения. Принцип МДО позволяет расширить критерий выбора модели преобразования с учетом выбросов.

Выброс не участвует в построении модели, поэтому описание выбросов следует отделить от опи-

сания всех остальных точек. Если координаты опорных точек описываются через невязки относительно общего преобразования, то координаты выбросов должны описываться непосредственно (или в более общем случае для описания выбросов должна использоваться отдельная модель).

Пусть все координаты описываются с некоторой точностью ε . Тогда для описания одной координаты x нужно $\log_2[D(x)/\varepsilon]$ бит, где $D(x)$ определяет диапазон, в котором меняется значение x , например, это может быть ширина изображения или СКО абсцисс опорных точек.

Для описания одной невязки требуется $\log_2[\sigma(e)/\varepsilon]$ бит ($-\log_2 \varepsilon = c$, см. уравнение (2)). При исключении одной точки из общей модели невязкам будет соответствовать новая случайная величина e' с другим значением $\sigma(e')$, которое вычисляется по всем точкам, кроме исключенной. Теперь для хранения невязок нужно $(N-1)\log_2[\sigma(e')/\varepsilon]$ бит, а для отдельного хранения координаты исключенной точки – $\log_2[D(x)/\varepsilon]$ бит. Таким образом, критерий МДО для исключения выброса в предположении о нормальном законе распределения невязок будет

$$(N-1)\log_2[D(e')/\varepsilon] + \log_2[D(x)/\varepsilon] > N\log_2[D(e)/\varepsilon]. \quad (5)$$

Априорная точность ε в левой и правой частях формулы (5), к счастью, сокращается. С учетом необходимости описания двух координат получаем критерий исключения выброса:

$$(N-1)[\log_2[\sigma(e'_x)] + \log_2[\sigma(e'_y)]] + \log_2[D(x)] + \log_2[D(y)] > N[\log_2[\sigma(e_x)] + \log_2[\sigma(e_y)]] \quad (6)$$

Этот критерий интересен по нескольким причинам. Он включает среднюю ошибку до и после исключения точки (обычно в таком критерии исключения выбросов, как K -сигма, используется одна из этих двух ошибок). В этот критерий в явном виде входит число точек (в критерии K -сигма рекомендуется брать разные значения K в зависимости от числа точек). И, наконец, в этом критерии отсутствуют эвристически задаваемые параметры (правда, длину описания отдельной точки корректно оценить достаточно сложно). К сожалению, более подробное обсуждение этого критерия выходит за рамки данной работы.

6. Экспериментальная проверка

Для проверки описанного выше метода выбора оптимальной модели пространственного преобразования были осуществлены две группы тестов: по синтезированным наборам опорных точек и по реальным изображениям.

В первом случае задавались различные пространственные преобразования, по которым далее проводилась генерация наборов опорных точек: точки первого изображения выбирались случайно, а точки второго изображения строились из них с использованием заданного преобразования с добавлением случайных ошибок. Поскольку истинное преобразование в каждом случае было известно, можно было оценить среднюю (по всему изображению) точность найденного по опорным точкам преобразования. Таким образом, синтезированные тесты давали возможность объективного сравнения выбранных тем или иным способом моделей преобразования.

Рассмотрим следующий пример. Пусть задано пространственное преобразование

$$\begin{cases} x'_i = 10 + 1,02608x_i - 0,08977y_i, \\ y'_i = -13 + 0,08977x_i + 1,02608y_i. \end{cases} \quad (7)$$

