

ПОСТРОЕНИЕ ПРОЦЕДУР ПРЕОБРАЗОВАНИЯ ДАННЫХ В СХЕМЕ ПОСЛЕДОВАТЕЛЬНОГО ОЦЕНИВАНИЯ МОДЕЛЕЙ ИСКАЖЕНИЙ ПО МАЛЫМ ФРАГМЕНТАМ ИЗОБРАЖЕНИЙ

В.А. Фурсов, А.В. Устинов

Институт систем обработки изображений РАН

Введение

Для решения задач оценивания (идентификации) характеристик во много их отношениях удобными являются линейные модели с постоянными (по пространственным координатам) параметрами. Если искажения нелинейны и (или) пространственно - зависимы, применяют кусочно-постоянную аппроксимацию множеством инвариантных к сдвигу линейных моделей на малых фрагментах изображений.

Простейшей линейной моделью с постоянными параметрами является фильтр с конечной импульсной характеристикой (КИХ - фильтр). Связь отсчетов входного (неискаженного) - $x(n_1, n_2)$ и выходного (искаженного) - $y(n_1, n_2)$ изображений для этой модели описывается соотношением [1]

$$y(n_1, n_2) = \sum_{\{k_1, k_2\} \in D} h(k_1, k_2) x(n_1 - k_1, n_2 - k_2) + \xi(n_1, n_2), \quad (1)$$

где $h(k_1, k_2)$, $x(n_1 - k_1, n_2 - k_2)$ - значения импульсной характеристики и входных отсчетов в точках опорной области D , а $\xi(n_1, n_2)$ - дискретная последовательность, включающая погрешности измерений и ошибки аппроксимации.

Из N отсчетов $y(n_1, n_2)$, принадлежащих выbranному на выходном изображении фрагменту, составим $N \times 1$ - вектор \mathbf{y} . Пусть соответствующая каждой компоненте вектора \mathbf{y} опорная область на входном изображении содержит M отсчетов. Составленные из этих отсчетов N векторов размерности $M \times 1$ обозначим \mathbf{x}_i , $i=1, N$, а соответствующий им $M \times 1$ -вектор параметров импульсной характеристики - \mathbf{h} . Тогда можно записать матричное соотношение

$$\mathbf{y} = \mathbf{X}\mathbf{h} + \xi \quad (2)$$

где $N \times M$ - матрица $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$, а $N \times 1$ - вектор ξ составлен из фигурирующих в (1) ошибок $\xi(n_1, n_2)$.

Известна [2] оценка по методу наименьших квадратов (МНК-оценка) $\hat{\mathbf{h}}$ вектора параметров \mathbf{h} модели (2):

$$\hat{\mathbf{h}} = [\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T \mathbf{y}. \quad (3)$$

Применение указанной оценки в данном случае связано с принципиальными трудностями. Если размеры выбранного фрагмента невелики, свой-

ство устойчивости статистических характеристик ошибок, на соответствующей ему малой выборке, не проявляется в полной мере. Поэтому даже если на изображении в целом ошибки аппроксимируются гауссовой плотностью, МНК-оценки на малом фрагменте не могут считаться вполне обоснованными.

Для преодоления указанных трудностей часто идут по пути построения робастных, например, оптимальных оценок для "наихудшего" в заданном классе распределения ошибок. Если априорная информация бедная, класс распределений, по-видимому, должен задаваться достаточно широким, что неизбежно приведет к неполному использованию информации, содержащейся в конкретном наборе данных.

На наш взгляд более плодотворным является адаптивный подход, заключающийся в уточнении модели ошибок по мере появления дополнительной информации, содержащейся в невязках. Настоящая работа посвящена исследованию круга вопросов, связанного с построением такой схемы обработки информации.

