

УДК 528.27

ИССЛЕДОВАНИЕ ТОЧНОСТИ ТРАНСФОРМИРОВАНИЯ РАСТРОВЫХ ИЗОБРАЖЕНИЙ

канд. техн. наук, доц. И.Е. РАК
(Белорусский национальный технический университет, Минск)

Рассматриваются вопросы, с которыми сталкиваются специалисты, занимающиеся созданием ГИС города инженерного назначения на основе архивного картографического материала. Исследованы методы преобразования и даны рекомендации по учету возможных ошибок исходных материалов, возникающих под влиянием нелинейных деформаций основы, на которой отпечатан исходный материал.

Необходимость проведения градостроительных и учетно-кадастровых работ, мониторинга земельных ресурсов, потребность власти на местах иметь возможность оперативного принятия управленческих решений определили развитие геоинформационных систем (ГИС) инженерного назначения.

Сегодня идея создания городской ГИС инженерного назначения, возникшая в начале развития компьютерных технологий, до сих пор практически не реализована ни в одном крупном городе СНГ. Главная причина этого (например, в Республике Беларусь) заключается не в отсутствии финансирования и не в отсутствии специалистов, готовых заниматься этой проблемой, а в том, что нет системного подхода к решению глобальной задачи по сбору и использованию уже накопленной информации, нет единого информационного пространства для обмена информацией между отдельными ведомствами [5].

И первым шагом к решению данной проблемы должны быть законодательные акты, закрепляющие концепцию создания и развития электронного векторного топографического плана города.

Кроме того, для создания ГИС инженерного назначения необходима современная, точная цифровая топографическая и картографическая информация. Но из-за того, что в городах имеется несколько местных систем координат, сети неравноточные, не согласованы друг с другом, совместно не уравниваются, формировать единое равноточное пространство территории города невозможно. Поэтому вопрос о создании единой системы координат города, которая бы учитывала особенности уже существующих систем и обеспечивала бы свободный доступ исполнителям, не противореча в то же время действующему законодательству о сведениях, составляющих государственную тайну, потребует согласованных решений [6].

Еще одна проблема – архивный картографический материал. Архивный материал – это планшеты, созданные 20–30 лет назад и, соответственно, имеющие на сегодняшний день устаревшую информацию. Поэтому назрела необходимость в разработке технических приемов, которые позволят исследовать содержание архивного картографического материала и привести его к современному состоянию. Наиболее эффективным методом перевода информации с бумажного носителя в цифровой является сканирование и трансформирование планшетов. В то же время для этого нужны четкие рекомендации по учету возможных ошибок исходных материалов, возникающих под влиянием нелинейных деформаций основы, на которой отпечатан исходный материал. Но отсканировать планшет – это еще не значит получить топографо-геодезическую основу данной территории.

Технология исследования точности трансформирования была опробована на двух издательских оригиналах на пластике (В-1 и Б-4 в системе координат 1963 года).

Сканирование было выполнено с геометрическим разрешением 85,3 мкм (300 dpi). Трансформирование выполнялось с помощью программных средств ГИС ArcGis и реализованных в ней методов аффинного преобразования и полиномиальных преобразований второй и третьей степени.

В качестве исходных и контрольных точек были использованы углы рамки трапеции, точки ее пересечения с линиями километровой сетки и узлы километровой сетки (рис. 1, 2).

Минимальное число контрольных точек (n) устанавливалось исходя из известного соотношения между точностью вычисления ожидаемой средней квадратической ошибки m_m и самой средней квадратической ошибкой (m):

$$\frac{m_m}{m} = \frac{1}{\sqrt{2n}}. \quad (1)$$

Потребовав, чтобы отношение m_m/m не превышало 1/7...1/5 (15...20 %), получим, что число контрольных точек n должно быть не менее 12.

Определение параметров преобразования одной координатной системы в другую выполняется методом наименьших квадратов, эффективность применения которого (как и надежность оценки точности уравнивательных вычислений) определяется наличием систематических ошибок. В связи с этим была выполнена статистическая обработка сформированных рядов остаточных расхождений координат то-

чек (Δ_x , Δ_y) для проверки случайного характера этих расхождений и гипотезы о нормальности распределения погрешностей соответствующего ряда с помощью известных критериев при 5 %-ном уровне значимости (то есть при доверительной вероятности 95 %).

