ПРОПОРЦИОНАЛЬНОЕ ПРОКРУСТОВО ПРЕОБРАЗОВАНИЕ В РАСПОЗНАВАНИИ РУКОПИСНЫХ ЦИФР И ЗНАКОВ

В.С. Муха, А.Н. Кузьков
Кафедра информационных технологий автоматизированных систем,
Белорусский государственный университет информатики и радиоэлектороники
Минск, Республика Беларусь
E-mail: mukha@bsuir.by, aliaksandr.kuzkou@gmail.com

Разработан алгоритм пропорционального Прокрустова преобразования. Алгоритм применен для распознавания рукописных цифр и знаков, написанных по образцу, предназначенному для всесоюзной переписи населения 1989 года. Приведены результаты распознавания.

Введение

Прокрустово преобразование состоит в том, чтобы преобразованием P подогнать одну $(k\times n)$-матрицу данных X к другой $(k \times n)$-матрице данных Y таким образом, чтобы результат преобразования $Z = PX$ как можно меньше отличался от Y. Здесь X можно ассоциировать с гостем Прокруста, персонажа легенды древней Греции , Y – с прокрустовым ложе, P – с «преобразователем»-Прокрустом [1].

Содержательный смысл Прокрустова преобразования состоит в следующем. Столбцы $(k\times n)$-матриц X и Y представляют собой упорядоченные точки в k-мерном Евклидовом пространстве E^k . Два упорядоченных набора точек представляют в этом пространстве некоторые конфигурации точек (фигуры). При $k = 2$ это фигуры на плоскости, при $k = 3$ – в трехмерном пространстве, и т.д. Задача состоит в том, чтобы с помощью Прокрустова преобразования как можно лучше подогнать одну фигуру к другой.

Ортогональное Прокрустово преобразование определяется выражением [1]

$$Z = GJ^T + cBX, \tag{1}$$

где B – ортогональная $(k \times k)$-матрица, $G^T = (g_1, g_2, ..., g_k)$, $J^T = (1_1, 1_2, ..., 1_n)$ – векторы-строки, c – скаляр, а задача ортогонального Прокрустова преобразования формулируется в виде

$$tr((Y-Z)^T(Y-Z)) \to \min_{B,G,c}, \quad B^TB = I. \tag{2}$$

Преобразование (1), (2) осуществляет смещение, поворот-отражение и масштабирование пространственной фигуры X с сохранением ее пропорций.

Ортогональное Прокрустово преобразование нашло применение в распознавании рукописных букв [2]. Однако оно является достаточно сложным в программной реализации. Поэтому в работе [3] предложено использовать в задачах распознавания линейное Прокрустово преобразование вида

$$Z = GJ^T + BX, \tag{3}$$

с критерием оптимальности вида

$$tr((Y-Z)^T(Y-Z)) \to \min_{B,G}. \tag{4}$$

Линейное Прокрустово преобразование (3), (4) легко программируется и хорошо себя зарекомендовало при распознавании печатных букв [3].

Вместе с тем при распознавании рукописных букв и цифр проявляются определенные недостатки линейного Прокрустова преобразования. С его помощью фигура может быть подогнана к «чужому» эталону лучше, чем к «своему». На рис. 1 приведен пример, когда изображение цифры 7 за счет сильного сжатия по горизонтали хорошо подгоняется к изображению эталона цифры 1.

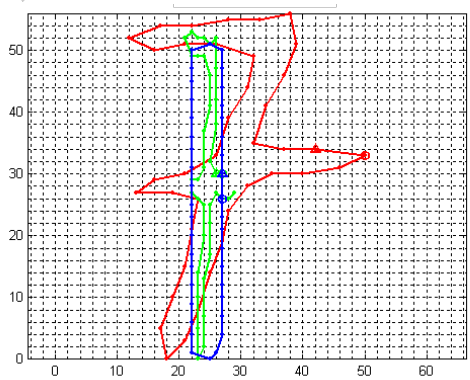

Рис. 1 – Линейное Прокрустово преобразование цифры 7 при ее подгонке к эталону цифры 1

На рис. 2 представлена подгонка изображения той же цифры 7 к изображению эталона цифры 7. Прокрустово расстояние преобразования на рис. 1 меньше по сравнению с расстоянием преобразования на рис. 2, вследствие чего цифра 7 при наличии эталонов цифр 1 и 7 распознается как единица. Такой эффект является следствием того, что линейное Прокрустово преобразование обладает большими возможностями непропорционального масштабирования. В линейном преобразовании, предложенном ниже и названном пропорциональным, удается в какой-то мере ограничить возможности непропорционального масштабирования и тем самым улучшить возможности распознавания для определенного класса задач.