1. Схема последовательного МНК-оценивания

Процедура оценивания строится в виде попеременно реализуемых этапов вычисления МНК-оценок (3) и преобразования исходных данных. В принципе, для улучшения качества оценок можно применять по крайней мере два типа преобразования: взвешивание данных

$$\tilde{\mathbf{X}} = \mathbf{G}\mathbf{X}, \quad \tilde{\mathbf{y}} = \mathbf{G}\mathbf{y}, \quad (4)$$

$$\mathbf{G} = \text{diag}(g_1, g_2, \dots, g_N)$$

и непосредственная корректировка вектора \mathbf{y} . В настоящей работе исследуется лишь преобразование типа взвешивания. Матрица $\mathbf{G} = \mathbf{G}(\hat{\xi})$ этого преобразования после каждого этапа МНК-оценивания строится по вектору невязок:

$$\hat{\xi} = \mathbf{y} - \mathbf{X}\hat{\mathbf{h}}. \quad (5)$$

Для получения начальной оценки при отсутствии априорной информации она задается единичной.

По схеме последовательного МНК- оценивания можно вычислять оценки максимального правдоподобия [3] $\hat{\mathbf{h}}$ и \hat{s} :

$$Q(\xi(\hat{\mathbf{h}}), \hat{s}) = \min_{\mathbf{h}, s} \sum_{i=1}^N F(\xi_i(\mathbf{h}), s),$$

если функция потерь

$$F(\xi(\mathbf{h}), s) = -\ln p(\xi(\mathbf{h}), s) \quad (7)$$

и соответствующие ей весовые коэффициенты

$$g(\xi, s) = (F_{\xi}^{\circ}(\xi(\mathbf{h}), s)\xi^{-1})^{1/2} \quad (8)$$

удовлетворяют условиям:

$$\begin{aligned} F_{\xi}^{\circ}(\xi(\mathbf{h}), s)\xi^{-1} > 0, \\ \lim_{\xi \rightarrow \infty} g(\xi, s) = 0, \lim_{\xi \rightarrow 0} g(\xi, s) < \infty. \end{aligned} \quad (9)$$

Необходимые условия минимума в задаче (6) выписываются в виде $(M+1)$ уравнений:

$$\begin{aligned} \left[\sum_{i=1}^N g^2(\xi_i(\hat{\mathbf{h}}), \hat{s}) x_i x_i^T \right] \hat{\mathbf{h}} - \\ - \left[\sum_{i=1}^N g^2(\xi_i(\hat{\mathbf{h}}), \hat{s}) x_i y_i \right] = 0 \end{aligned} \quad (10)$$

$$\sum_{i=1}^N F_{\xi}^{\circ}(\xi_i(\hat{\mathbf{h}}), \hat{s}) = 0. \quad (11)$$

Нетрудно заметить, что попеременное решение уравнений (10) и (11) по существу и является схемой последовательных МНК-оценок.

Приведенные выше соотношения показывают возможность однозначного построения весовой матрицы преобразования \mathbf{G} , если задана априорная гипотеза в виде плотности распределения. В связи с этим возникает следующий вопрос: возможна ли оценка параметрического семейства распределений ошибок, по полученному на промежуточном этапе вектору невязок (5).

Ответ на поставленный вопрос дает качественная теория идентификации [4]. Можно показать, что

$$\mathbf{T}_0^T \xi = \mathbf{T}_0^T \hat{\xi}, \quad (12)$$

где \mathbf{T}_0 - матрица размерности $N \times (N-M)$, составленная из $N-M$ собственных векторов, соответствующих нулевым собственным значениям матрицы $\mathbf{X}_{M-1} \mathbf{X}_{M-1}^T$, т.е. проекции вектора ошибок ξ и вектора невязок $\hat{\xi}$ на выбранный базис нуль-пространства матрицы \mathbf{X} совпадают. В частности, если вектор ошибок принадлежит нуль-пространству, он совпадает с вектором невязок. Ясно, что в некоторой окрестности существует множество (не нулевой меры) векторов ошибок, достаточно близких к соответствующим векторам невязок. Признаком принадлежности вектора ошибок к этому множеству может служить сравнительно большое значение нормы вектора невязок. Если это так, то остается решить вопрос с выбором множества параметрических семейств распределений и меры различения гипотез.