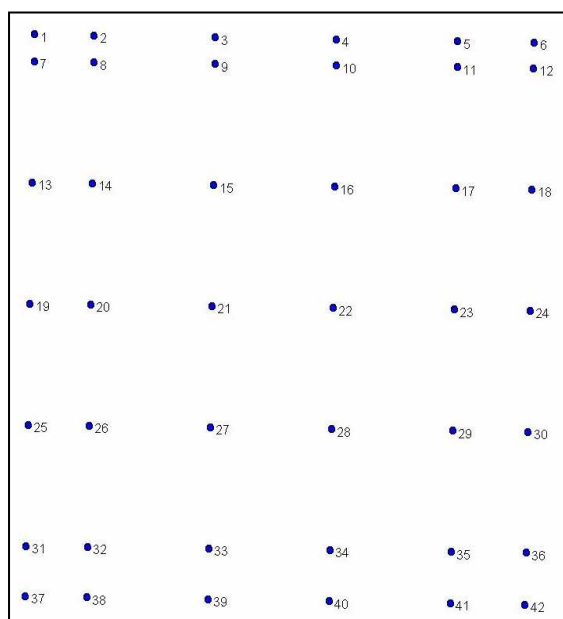


Рис. 1. Схема размещения опорных и контрольных точек на трапеции В-1

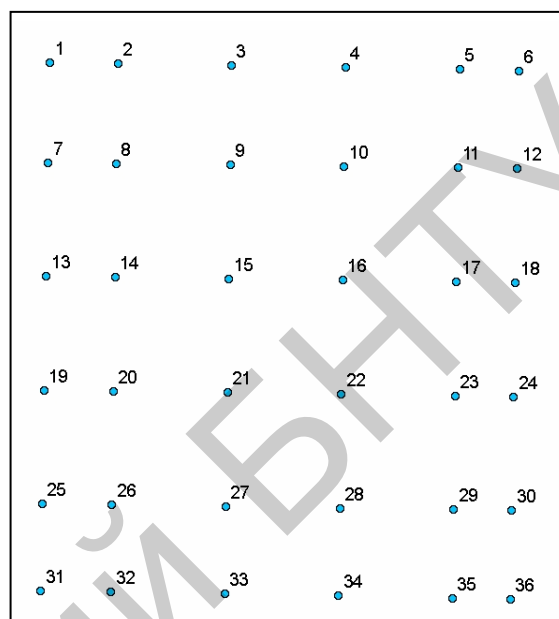


Рис. 2. Схема размещения опорных и контрольных точек на трапеции Б-4

Таким образом, всего на трапеции В-1 и Б-4 было выбрано 42 и 36 опорных точек соответственно, что позволило в последующем выполнить определение параметров трансформирования растровых изображений несколькими способами с различным числом исходных и контрольных точек (табл. 1).

Таблица 1

Исходные и контрольные точки

Способ трансформирования	Число точек на листе					
	В-1			Б-4		
	вариант	исходный	контрольный	вариант	исходный	контрольный
Аффинный	1	4	38	10	4	32
	2	8	34	11	8	28
	3	12	30	12	12	24
	4	22	20	13	20	16
Полиномиальный 2-й степени	5	8	34	14	8	28
	6	12	30	15	12	24
	7	22	20	16	20	16
Полиномиальный 3-й степени	8	12	30	17	12	24
	9	22	20	18	20	16

Первичный анализ результатов измерений

Для оценки точности преобразования дискретных точек исходных оригиналов по расхождениям эталонных и вычисленных координат контрольных точек, в каждом эксперименте были вычислены средние (v) и средние квадратические ошибки (m) по координатным осям и в плане:

$$\left. \begin{aligned} m_x &= \sqrt{\frac{\sum \Delta_x^2}{n}}, \quad m_y = \sqrt{\frac{\sum \Delta_y^2}{n}}, \quad m_s = \sqrt{\frac{\sum (\Delta_x^2 + \Delta_y^2)}{n}}; \\ v_x &= \frac{\sum |\Delta_x|}{n}, \quad v_y = \frac{\sum |\Delta_y|}{n}, \quad v_s = \frac{\sum \sqrt{(\Delta_x^2 + \Delta_y^2)}}{n}. \end{aligned} \right\} \quad (2)$$

