

ПРИНЦИП СОГЛАСОВАННОСТИ ОЦЕНОК В ЗАДАЧЕ ИДЕНТИФИКАЦИИ МОДЕЛЕЙ ЦВЕТОВОСПРОИЗВЕДЕНИЯ

А.В. Никоноров, С.Б. Попов, В.А. Фурсов
Институт систем обработки изображений РАН,
Самарский государственный аэрокосмический университет

Аннотация

Приводится общая постановка задачи идентификации моделей цветовоспроизведения. Для линейных моделей описывается метод идентификации, основанный на принципе согласованности оценок. Предложена процедура сокращенного перебора при помощи генетических алгоритмов оптимизации. Проведено сравнение с полным перебором вариантов.

Постановка задачи идентификации модели цветообразования

Рассмотрим задачу формирования цвета при офсетной печати в несколько красок, которые в дальнейшем будем называть базовыми. Оставив пока в стороне систему восприятия цвета человеком, будем работать со спектральными коэффициентами отражения печатных оттисков в видимом диапазоне длин волн, в дальнейшем для краткости называемых спектрами. Спектр печатного оттиска образуется комбинацией базовых красок, взятых в определенных концентрациях и наложенных в заданной последовательности на печатную основу. Базовые краски характеризуются их спектрами в стопроцентных концентрациях, печатная основа также имеет свой спектр. Тогда образование спектра красочной смеси \mathbf{r} описывается некоторым физическим законом \mathfrak{F} в зависимости от спектров базовых красок \mathbf{r}_{bi} ($i = \overline{1, N}$), спектра печатной основы \mathbf{r}_p и набора концентраций красок \mathbf{a} :

$$\mathbf{r} = \mathfrak{F}(r_p, r_{p1}, \dots, r_{pN}, \mathbf{a}). \quad (1)$$

На практике спектры красочных смесей измеряются в диапазоне от 380нм до 700нм с шагом в 10нм. Тогда соотношение (1) можно записать в операторной форме:

$$\mathbf{r} = \mathfrak{F}(\mathbf{R}_p, \mathbf{a}), \quad (2)$$

$$\text{где } \mathbf{r} = \begin{bmatrix} r^1 \\ r^2 \\ \dots \\ r^{32} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} r_p^1 & r_{p1}^1 & \dots & r_{pN}^1 \\ r_p^2 & r_{p1}^2 & \dots & \\ \dots & \dots & \dots & \\ r_p^{32} & \dots & r_{pN-1}^{32} & r_{pN}^{32} \end{bmatrix},$$

$$\mathbf{a} = [a_1, a_2, \dots, a_M].$$

Будем искать наилучшее приближение \mathfrak{F} в параметрически заданном семействе $\{F(\mathbf{R}_p, \mathbf{a}, \mathbf{c})\}$. Конкретное значение параметра \mathbf{c}^* определяет конкретный вид оператора $F(\mathbf{R}_p, \mathbf{a})$.

В зависимости (2) участвуют разнородные параметры скаляры a_i и вектора \mathbf{r}_i . Для удобства построения модели (2) модифицировать и представить зависимость спектра смеси не от спектров \mathbf{r}_i чистых базовых красок и теоретических плотностей краски a_i , а от спектров базовых красок, соответствующих концентрациям $a_i - \mathbf{r}_i$. Тогда получим $\mathbf{r} = \mathfrak{F}(\mathbf{R}, \mathbf{a})$, а поиск приближение ищется в семействе $\{F(\mathbf{R}, \mathbf{c})\}$ [4, 5].

Задачу идентификации модели цветообразования можно формализовать как минимизацию функционала – критерия качества подобранной зависимости

$$I(F) = \int \Phi(\mathbf{r}, F(\mathbf{R}, \mathbf{c})) d\mathbf{R}. \quad (3)$$

Лучшим приближением зависимости (2) по выбранному нами критерию будет значение $F(\mathbf{R}, \mathbf{a}, \mathbf{c}^*)$, доставляющее минимум этому функционалу.