2. Анализ распределений ошибок

В таблице 1 приведены некоторые из известных симметричных относительно центра ξ_0 параметрических семейств распределений, удовлетворяющих условиям (9). Для упрощения специально отобраны плотности, зависящие лишь от одного (при $\xi_0=0$) параметра масштаба $-s$. Во втором столбце приведены выражения для плотностей распределений, в третьем - соответствующие этим плотностям функции потерь (7), а в четвертом - выражения для квадратов весовых коэффициентов (8). Распределения 4 и 5 получаются одно из другого заменой s на $2s$. Поэтому при проведении экспериментов далее используется только семейство 4.

В качестве меры для различения гипотез о принадлежности эмпирической функции распределения $F_{эмп}(\xi)$ к одному из указанных однопараметрических семейств $F(s, \xi)$ примем среднеквадратичный критерий вида:

$$J_{\mathcal{L}}(s) = \left\{ \sum_{i=1}^{N-1} [F(s, \eta_i) - p_i]^2 * (\xi_{i+1} - \xi_i) \right\}^{1/2} \quad (13)$$

Здесь ξ_i - отсчеты вариационного ряда, построенного из компонент вектора ошибок ξ , $\eta_i \in [\xi_i; \xi_{i+1}]$, $p_i, i=1, N$ - отсчеты эмпирической функции распределения.

Выбор критерия (13) оправдывается тем, что "весовые" множители $\xi_{i+1} - \xi_i$ по смыслу близки к

множителям $\frac{n}{p_i}$ при вычислении статистики

χ^2 (если N - число разрядов, а n - объем выборки). В данном случае ввиду малого числа наблюдений, использование свойств χ^2 -распределения, конечно, необоснованно. Более того, теряет смысл и разбиение выборки на разряды, поэтому далее мы полагаем $N=n$. Однако указанная аналогия может быть полезна при необходимости верификации результатов для больших N .

Для проверки эффективности критерия (13) проводился эксперимент, заключающийся в генерации малых выборок с заданными в таблице 1 законами распределения и последующей их аппроксимацией набором тех же параметрических семейств функций. При построения выборки в качестве случайного значения ξ_j принимался квантиль, соответствующий вероятности $p_i = \frac{i-0.5}{N}, i=1, N$.

При этом способе построения выборки длина отрезка интегрирования $[a; b]$ растёт с ростом N как $\sim \sqrt{\ln N}$ при распределении 1, $\sim \ln N$ при распределении 2,4 и $\sim N$ при распределении 3. Поэтому оценка сверху погрешности традиционных квадратурных формул может не выполняться. Тем не менее отно-

шение $(b-a)/N$ для распределений 1,2 и 4 стремится к нулю, поэтому выражение под корнем в правой части (13) является интегральной суммой для интеграла.

$$J(s) = \int_{-\infty}^{+\infty} [F(s, \xi) - F_{эмп}(\xi)]^2 d\xi.$$

При этом вместо кажущегося естественным выбора η_i по формуле центральных прямоугольников: $\eta_i = (\xi_i + \xi_{i+1})/2$, величины $\eta_i \in [\xi_i; \xi_{i+1}]$, опре-

делялись равными квантилям соответствующих вероятностей:

$$p_{li} = \frac{i}{N}, i = \overline{1, N-1}.$$

Это позволяет исключить методическую ошибку аппроксимации, в частности, обеспечивает равенство нулю критерия (13) в случае, когда генерируемое и аппроксимирующее семейства совпадают.

Таблица 1.