После этого все точки, на которых остаточные невязки превышали утроенное значение установленной Инструкцией [3] средней погрешности положения точек плановой съемочной сети относительно пунктов государственной геодезической сети (0,3 мм или 3,0 м), были исключены как ошибочные. Уточненные значения средних и средних квадратических погрешностей преобразования были вычислены по формулам (2) повторно. Всего по этому основанию исключена 21 точка (5 % от общего их числа).

Анализ полученных данных позволяет сделать следующие выводы:

- 1) установленные по контрольным точкам погрешности преобразования во всех вариантах близки между собой (кроме вариантов 10 и 11, где эти погрешности оказались несколько больше), а различия между ними не превышают 15...30 % от их величины, что соответствует расчетным значениям надежности оценки;
- 2) исключение из обработки контрольных точек с остаточными расхождениями, превышающими 0,3 мм в масштабе плана, улучшило средние и средние квадратические погрешности на 15...20 %;
- 3) во всех случаях средние квадратические ошибки планового положения на 15...20 % больше средних, что подчеркивает незначительное влияние ошибок систематического характера и свидетельствует в пользу нормального распределения;
- 4) все измерения выполнены с надлежащей точностью, заметного влияния систематических ошибок не обнаружено. Результаты обработки данных могут быть использованы для выработки обоснованных рекомендаций;
- 5) при использовании полиномов второй и тем более третьей степени и размещении опорных точек по периметру изображения значения остаточных расхождений координат точек существенно уменьшаются;
- 6) при обработке трапеции **В-1** средние и средние квадратические ошибки контрольных точек по оси ординат оказались на 10...25 % больше, чем ошибки абсцисс. Для трапеции **Б-4** ситуация обратная: ошибки абсцисс превышают ошибки ординат.

Так как в соответствии с требованиями Инструкции [3] средняя погрешность положения точек плановой съемочной сети относительно опорной геодезической сети не должна превышать 0,1 мм, то привязка растровой копии должна быть выполнена так, чтобы был обеспечен указанный допуск. Следовательно, погрешность положения точки привязанного к требуемой координатной системе и трансформированного по параметрам преобразования цифрового оригинала должна быть также 0,1 мм, а предельная – 0,2 мм. Применительно к масштабу 1:10000 этот допуск составит 1 и 2 м соответственно.

Однако сканирование издательских оригиналов выполнено с геометрическим разрешением 0,085 мм, что увеличивает ошибку измерения растровой копии и делает ее сопоставимой с ошибкой исходного оригинала (0,1 мм). Поэтому суммарную ошибку положения точки на исходном оригинале и ошибку измерения ее координат по растру можно полагать равной

$$M = \sqrt{(0,100)^2 + (0,085)^2} \approx 0,131 \text{ мм.}$$

С учетом этого средняя ошибка преобразования составит 1,3 м, а предельная – 2,6 м. В этот допуск укладываются все выполненные измерения, что свидетельствует об их достаточно высоком качестве. Причем, как и следовало ожидать, применение полиномиальных преобразований уменьшает эту погрешность на 20 % и более.

Следовательно, для приведения в соответствие точности исходного и цифрового оригиналов геометрическое разрешение растровой копии нужно уменьшить до пренебрегаемо малой величины, и оно должно составлять около 40 % от средней ошибки положения пунктов плановой сети, то есть иметь разрешение 500...600 dpi.

Статистические исследования остаточных расхождений координат контрольных точек

Статистические характеристики результатов экспериментальных работ были выполнены для выявления случайности погрешностей измерений, оценки достоверности найденных критериев, обоснованности используемых или предлагаемых технологических допусков и правомерности применения метода наименьших квадратов для определения вероятнейших значений неизвестных, которыми в контексте выполненных исследований являются параметры трансформирования исходной картографической основы – растровых копий соответствующих материалов.