В качестве критерия качества может быть выбрано расстояние между \mathbf{r} и $F(\mathbf{R}, \mathbf{a}, \mathbf{c})$ в метрике L^2 или L^∞ . Для L^2 получим следующую модификацию (3)

$$I(F) = \int (\mathbf{r} - F(\mathbf{R}, \mathbf{c}))^2 d\mathbf{R}. \quad (4)$$

В таком случае будет получено лучшее приближение физической природы процесса цветообразования при офсетной печати.

Однако в качестве критерия может быть использована близость цветов соответствующих реальному и рассчитанному на модели спектрам. Если описать переход от спектров красочных смесей к цветам этих смесей как: $\mathbf{q} = Q(\mathbf{R}, \mathbf{c})$, где \mathbf{q} принадлежит некоторому цветовому пространству, в общем случае с римановой метрикой $\|\cdot\|_R$. В таком случае (3) принимает вид

$$I(F) = \|Q(\mathbf{r}) - Q(F(\mathbf{R}, \mathbf{c}))\|_R. \quad (5)$$

В частности, в качестве $Q(\mathbf{R})$ может быть использован переход от спектров отражения в пространство Lab [1, 2, 5]. В этом пространстве задана евклидова метрика.

Минимизация соотношений (4) и (5) в общем случае не эквивалентна. Следует также отметить, что работа с функционалом в пространстве спектров приводит к тому, что полученное значение $F(\mathbf{R}, \mathbf{c}^*)$ является приближением физического закона смешивания красок при печати. Использование же критерия качества в цветовом пространстве ведет к получению модели видеoinформационного тракта: печатный оттиск – человеческое зрение. Так как конечной оценкой качества построенной модели цветообразования будет человеческая экспертная оценка, то использование критерия (5), согласованного с этой экспертной оценкой, предпочтительно. Однако использование (5) приводит к увеличению вычислительной сложности.

В данной статье рассмотрим линейную регрессионную модель для F

$$\mathbf{r} = \mathbf{R}\mathbf{c} + \xi. \quad (6)$$

Оптимальное по среднеквадратичному отклонению (СКО) оценивание в предположении того, что случайная помеха ξ имеет нормальную природу, выполняется по методу наименьших квадратов (МНК). Однако в случае отличного от СКО критерия качества (5) или отличной от нормальной природы помехи МНК в чистом виде теряет оптимальные качества [4].

Общая схема решения на основе принципа согласованности

В работе [6] на примере была продемонстрирована эффективность применения принципа согласованного оценивания параметров регрессии при малом числе экспериментальных данных. Для линейной по параметрам регрессионной задачи (6) в работе [4] было проведено теоретическое обоснование метода. Для задачи (6) в этой работе был предложен следующий алгоритм формирования согласованных оценок: на системе (1) выбирается несколько вариантов – переопределенных систем меньшей размерности, получаемых из исходной вычеркиванием некоторых строк. Порядок выбора вариантов предполагается заданным априорно. На каждом из вариантов производится перебор всех подсистем некоторой заданной меньшей размерности. На каждой из подсистем выполняется МНК оценивание вектора параметров \mathbf{c} . По некоторому критерию взаимной близости выбираются наиболее согласованные оценки. Такой метод позволяет выделить подпространство строк матрицы \mathbf{X} , наиболее свободное от шума, и, таким образом уменьшить невязку в соотношении (6).

В случае, когда нельзя определить подсистемы-варианты исходя из априорной информации, можно организовать нахождение согласованных оценок при помощи полного перебора переопределенных подсистем.

Переборная реализация переборного оценивания по принципу согласованности

Пусть в общем случае матрица \mathbf{R} в задаче (6) имеет размерность $M \times N$. Тогда согласованное оценивание будет выполняться на подсистемах

$$\mathbf{r}_n = \mathbf{R}_n \mathbf{c}_n + \xi_n, \quad (7)$$

где \mathbf{R}_n – матрица $M_n \times N$, $N < M_n < M$, строки матрицы \mathbf{R}_n формируются из строк матрицы \mathbf{R} , взятых в произвольном порядке. Аналогичным образом из \mathbf{r} и ξ получаются соответственно \mathbf{r}_n и ξ_n

Каждая подсистема кодируется в виде бинарного \mathbf{b}_k вектора, в котором 1 в k -том разряде означает присутствие k -той строки исходной матрицы в системе, 0 – ее отсутствие. Норма Хэмминга вектора \mathbf{b}_k равна M_n .