	Плотность распределения	Функция потерь	Весовые коэффициенты
1	$f(\xi) = \frac{1}{s\sqrt{2\pi}} e^{-\frac{\xi^2}{2s^2}}$	$\ln(\sqrt{2\pi}s) + (-\xi^2/2s^2)$	$\frac{1}{s^2}$
2	$f(\xi) = \frac{1}{4s} (1 + \frac{ \xi }{s}) e^{-\frac{ \xi }{s}}$	$\ln 4s + \xi /s - \ln(1 + \xi /s)$	$1/s \left(s + \frac{ \xi }{s} \right)$
3	$f(\xi) = \frac{s}{\pi} * \frac{1}{s^2 + \xi^2}$	$\ln(s^2 + \xi^2) + \ln(\pi/s)$	$\frac{2}{s^2 + \xi^2}$
4	$f(\xi) = \frac{1}{2s \cdot ch^2\left(\frac{\xi}{s}\right)}$	$2 \ln Sech(\xi/s) + \ln 2s$	$\frac{1}{s\xi} th\left(\frac{\xi}{s}\right)$
5	$f(\xi) = \frac{1}{s} * \frac{e^{-\frac{\xi}{s}}}{\left(1 + e^{-\frac{\xi}{s}}\right)^2}$	$\frac{\xi}{s} + 2 \ln\left(1 + \exp\left(-\frac{\xi}{s}\right)\right) + \ln s$	$\frac{1}{2s\xi} th \frac{\xi}{2s}$

Аппроксимация построенных для каждого набора данных эмпирических функций распределения осуществлялась по принципу “каждая” “каждой”. Для оценки параметра масштаба s методом дихотомии решалось уравнение $dJ_{\theta}(s)/ds = 0$. Результаты

исследования для $N=25$ приведены в таблице 2. Целые числа 1-4 в обозначениях строк и столбцов соответствуют номерам исходных распределений в таблице 1. При дальнейшем увеличении N они практически не изменяются.

Таблица 2.

Вид распределения	1		2		3		4	
	s	$J_{\theta}(s)$	s	$J_{\theta}(s)$	s	$J_{\theta}(s)$	s	$J_{\theta}(s)$
1 (s=1.000)	1,000	0,00011	0,554	0,02213	0,543	0,07834	1,185	0,01328
2 (s=0.546)	1,071	0,03682	0,546	0,00020	0,510	0,06869	1,172	0,00982
3 (s=0.480)	1,665	0,21448	0,578	0,14051	0,480	0,00005	1,215	0,14803
4 (s=1.176)	1,043	0,02247	0,548	0,00967	0,521	0,07447	1,176	0,00005

Результаты моделирования подтверждают предположение о возможности непосредственного использования критерия (13) для оценки априорных гипотез при небольших N (≥ 25), хотя χ^2 - распределение при этом, возможно, не существует.

Значение критерия при аппроксимации множества компонент вектора ошибок “своим” параметрическим семейством действительно минималь-

но и отличается от нуля, по-видимому, лишь из-за погрешностей округления.

Значения s и $J_{\theta}(s)$ для распределения 4 оказываются близкими к аналогичным величинам для распределения 2. Поэтому далее мы будем рассматривать лишь хорошо различимые классы 1, 2 и 3. Кроме прочего, этот выбор связан также с простотой вычисления соответствующих этим классам весовых коэффициентов.

3. Оценка распределения ошибок по распределению невязок

Из опирающихся на соотношение (12) интуитивных соображений качественного характера ясно, что близость ошибок и невязок будет иметь место для векторов ошибок, у которых норма вектора невязок (сопряженность с нуль-пространством) в среднем выше. Эти соображения подтверждаются экспериментом, который проводился по следующей схеме.

Осуществлялась идентификация коэффициентов импульсной характеристики искажающего фильтра по входному (неискаженному) и выходному (искаженному и зашумленному с заданным распределением шума) изображениям, показанным на рисунке 1. Размеры опорной области (соответствующие характеристике размазывающего фильтра) задавались 5×5 , а размеры фрагментов, на которые разбивалось изображение, - 25×25 соответственно. Опорные области на фрагменте размещались "без перекрытия", поэтому размерности векторов ошибок и невязок были 25×1 . Реализации шума указанной длины для каждого заданного параметрического семейства распределений генерировались как описано выше.