Технология таких исследований достаточно хорошо отработана, описана в технической литературе и сводится к проверке с помощью соответствующих критериев гипотезы о нормальном распределении исследуемых погрешностей. Такая оценка выполняется по результатам сравнения параметров кривой эмпирического (реального) и теоретического распределения с помощью тех или иных критериев.

Проверка гипотезы о нормальном распределении рядов ошибок с помощью мощных критериев Пирсона, Колмогорова, Ястремского и целого ряда иных выполняется только для оценки достаточно больших рядов, с числом ошибок несколько десятков или сотен [1]. В условиях эксперимента это требо-

вание не выполнялось, поэтому было принято решение ограничить исследования оценкой основных параметров кривой распределения и сравнением их с теоретическими значениями.

К основным параметрам, определяющим форму кривой распределения погрешностей, относятся: математическое ожидание (M); начальные моменты первого – четвертого порядков ($v_1 \dots v_4$); центральные моменты второго – четвертого порядков ($\mu_2 \dots \mu_4$); асимметрия (S); эксцесс (E); дисперсия (σ); среднее (v) и среднее квадратическое (m) значения; отношение (z) средней погрешности к средней квадратической, определяемые по приведенным ниже формулам [1], а также иные параметры (например, параметры «спрямленной» диаграммы).

$$M = \sum \Delta_i; \quad (3)$$

$$\left. \begin{aligned} v_1 &= \frac{1}{n} \sum \Delta_i, & v_3 &= \frac{1}{n} \sum \Delta_i^3, \\ v_2 &= \frac{1}{n} \sum \Delta_i^2, & v_4 &= \frac{1}{n} \sum \Delta_i^4; \end{aligned} \right\} \quad (4)$$

$$\left. \begin{aligned} \mu_2 &= \sigma^2 = v_2 - v_1^2, \\ \mu_3 &= v_3 - 3v_2v_1 + 2v_1^3, \\ \mu_4 &= v_4 - 4v_3v_1 + 6v_2v_1^2 - 3v_1^4; \end{aligned} \right\} \quad (5)$$

$$\left. \begin{aligned} S_k &= \frac{\mu_3}{\sigma^2} = \frac{\mu_3}{\mu_2}, \\ E_k &= \frac{\mu_4}{\sigma^4} - 3, \\ z_k &= \frac{v}{m}, \end{aligned} \right\} \quad (6)$$

где Δ_i – погрешность измерений (элемент исследуемого ряда); v , m – средняя и средняя квадратическая погрешности соответственно.

Тесноту размещения погрешностей относительно центра группирования (их «разброс») характеризуют математическое ожидание, дисперсия и средняя погрешность, форму кривой эмпирического распределения – эксцесс, а ее положение относительно центра группирования – асимметрия. Как известно, при нормальном распределении исследуемых погрешностей значения асимметрии и эксцесса равны нулю, а отношение средней погрешности к средней квадратической составляет 0,8.

Исследуемое (эмпирическое) распределение может отличаться от нормального с заданной степенью приближения к нему, поэтому перечисленные величины должны заключаться в некотором диапазоне допустимых значений, границы которого находятся по следующим формулам [7]:

$$\left. \begin{aligned} |S_k| &\leq 3\sigma_s, & \sigma_s &\approx \sqrt{\frac{6}{n+3}}, \\ |E_k| &\leq 3\sigma_m, & \sigma_m &\approx \sqrt{\frac{24(n-3)}{(n+3)(n+5)}}, \end{aligned} \right\} \quad (7)$$

где n – число погрешностей в исследуемой выборке; z_a , z_b – критические значения отношения z , выбираемые из [1, табл. 19] в зависимости от числа погрешностей и уровня значимости (при оценке асимметрии и эксцесса уровень значимости учитывается косвенно, с помощью коэффициента при стандарте σ , значение которого может колебаться от 2,0 до 3,0).

Заметим, что начальные моменты первого порядка характеризует величину систематической погрешности.