Прямой перебор всех возможных подсистем в задаче (1) требует больших вычислительных затрат,

что приводит к практической невозможности решения таких задач на персональных компьютерах.

Сокращение временных затрат возможно либо за счет использования высокопроизводительной вычислительной техники либо за счет алгоритмического сокращения перебора. Рассмотрим подробнее оба варианта.

Принципы параллельной реализации согласованного оценивания

Было показано, что вычислительная сложность МНК оценки для подсистем сильно различается, поэтому при параллельной реализации перебора на кластере необходимо загружать все процессоры задачами одинаковой размерности, при этом вычислительные затраты на решение каждой задачи будут одинаковыми. Общее количество задач размерности M_n , равное числу сочетаний из M по M_n , можно разделить между всеми процессорами кластера поровну. Выдачу заданий каждому процессору в виде набора векторов \mathbf{b}_k должен проводить процессор-планировщик. При гетерогенной структуре кластера объем всех наборов одинаков. Иначе на некоторой размерности проводится калибровка скорости вычислений на каждом процессоре, в дальнейшем размер наборов должен быть скорректирован обратно пропорционально скорости вычислений.

Перебор \mathbf{b}_n в порядке возрастания в двоичном коде или коде Грея не может обеспечить постоянство нормы Хэмминга для всех вариантов. Это приводит к появлению большого количества «холостых» итераций перебора в случае, когда необходимо организовать совместную обработку подсистем одинаковой размерности.

Осуществить перебор \mathbf{b}_i , где i изменяется от 0 до $I = C_N^{M_n} - 1$, с сохранением нормы Хэмминга позволяет алгоритм, состоящий из следующих шагов.

1. Пусть стартовым значением процедуры перебора будет $\mathbf{b}_0 = \underbrace{0 \dots 0}_{N-M_n} \underbrace{1 \dots 1}_{M_n}$.

2. На i -ом шаге выполняется следующая процедура. Пусть l – наименьший разряд, для которого $\mathbf{b}_{i-1}^{l-1} = 1$ и $\mathbf{b}_{i-1}^l = 0$. Тогда $\mathbf{b}_i^j = \mathbf{b}_{i-1}^j$ при $j > l$, $\mathbf{b}_i^l = 1$ и $\mathbf{b}_i^{l-1} = 0$. Если $\mathbf{b}_{i-1}^0 = 0$ и $l > 1$, тогда $\mathbf{b}_i^{l-1-j} = \mathbf{b}_{i-1}^j$ при $j < l - 1$.

Если необходимого l не найдено, значит, перебор завершен и получен вектор $\mathbf{b}_l = \underbrace{1 \dots 1}_k \underbrace{0 \dots 0}_{N-k}$. Тогда

размерность M_n увеличивается на 1 и выполняется переход на шаг 1.

3. Если не получено \mathbf{b}_1 – переход на шаг 2.

В таблице 1 приведена последовательность \mathbf{b}_i для $N = 5$ и $k = 3$, $I = C_5^3 - 1 = 19$.

Задача оценивания выполняется для каждой переопределенной системы, определяемой вектором \mathbf{b}_i . Оценивание для нескольких значений размерности переопределенных систем, можно выполнить для каждой размерности при помощи описанной

процедуры. Назовем такой алгоритм поиска согласованных оценок алгоритмом с последовательным увеличением размерности.

Таблица 1. Формирование двоичного представления переопределенных систем

i	b_i	i	b_i	I	b_i
0	000111	7	011001	14	101010
1	001011	8	011010	15	101100
2	001101	9	011100	16	110001
3	001110	10	100011	17	110010
4	010011	11	100101	18	110100
5	010101	12	100110	19	111000
6	010110	13	101001		

Идея применения генетического алгоритма при получении согласованных оценок

Кроме использования высокопроизводительной вычислительной техники возможно алгоритмическое сокращение перебора при расчете оценок с использованием эволюционных методов поиска оптимального значения функции взаимной близости. Особенно эффективна данная методика, когда достаточно получить оценку параметров, дающую не обязательно глобальный оптимум критерия взаимной близости, а, хотя бы, гарантированно близкое к нему значение. Такими свойствами обладает оценивание с использованием генетического алгоритма оптимизации.