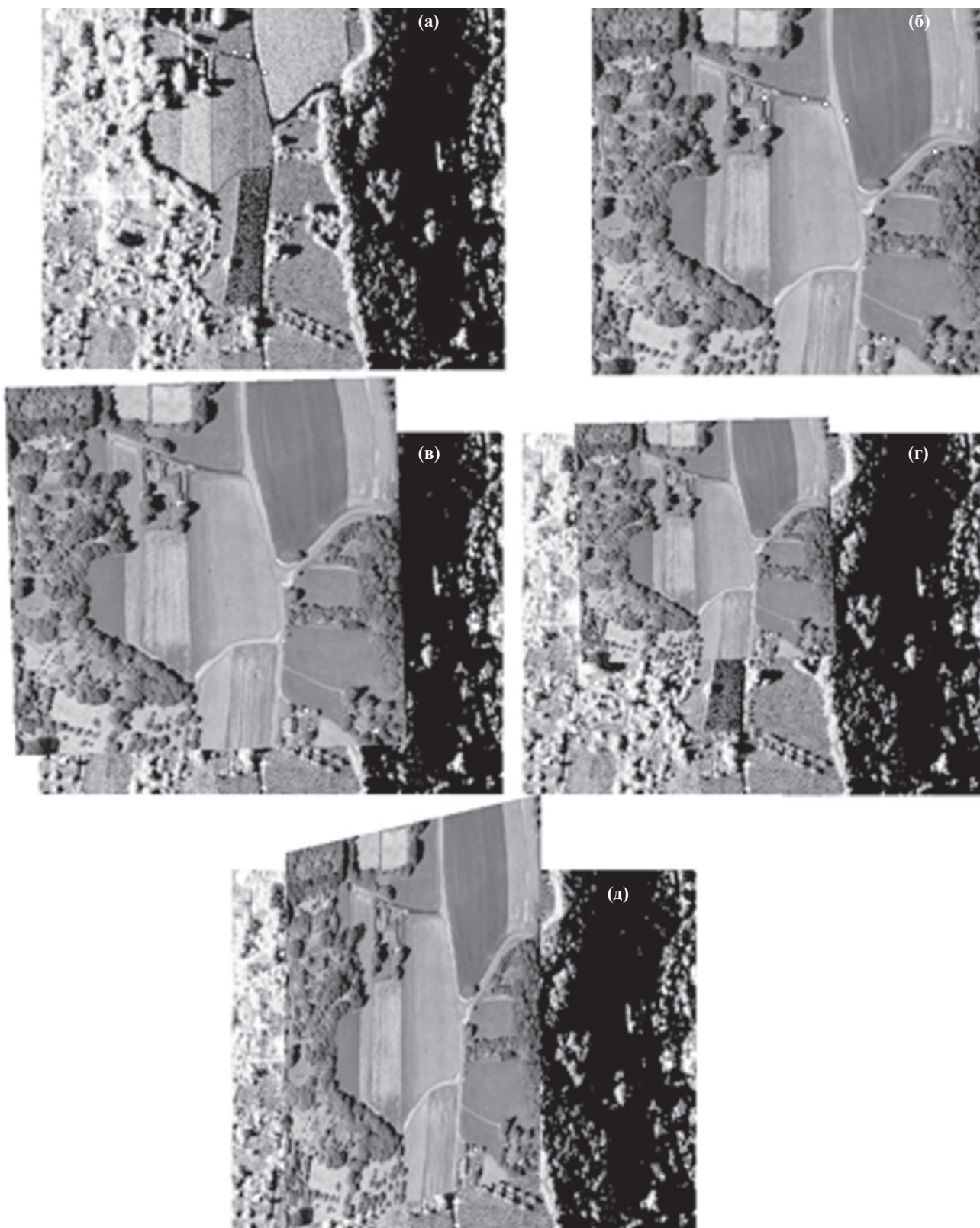
Это преобразование подобия с масштабом 1,03 и углом поворота 5° . И пусть задан набор из 6 пар опорных точек. Координаты точек на первом изображении выбираются случайно, а координаты соответствующих им точек второго изображения вычисляются по указанному преобразованию с добавлением ошибки, равномерно распределенной в диапазоне $[-5, 5]$ пикселей.

В табл. 1 приведены значения средней ошибки E_{emp} , вычисленной для заданных опорных точек, средней истинной ошибки по всему изображению E_{real} и длины описания DL для разных классов преобразований, параметры которых получены с помощью МНК.

Как видно из табл. 1, ошибка совмещения, измеренная по опорным точкам, уменьшается при увеличении числа параметров преобразования, однако реальная ошибка при этом может начать возрастать, когда оценивается преобразование, более сложное, чем преобразование, которое было использовано для порождения точек. Также видно, что хотя длина описания не строго соотносится с реальной ошибкой, их минимумы совпадают. При тестировании на многих примерах было установлено, что критерий длины описания в большинстве случаев позволяет выбрать тот класс преобразований, при котором истинная ошибка оказывается минимальной.

Таблица 1. Зависимость длины описания от модели преобразования

Преобразование	Число параметров	E_{emp} , пиксел	E_{real} , пиксел	DL , бит
Сдвига	2	19,3	22,6	54,0
Движения	3	6,1	6,5	48,6
Подобия	4	3,4	1,4	47,7
Аффинное	6	1,8	3,9	59,8
Билинейное	8	1,5	4,8	75,8
Квадратичное	12	0,0	25,3	108,0



Результаты совмещения радиолокационного (а) и оптического (б) изображений по указанным на них опорным точкам посредством преобразования движения (в), подобия (г) и аффинного преобразования (д). Среднее значение невязок для опорных точек составляет: в – 20,8, г – 2,9, д – 2,6 пикселей; значение DL : в – 58,0, г – 46,8, д – 63,8 бит.