Для проверки гипотезы нормального распределения довольно часто используют так называемый способ «спрямленных» (выпрямленных) диаграмм (рис. 3), для чего исследуемые погрешности размещают в ряд по возрастанию их значений.

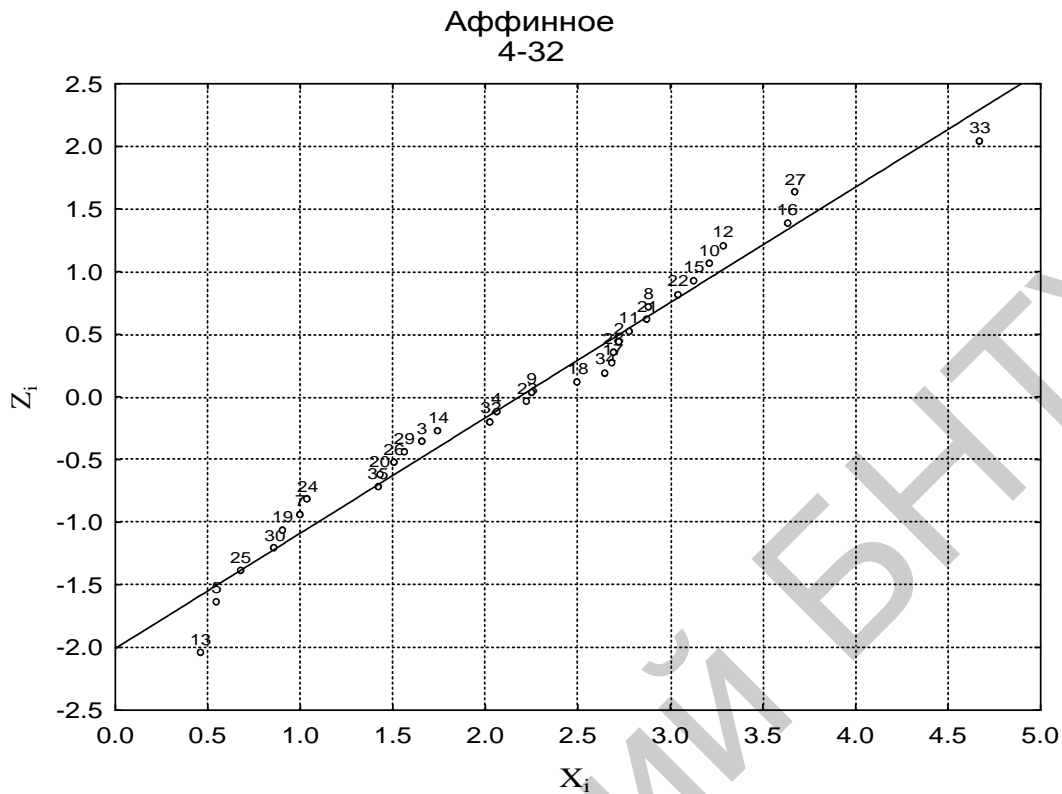


Рис. 3. Пример использования способа «спрямленных» диаграмм

Каждому элементу i этого ряда соответствует функция распределения, вычисляемая по формуле:

$$\bar{F}(z_i) = \frac{i-0,5}{n}, \quad (8)$$

по значению которой из выражения

$$F(\bar{z}_i) = \frac{2}{2\pi} \int_0^{\pi} e^{\frac{z^2}{2}} dz + 0,5 = \bar{F}(z_i) \quad (9)$$

по вспомогательной таблице в [1] находят величину z .

Гипотеза нормальности распределения считается непротиворечащей результатам измерений, если удаление точек от прямой, соответствующей желаемому (нормальному) распределению, не превышает вычисленного для принятого уровня значимости величины

$$|\bar{z} - z_i| \frac{t_i}{\phi(z_i)} \sqrt{\frac{[0,5 + \Phi(z_i)] \cdot (0,5 - \Phi(z_i))}{n}}, \quad (10)$$

где $\phi(z_i)$ – нормальная плотность распределения; t_i – критерий, выбираемый в зависимости от принятого уровня значимости (доверительной вероятности) и числа погрешностей; n – общее число членов ряда.