Реализацию базового генетического алгоритма (ГА) можно представить как итерационный процесс, включающий несколько этапов:

1. генерация начальной популяции,
2. воспроизводство «потомков»:
 - выбор родительской пары,
 - реализация оператора скрещивания,
 - реализация оператора мутации,
3. процедура отбора по наилучшему значению функции приспособленности и формирование на его основе нового поколения,
4. если не выполнено условие останова, то перейти к п.2.

Реализация согласованного оценивания с использованием технологии ГА отличается от классической схемы ГА. Отличие заключается в том, что при согласованном оценивании необходимо организовывать перебор подсистем с одинаковой размерностью. Рассмотрим модификацию классической схемы ГА, учитывающую эту особенность.

Начальная популяция формируется из взятых случайным образом бинарных векторов b_i с заданной нормой Хэмминга M_n . В терминах ГА каждая подсистема образует особь, причем записанную не в коде Грея, как принято в ГА, а в обычном двоичном коде.

Критерий взаимной близости, используемый при оценивании по принципу согласованности, принимается в качестве функции приспособленности.

Можно модифицировать операторы скрещивания и мутации таким образом, чтобы получаемая i -тая подсистема имела размерность $d_i \in [D_1; D_2]$. При близких значениях D_1 и D_2 сложность оценивания на наборах подсистем обрабатываемых каждым процессором будет приблизительно одинаковой.

Модифицированный оператор определяется как суперпозиция двух операторов: оператора изменяющего размерность подсистемы соответствующей особи (т.е. норму Хэмминга соответствующего бинарного вектора) и оператора инвертирующего значения генов особи без изменения нормы Хэмминга. Действие первого оператора заключается в следующем.

Пусть d_{i-1} норма Хэмминга особи до мутации. Случайным образом выбирается d_i новая норма Хэмминга. Если изменение нормы $\Delta_i = d_i - d_{i-1}$ положительно, то Δ_i случайно выбранных бит бинарного вектора меняются с 0 на 1, если Δ_i отрицательно, то наоборот, значения бит меняются с 1 на 0.

После изменения нормы Хэмминга применяется второй оператор, случайным образом меняющий k бит с 0 на 1 и столько же бит с 1 на 0.

Рассмотрим оператор скрещивания. Пусть b_1 и b_2 родительская пара бинарных векторов, участвующая в скрещивании. Рассмотрим следующие бинарные вектора $w_1 = b_1 \wedge \neg b_2$ – вектор, содержащий единичные биты на тех, где они есть в первом родительском векторе, то же самое для второго родительского вектора $w_2 = b_2 \wedge \neg b_1$, и $w_{12} = b_1 \wedge b_2$. Выбирается норма d_i . Выбираются случайные числа $0 \leq k_1 \leq \min(d_i/2, \|w_1\|_H)$, $0 \leq k_2 \leq \min(d_i/2, \|w_2\|_H)$ и $k_3 = d_i - k_1 - k_2$. Строятся вектора w_1', w_2', w_{12}' , содержащие, соответственно, k_1, k_2 и k_3 случайно выбранных единиц из векторов w_1, w_2, w_{12} . Результирующий вектор-потомство образуется как $b = w_1' \wedge w_2' \wedge w_{12}'$.

При параллельной организации ГА каждый процессор реализует свою популяцию, в которой происходит отбор. С некоторой периодичностью происходит объединение популяций и отбор лучших значений на всем кластере. Модификация ГА для сохранения нормы Хэмминга в определенных границах позволяет обеспечить равномерную загрузку процессоров кластера.

Численный пример

С использованием описанного метода было выполнено оценивание параметров смеси двух красок голубой и пурпурной в концентрациях 50% и 60% соответственно, при этом $N=2$. В качестве критерия взаимной близости использовался цветовой контраст в пространстве lab [5]

$$\Delta E = \sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2}. \quad (7)$$

МНК оценка параметров модели дала ошибку $\Delta E=2,3192$.