После оценки на каждом фрагменте импульсной характеристики вычислялись невязки (5). Далее для каждой реализации вектора невязок строилась эмпирическая функция распределения, которая затем аппроксимировалась каждым из указанных выше трех (1, 2 и 3) параметрических семейств распределений и подсчитывалось значение критерия (13). Всего на изображении имелось 100 фрагментов, среди которых 80 информативных [5]. При моделировании векторов ошибок искусственно вводилась рандомизация, а результаты усреднялись по

100 экспериментам. В таблице 3 приведены данные о среднем количестве случаев K_{cp} , в которых каждое распределение оказывалось наилучшей, в смысле минимума критерия (13), аппроксимацией.

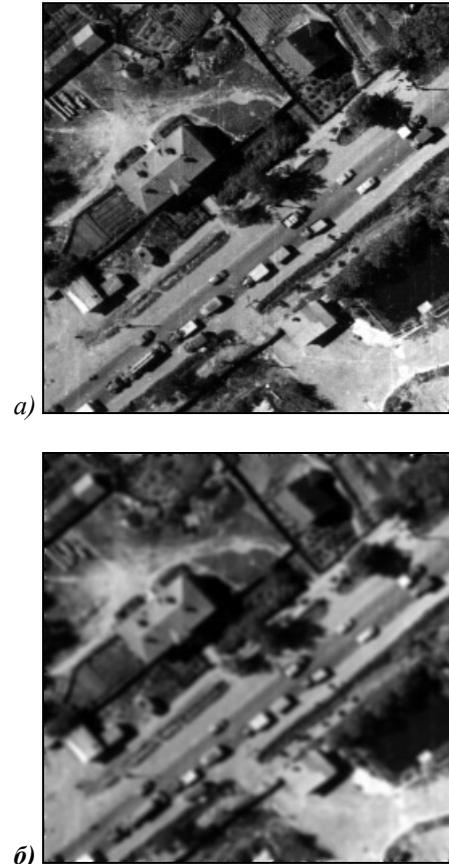


Рис. 1. Тестовые изображения: а – входное (неискаженное), б – выходное (искаженное и зашумленное).

Таблица 3.

Распределение шумов	Априорная гипотеза	Количество решений K_{cp} по изображению	Средняя норма вектора невязок
	1	43,4	4,439
1	2	34,7	4,405
	3	1,9	3,578
	1	22,6	4,614
2	2	53,1	4,753
	3	4,3	4,555
	1	0,3	2,236
3	2	20,9	9,843
	3	58,8	10,747

Приведенные данные хорошо согласуются с соображениями качественного характера, основанными на соотношении (12).

4. Пример построения процедуры обработки

С использованием полученных результатов построена процедура оценивания с преобразовани-

ем типа взвешивания, в которой после каждого этапа МНК - оценивания по вычисляемому для вектора невязок (5) критерию (13) принимается решение об использовании одного из типов преобразований (2 или 3). Эксперимент проводился на тех же тестовых изображениях (рис. 1), так что параметры модели искажений были известны и результаты оценивания можно было сравнить с их истинными значениями.

В таблице 4 приведены полученные в эксперименте (усредненные по всем информативным фрагментам на изображении) относительные погрешности оценивания (в процентах) при различных типах распределений ошибок в исходных данных и различных значениях параметра масштаба s .

Для каждого распределения приведена относительная погрешность ε_1 МНК-оценок и относительная погрешность ε_3 оценок на 3 шаге процедуры последовательного МНК-оценивания с промежуточными преобразованиями данных типа взвешивания.

Таблица 4.