Таблица 2. Влияние числе бит, выделенного для записи угла поворота, на длину описания

$L(a_i = \text{угол}),$ бит	$E_{\text{emp}},$ пиксел	$E_{\text{real}},$ пиксел	$DL,$ бит
5	15,8	10,9	68,4
6	6,4	4,5	54,6
7	5,2	2,4	53,1
8	3,4	1,4	47,7
9	3,1	1,7	48,3
10	3,0	2,0	49,7

Отметим, однако, что не всегда класс выбранного преобразования соответствует преобразованию, использованному для порождения тестовых данных, что особенно характерно для случая малых выборок или больших шумов. Это и понятно, поскольку такие выборки могут просто не содержать информации о некоторых параметрах преобразования. Важно, что на основе принципа МДО выбирается не “истинное”, а наилучшее (по своим экстраполяционным свойствам) преобразование, которое может быть построено по имеющимся данным.

В табл. 2 приведены значения оцененной и истинной ошибок совмещения, а также длина описания для разного числа бит, отводимого на параметр угла поворота для преобразования подобия, построенного для предыдущего примера. Как видно, оптимальной (для использованного нами набора точек) оказывается точность в 2^{-8} рад, что также соответствует минимуму истинной ошибки. Для нашего набора угол поворота был найден как $5,2^\circ \pm 0,22^\circ$.

Естественно, достижение минимума истинной ошибки при минимуме длины описания не гарантировано, а носит вероятностный характер (вероятность выбрать “плохое” решение в рамках информационного подхода падает экспоненциально при увеличении количества информации). Оцененная точность определения параметров также может содержать некоторую ошибку. Удивительно то, что даже округление найденных параметров до значащих разрядов может увеличивать результирующую точность совмещения. Однако тестирование показало, что при негауссовом распределении внесенных в координаты точек ошибок не следует доверять полученной точности определения параметров, хотя формула (4) все еще позволяет выбрать класс пространственного преобразования подходящей сложности.

Тестирование также показало полезность выбора класса преобразования при совмещении реальных изображений. На рисунке представлен пример, когда такой выбор позволяет ощутимо снизить ошибки совмещения. При увеличении числа опор-

ных точек выигрыш от выбора оптимального класса пространственного преобразования уменьшается, но продолжает быть достоверным.

Заключение

Принцип минимальной длины описания позволяет осуществить выбор между пространственными преобразованиями, содержащими разное число параметров. При использовании таких критериев, как среднеквадратичное отклонение, возникает тенденция выбирать как можно более сложное пространственное преобразование, которое, однако, не всегда позволяет достигнуть максимально возможной точности.

Предложенный нами информационный критерий позволяет избежать этой проблемы, что ведет к повышению точности совмещения. Это особенно актуально при структурном совмещении изображений, подверженных различного рода изменчивости (например, сезонно-суточной), в результате которой на этих изображениях отождествляется малое число опорных точек и их координаты могут быть определены с большими ошибками.

В дальнейшем этот подход может быть распространен на пространственные преобразования непараметрического типа, что актуально для задач совмещения биомедицинских изображений и изображений трехмерных сцен.

ЛИТЕРАТУРА

1. *Vitanyi P.M.B., Li M.* Minimum description length induction, Bayesianism, and Kolmogorov complexity // *IEEE Transactions on Information Theory*. 2000. V. 46. № 2. P. 446–464.
2. *Brown L.G.* A survey of Image Registration Techniques // *ACM Computing surveys*. 1992. V. 24. P. 325–376.
3. *Dowman I.J.* Automating image registration and absolute orientation: solutions and problems // *Photogrammetric Record*. 1998. V. 91. № 16. P. 5–18.
4. *Потапов А.С., Малышев И.А., Луцив В.Р.* Совмещение аэрокосмических изображений с субпиксельной точностью методом локальной корреляции // *Оптический журнал*. 2004. Т. 71. № 5. С. 31–36.
5. *Potapov A.S., Lutsiv V.R.* Information-theoretic approach to image description and interpretation // *Proc. SPIE*. 2003. V. 5400. P. 277–283.
6. *Potapov A.S.* Image matching with the use of the minimum description length approach // *Proc. SPIE*. 2004. V. 5426. P. 164–175.
7. *Стратонович Р.Л.* Теория информации. М.: Сов. радио, 1975. 424 с.
8. *Hansen M., Yu B.* Model Selection and the Principle of Minimum Description Length // *Technical Memorandum, Bell Labs, Murray Hill. N.J.* 1998. 13 p.