При выполнении расчетов были использованы рекомендации и вспомогательные таблицы, имеющиеся в работе [1, с. 99–100, 180, 182–183].

Результаты статистической обработки данных и их анализ

Статистический анализ полученных данных включал проверку гипотезы о нормальном распределении исследуемых погрешностей по следующим параметрам:

- по размаху выборки, определяемому величинами средних, средних квадратических погрешностей и их отношению;
- по форме кривой, определяемой значениями асимметрии и эксцесса, зависящими от величин погрешностей и их частот (с учетом знаков);
- по параметрам «спрямленных» диаграмм.

Параметры размаха выборки были вычислены без учета исключенных из обработки точек, что несколько изменило отношения z_k в сторону увеличения. При этом критические значения z_a, z_b найдены по таблице 19 в [1] для 5 %-ного уровня значимости в зависимости от числа исследуемых погрешностей.

Полученные данные показали, что по исследуемому параметру гипотеза о нормальном распределении не отвергается для рядов, полученных при использовании полиномов второй и третьей степени. Для рядов погрешностей, полученных при использовании аффинных преобразований, отношения z_k близки к верхнему пределу критических значений z_b , превышая их на 0,01...0,02 (соответствующие значения в таблице выделены). Одной из возможных причин этого является достаточно малое число измерений в рядах.

Параметры, определяющие форму кривой, вычислены по остаточным расхождениям эталонных и вычисленных координат (а не по невязкам планового положения, не имеющих знака).

Полученные данные позволили сделать следующие основные выводы:

- во всех случаях уклонения асимметрии и эксцесса от их теоретических значений не превышают утроенных значений стандартов соответствующих ошибок и должны быть признаны несущественными;
- значения математических ожиданий во всех случаях соответствуют систематическим погрешностям, составляющим несколько сантиметров, что никак не может признаваться существенным;
- во всех случаях значения дисперсии, вычисленной по формуле

$$\sigma_s = \sqrt{\sigma_x^2 + \sigma_y^2} \quad (11)$$

и соответствующей средней квадратической погрешности m_s , вычисленной по формуле (2), отличаются на величину, не превышающую стандарта среднего квадратического отклонения σ_σ .

Изложенное свидетельствует о том, что вычисленные значения асимметрии и эксцесса не противоречат нормальному распределению исследуемых погрешностей.

Способ спрямленных диаграмм, использованный для исследования результатов обработки данных двух трапеций, показывает, что ни в одном из исследованных рядов измерений гипотеза о нормальности исследуемых выборок не опровергается по малости уклонений точек от соответствующих прямых.

Подводя итог статистических исследований, следует отметить, что все использованные критерии проверки гипотезы о нормальном распределении, за исключением отношения средней погрешности к средней квадратической для вариантов аффинных преобразований, не отвергают ее при 5 %-ном уровне значимости. Что же касается аффинных преобразований, то выход эмпирических значений отношений z_k средних погрешностей к средним квадратическим не может служить серьезным основанием для опровержения проверяемой гипотезы как по малости величины отклонения, так и в связи с малым числом исследуемых погрешностей, число которых для уверенной оценки должно быть более 30 [1].

Изложенное дает основание полагать, что выполненные измерения не содержат заметных систематических погрешностей, а остаточные расхождения координат контрольных точек подчиняются закону нормального распределения, что позволяет использовать их для установления соответствующих допусков и выработки рекомендаций.