Исходное распределение 1											
s	0,01	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9	1,0
ε_1	0,2766	2,272	5,073	7,749	9,589	10,37	16,07	16,01	23,22	25,28	24,47
ε_3	0,2768	2,274	5,062	7,782	9,589	10,42	16,04	16,09	23,21	25,47	24,66
Исходное распределение 2											
s	0,01	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
ε_1	0,4251	2,601	5,005	7,806	9,888	12,71	15,53	15,17	21,82	22,20	26,31
ε_3	0,4271	2,621	5,012	7,836	9,888	12,80	15,54	15,44	21,78	22,51	26,41
Исходное распределение 3											
s	0,01	0,05	0,1	0,15	0,2	0,25	0,3	0,35	0,4	0,45	0,5
ε_1	1,140	6,492	12,69	20,76	29,92	28,24	39,45	42,33	56,62	56,08	68,47
ε_3	0,773	4,551	9,15	14,72	23,60	20,83	26,59	22,28	35,81	37,55	41,95

Указанный в таблице выбор значений параметра s связан с тем, что при аппроксимации функции распределения (см. таблицу 2) оптимальное значение s для распределений 2 и 3 примерно вдвое меньше, чем для распределения 1. При аппроксимации по методу моментов для распределения 2 значение s будет точно в два раза меньше.

Заключение

Проведенные исследования показывают принципиальную возможность существенного повышения точности оценивания по малому числу наблюдений с использованием стратегии последовательного МНК-оценивания. В данном случае для анализа вектора невязок и принятия решения о типе преобразования взвешивания использовался критерий по смыслу максимально близкий к известным в теоретической статистике. При построении простых в вычислительном отношении практических процедур оценивания, возможно, более предпочтительным окажется применение критериев, основанных на непосредственном вычислении моментных и других

характеристик эмпирического распределения вектора невязок.

Литература

1. Даджион, Мерсеро. Цифровая обработка многомерных сигналов 1984, 488 с.
2. Демиденко Е.З. Линейная и нелинейная регрессии. 1983. М., "Финансы и статистика".
3. В.А. Фурсов. Идентификация линейной модели по малой выборке с совместной оценкой параметров и класса функций потерь. - Труды VI совещания по непараметрическим и робастным методам статистики в кибернетике. Томск, 1987, ч. II, с. 425-433.
4. Шамриков Б. М. Сравнительный анализ точности параметрической идентификации динамических объектов в разомкнутых и замкнутых автоматических системах. // Известия АН СССР. Техническая кибернетика. - 1986 г., №3, с.143-150.
5. Фурсов В.А. Идентификация оптических искажающих систем с отбором информативных фрагментов изображений. Компьютерная оптика. Вып. 14-15, 1995, с. 78-79.

The construction procedures of data conversion in the scheme of a sequential least-squares estimation of models of distortions on small fragments of images

V.A. Fursov, A.V. Ustinov
Image Processing Systems Institute of RAS

Abstract

The linear models with constant (space-wise) parameters are convenient in many respects for solving problems of evaluation (identification) of characteristics. Piecewise constant approximation by a set of shift-invariant linear models on small fragments of images is applied, if the distortions are nonlinear and (or) spatially dependent.

Citation: Fursov VA, Ustinov AV. The construction procedures of data conversion in the scheme of a sequential least-squares estimation of models of distortions on small fragments of images. *Computer Optics* 1997; 17: 111-115.

References

- [1] Dudgeon D, Mersereau R. *Multidimensional Digital Signal Processing*; Moscow: Mir Publisher; 1984; 488.
- [2] Demidenko EZ. *Linear and nonlinear regression*; Moscow: Finance and Statistics; 1983.
- [3] Fursov VA. Identification of a linear model based on a small sample with simultaneous estimation of parameters and class of loss functions. *Proceedings of the VI conference on nonparametric and robust methods of statistics in cybernetics*. Tomsk; 1987; 2: 425-433.
- [4] Shamrikov BM. Comparative analysis of the accuracy of parametric identification of dynamic objects in open-loop and closed-loop automatic control systems. *Izv. AN USSR; Tech. Cybernetics*; 1986; 3: 143-150.
- [5] Fursov VA. Identification of optical distorting systems by selecting image informative fragments. *Computer Optics* 1995; 14-15(1): 79-89.