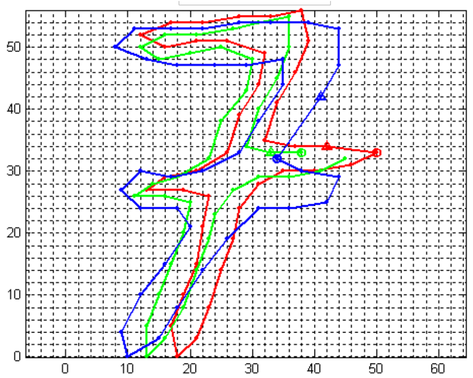

Рис. 2 – Линейное Прокрустово преобразование цифры 7 при ее подгонке к эталону цифры 7

I. Пропорциональное линейное Прокрустово преобразование

Определим Прокрустово преобразование $(k \times n)$-матрицы X к $(k \times n)$-матрице Y в виде

$$Z = GJ^T + BX, \tag{5}$$

где $B = (b_{i,j})$ – произвольная $(k \times k)$-матрица, $G^T = (g_1, g_2, ..., g_k)$, $J^T = (1_1, 1_2, ..., 1_n)$ – векторы-строки, а задачу пропорционального Прокрустова преобразования сформулируем в виде

$$tr((Y - Z)^T (Y - Z)) \to \min_{B,G},$$

$$b_{m,m} - b_{m+1,m+1} = 0 \quad \forall m = \overline{i, k-1}. \tag{6}$$

Преобразование (5), (6) назовем пропорциональным линейным Прокрустовым преобразованием.

Решение задачи (5), (6) определяется выражением

$$Z = \overline{\nu}_y + B(X - \overline{\nu}_x), \tag{7}$$

где $\overline{\nu}_x = (\nu_x, \nu_x, ..., \nu_x,)$, $\overline{\nu}_y = (\nu_y, \nu_y, ..., \nu_y,)$ – $(k \times n)$-матрицы с повторяющимися столбцами вида

$$\nu_x = \left(\frac{1}{n} \sum_{j=1}^{n} x_{i,j} \right), \quad \nu_y = \left(\frac{1}{n} \sum_{i=1}^{n} y_{i,j} \right).$$

Элементы $b_{1,1}$ и $b_{i,j}$, $i, j = \overline{1, k}$, $i \neq j$, матрицы B определяются как решение системы линейных алгебраических уравнений следующего вида

$$\begin{cases} b_{1,1} \sum\limits_{m=1}^{k} \overline{\mu}_{m,m} + \sum\limits_{m=1}^{k} \sum\limits_{\substack{l=1,\\ l \neq m}}^{k} b_{m,l} \overline{\mu}_{l,m} = \sum\limits_{m=1}^{k} \mu_{m,m}, \\ b_{1,1} \overline{\mu}_{i,j} + \sum\limits_{\substack{l=1,\\ l \neq i}}^{k} b_{m,l} \overline{\mu}_{l,m} = \mu_{i,j}, \; i \neq j = \overline{1, k}, \end{cases} \tag{8}$$

в которых

$$\mu = (\mu_{i,j}) = \left(\frac{1}{n} \sum_{l=1}^{n} y_{i,l} x_{j,l} \right),$$

$$\overline{\mu} = (\overline{\mu}_{i,j}) = \left(\frac{1}{n} \sum_{l=1}^{n} x_{i,l} x_{j,l} \right), \quad i, j = \overline{1, k}.$$

Алгоритм пропорционального Прокрустова преобразования (7) запрограммирован в виде m-файла-функции Матлаб.

II. Пропорциональное линейное Прокрустово преобразование в распознавании цифр и знаков

Пропорциональное линейное Прокрустово преобразование (7) применялось, наравне с ортогональным (1), (2) и линейным (3), (4) преобразованиями, к распознаванию рукописных цифр и некоторых знаков. Распознавались цифры и знаки, написанные от руки по образцам написания, рекомендованным для всесоюзной переписи населения 1989 года (рис. 3). В качестве эталонов использовались образцы, взятые из рис. 3.

Рис. 3 – Образцы написания цифр и знаков, рекомендованные к всесоюзной переписи населения 1989 г.

С помощью линейного Прокрустова преобразования правильно были распознаны 28 цифр и знаков из 36, то есть 78%. В этом случае некоторые из цифр 3, 7, 8, 9 были распознаны как единица вследствие особенностей линейного Прокрустова преобразования, проиллюстрированных графически на рис. 1. С помощью ортогонального и пропорционального линейного преобразований 36 предъявленных на распознавание цифр и знаков были распознаны стопроцентно. Таким образом, предложенное пропорциональное линейное Прокрустово преобразование оказывается хорошей альтернативой ортогональному Прокрустову преобразованию.

1. Schneider, W. Jesper. Matrix comparison, part 2: Measuring the resemblance between proximity measures or ordination results by use of the mantel and procrustes statistics /W. Jesper Schneider, Borlund Pia // Journal of the American Society for Information Science and Technology. – Volume 58. – Issue 11. – September 2007. – P 1596–1609.
2. Муха, В.С. Прокрустово преобразование в распознавании рукописных букв / В.С. Муха, А.Н. Кузьков // Информационные технологии и системы. Материалы международной научной конференции (БГУИР, Минск, Беларусь, 23 октября 2013). – Минск, БГУИР, 2013. – С. 298–299.
3. Муха, В.С. Линейное прокрустово преобразование в распознавании букв /В.С. Муха // Автоматика и вычислительная техника. – 2012. – № 3. – С. 36 – 48.