Выводы и предложения

Приведенные выше результаты исследования некоторых способов координатных преобразований с целью последующего трансформирования растрового изображения и подтверждение гипотезы о нормальном распределении остаточных расхождений координат контрольных точек позволяют сделать несколько выводов и рекомендаций, касающихся отработки и совершенствования технологии создания геоинформационных систем на растровой картографической основе:

- 1) при сканировании исходных материалов геометрическое разрешение выходного изображения не должно вносить дополнительных искажений, сохраняя точность положения объектов в соответствии с требованиями действующих нормативных документов. Это можно обеспечить в случае, если размер пикселя растрового оригинала будет не более 30...40 % от средней погрешности положения опорных точек на исходном оригинале. Применительно к составительскому оригиналу это соответствует геометрическому разрешению порядка 30...40 мкм или 500...600 dpi;
- 2) в качестве исходных и контрольных опорных точек желательно принимать углы рамки трапеции, точки пересечения линий координатной сетки и ее выходы на рамку;
- 3) при большом числе исходных опорных точек, равномерно размещенных по рамке трапеции, все исследованные способы координатных преобразований обеспечивают получение результатов одинаковой точности, не зависящей от положения контролируемой точки;
- 4) при использовании в качестве исходной основы составительских оригиналов на пластике допустимо использование аффинных преобразований по четырем точкам, совмещенным с углами трапеции. Контроль результатов такого преобразования можно выполнить по четырем – пяти точкам, расположенным по серединам северной, восточной, южной и западной рамок трапеции;

5) полиномы второй и третьей степени целесообразно использовать в случаях, когда есть основание предполагать возможность наличия неравномерной или случайной деформации исходного оригинала. Исправить ее влияние можно не менее чем по восьми точкам, размещенным по углам рамки и на ее серединах. В этом случае контрольные точки можно располагать в центре трапеции и вблизи ее рамки;

6) применение полиномиальных преобразований при размещении исходных с нарушением их симметрии относительно центра листа приводит к искажениям, что следует считать недопустимым;

7) точность координатных преобразований может быть установлена по четырем – пяти контрольным опорным точкам, размещенным в местах, наиболее удаленных от исходных точек. Надежность оценки точности по такому числу точек может составить 20...25 %;

8) для определения параметров растровых оригиналов, полученных с рекомендованным в пункте 1 геометрическим разрешением, величина остаточной невязки на исходных точках не должна превышать 0,1 мм, а предельная – 0,2 мм в масштабе плана. Средняя невязка, найденная по контрольным точкам, не должна превышать 0,12...0,17 мм в масштабе плана;

9) трансформирование растрового изображения должно выполняться возможно меньшими фрагментами, хотя это и требует дополнительных затрат времени.

ЛИТЕРАТУРА

1. Видуев, Н.Г. Вероятностно-статистический анализ погрешностей измерения / Н.Г. Видуев, Г.С. Кондра. – М.: Недра, 1969.
2. Инструкция по топографической съемке в масштабах 1:5000, 1:2000, 1:1000, 1:500. – М.: Недра, 1985.
3. Калинин, А.С. Технологии организации данных при ведении дежурных планов городов при использовании систем CREDO третьего поколения / А.С. Калинин // Автоматизированные технологии изысканий и проектирования: материалы конф. – М., 2007.
4. Пыко, Т.В. Современное состояние и перспективы развития топографо-геодезического картографического производства Республики Беларусь / Т.В. Пыко // Материалы Первого конгресса геодезистов и картографов. – Минск, 2002.
5. Рак, И.Е. Разработка и исследование методики топографо-геодезических работ для использования ГИС-технологий в средних и малых городах Беларуси / И.Е. Рак // Изв. вузов. Геодезия и аэрофото-съемка. – 2008. – № 3.
6. Федосеев, Ю.Е. Создание цифровых карт масштаба 1:2000 с применением устаревших архивных данных / Ю.Е. Федосеев, Н.М. Никитина // Геопрофи. – 2004. – № 3, 4.
7. Федосеев, Ю.Е. Некоторые проблемы и технологии создания топоосновы для комплексной модели населенного пункта / Ю.Е. Федосеев, И.Е. Рак // Изв. вузов. Геодезия и аэрофото-съемка. – 2007. – № 3.

Поступила 13.05.2015

INVESTIGATING THE ACCURACY OF TRANSFORMING OF THE RASTER IMAGES

I. RAK

The article discusses the issues faced by professionals involved into creation of GIS city for engineering purposes on the basis of archival cartographic material. Also here explored methods of the conversion and given recommendations for registration possible mistakes arising under the influence of nonlinear deformation of the base which is used for printing the source material.