Полный перебор подсистем с размерностями от 6 до 10 включительно дал минимальное значение $\Delta E=0,01347$. На рис. 1 показан график зависимости цветового контраста от размерности подсистемы.

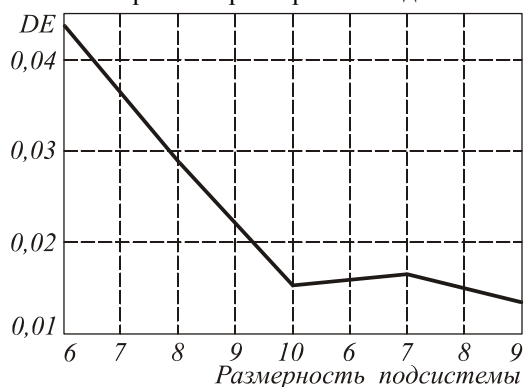


Рис. 1. Зависимость ошибки от размерности системы при полном переборе

При таком переборе было выполнено 27 000 000 итераций.

Та же задача решалась с использованием генетических алгоритмов, при этом было выполнено всего лишь 17 000 итераций. Остальные параметры генетического алгоритма были следующие. Размер популяции до отбора 320, размер после отбора 80. Селекция при скрещивании равновероятная, вероятность мутации 0,07. Оптимизация выполнялась в течение 50 эпох. На рис. 2 приведена зависимость лучшего значения ошибки в популяции от эпохи.

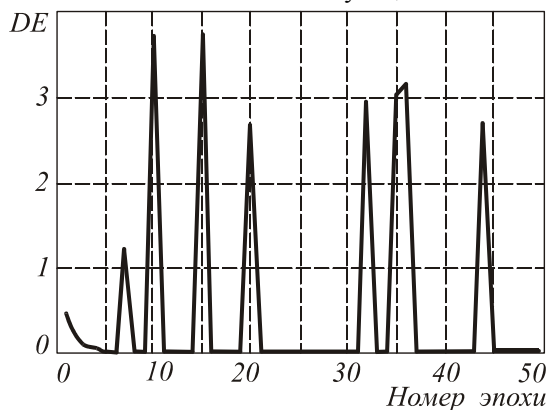


Рис. 2. Зависимость ошибки от эпохи генетического алгоритма

Полученное при этом значение ошибки модели составило $\Delta E=0,0239$, что всего лишь в два раза хуже, чем при полном переборе. Такой результат позволяет сделать вывод о том, что предлагаемый метод реализации оценивания с использованием генетического алгоритма оптимизации дает очень существенное уменьшение вычислительной сложности при незначительном (при определенных требованиях) снижении точности. В частности, при решении задачи идентификации системы цветовоспроизведения такое снижение является приемлемым, так как такая точность сравнима со значением повторяемости измерительного прибора, которая составляет $\Delta E=0,01$.

Заключение

Рассмотренный пример показывает целесообразность применения предложенных методов формирования согласованных оценок параметров моделей. Два варианта вычислительной реализации позволяют управлять соотношением «время на расчет» – «ошибка аппроксимации» модели.

Литература

1. Джадд Д., Вышецки Г. Цвет в науке и технике // М., Мир, 1978. - 580 с.
2. Шашлов Б. А. Цвет и цветовоспроизведение // М., Мир книги, 1995. - 316 с.
3. Фурсов В.А. Идентификация моделей систем формирования изображений по малому числу наблюдений. - Самара: ИПО СГАУ, 1998. - 128 с.
4. Фурсов В.А. Проблемы вычисления оценок по малому числу наблюдений // Лекции молодежной школы «Математическое моделирование 2001», Самара, 13-16 июня 2001. С. 56-63.
5. Никоноров А.В., Попов С.Б. Сравнительный анализ моделей цветообразования при офсетной многокрасочной печати // Компьютерная оптика. №23. 2002.
6. Никоноров А.В., Попов С.Б., Фурсов В.А. Идентификация моделей цветовоспроизведения // 6-я международная конференция «Распознавание образов и анализ изображений, новые информационные технологии» Великий Новгород 21-26 октября 2002. С. 431-436.