К ВОПРОСУ О МОДЕЛИРОВАНИИ МЕХАНИЗМОВ НОРМАЛИЗАЦИИ ЗРИТЕЛЬНЫХ ОБРАЗОВ

Е. П. Путятин, И. В. Шульгин, В. П. Юрченко, О. М. Абрамов

Введение

Пусть имеется некоторое множество M зрительных картин, заданных в виде функций распределения яркости $B(x, y)$ в поле зрения. Произведем разбиение множества M на подмножества $m_1, m_2, \ldots, m_i$ так, что 1) $m_i = \varnothing$; 2) $U_{im_i} = M$; 3) m_i и m_j либо не пересекаются, либо совпадают. Зрительные картины $B_1, B_2 \in M$ будем считать эквивалентными и обозначим $B_1 \sim B_2$, если существует такое m_i, что $B_1 \in m_i$ и $B_2 \in m_i$. Например, класс эквивалентности m_i может образовывать одинаковые, но смещенные по осям зрительные картины.

Если имеется некоторая зрительная картина B, подлежащая распознаванию, то в общем случае эта задача сводится к определению класса m_i, к которому принадлежит B. Фундаментальным алгоритмом распознавания является перебор. В этом случае элемент B сравнивают со всеми элементами множества M.

Задача распознавания существенно упрощается при сравнении не со всеми элементами M, а только с представителями в каждом классе эквивалентности m_i. Для осуществления данной возможности нужно построить оператор F, действующий из M в M, такой, что для каждого элемента $B \in m_i$ выполняется соотношение $F(B) = B_{i_0}$, где B_{i_0} — некоторый представитель, который будем называть эталоном или же нормальным элементом в классе эквивалентности m_i. В построении оператора F и состоит задача нормализации, которая решается в данной работе для сдвига, вращения и масштабных преобразований зрительных картин.

В литературе известны некоторые алгоритмы автоматической нормализации зрительных картин, например, метод интегральных моментов [1, 2]. В настоящей работе отмечаются необходимые и достаточные условия приведения эквивалентных картин к эталонному виду. Анализ этих условий позволяет говорить об известных алгоритмах как о частных случаях.

Математические модели нормализации

Пусть имеется некоторое множество преобразований G, обладающее свойствами группы с групповой операцией-суперпозицией. Построим разбиение множества M зрительных картин по заданной группе преобразований. В один класс эквивалентности m_i отнесем те и только те картины B_1 и B_2, для которых существуют преобразования S_1, $S_2 \in G$ такие, что

$$B_1 \circ S_1 = B_2 \circ S_2. \tag{1}$$

Примерами групп преобразований являются следующие: G^1 — группа сдвигов зрительной картины $B_i \in M$ по одной из осей координат; G^2 — группа поворотов; G^3 — группа изменения масштаба по одной из осей; G^4 — группа преобразований подобия и др. Каждая группа $G^i (i = 1, ..., 6)$ порождает разбиение

$$M = \{m_1^i, m_2^i, \ldots, m_i^i, \ldots\}.$$

Все перечисленные группы являются однопараметрическими, т. е. каждому преобразованию S из группы соответствует одно и только одно вещественное число t, которое полностью определяет преобразование.

В классе эквивалентности, к которому принадлежат зрительные картины B_i, возьмем некоторую эталонную картину B_0. Тогда для однопараметрической группы оператор F будет иметь вид

$$B_i \circ S_{t_i} = B_0 \circ S_{t_0}; \quad B_0 = B_i \circ S_{t_i} \circ S_{t_0}^{-1} = B_i \circ S_t, \tag{2}$$

где t_i и t_0 — вещественные числа, характеризующие преобразования S_i и S_0 зрительных картин. Параметр t является функционалом соответствующих зрительных картин

$$t = \Phi(B_i, B_0). \tag{3}$$

Аналогично для двух эквивалентных зрительных картин $B_1 \circ S_{t_1} = B_2 \circ S_{t_2}$ имеем

$$B_1 = B_2 \circ (S_{t_2} \circ S_{t_1}^{-1}) = B_2 \circ S_t. \tag{4}$$

Здесь параметр t зависит от t_1 и t_2

$$t = P(t_1, t_2) = P[\Phi(B_1, B_0), \Phi(B_2, B_0)], \tag{5}$$

где P — некоторая функция. Функционалы $\Phi(B_1, B_0)$ и $\Phi(B_2, B_0)$ отображают двумерное пространство функций яркости B (x, y) в одномерное пространство вещественных чисел.

Рассмотрим некоторые конкретные группы преобразований.

1. *Группа* G^1 *сдвигов по одной из осей координат*. Пусть имеется зрительная картина $B(x, y)$, которая дважды подвергается преобразованиям сдвига по оси x: один раз на величину l_1 и второй — на величину l_2. Очевидно, что результирующее преобразование S_l для

$$B \circ S_{l_1} \circ S_{l_2} = B \circ S_l$$

полностью характеризуется параметром l, равным сумме параметров преобразования l_1 и l_2. Таким образом, группа сдвигов G^1 изоморфна группе вещественных чисел по сложению.

На основании соотношений (2) и (3) и учитывая изоморфизм по сложению, получаем необходимые и достаточные условия эквивалентности зрительных картин B при фиксированном эталоне B_0

$$\Phi(B) - \Phi(B_0) = l, \tag{6}$$

где l — величина сдвига для зрительных картин

$$B_0(x, y) = B(x - l, y). \tag{7}$$

Учитывая соотношение (2), оператор F автоматического приведения зрительной картины $B(x, y)$ к эталонному виду (оператор центрирования) будет записан в виде

$$B_0(x, y) = B[x + \Phi(B), y]. \tag{8}$$

В качестве функционалов $\Phi(B)$ могут быть приняты все те, для которых выполняется условие (6): центр тяжести

$$x_c = \frac{\iint\limits_D B(\xi, \eta)\, \xi d\xi d\eta}{\iint\limits_D B(\xi, \eta)\, d\xi\, d\eta}, \tag{9}$$

где D — область интегрирования зрительной картины; «логарифмический центр тяжести»

$$x_e = \ln \iint\limits_D B(\xi, \eta)\, l^{\xi} d\xi\, d\eta; \tag{10}$$

параметры любых соответствующих характерных точек на зрительной картине (углы излома, максимальная яркость и т. д.). Вопрос о выборе конкретного вида функционала $\Phi(B)$ приобретает первостепенное значение при технической реализации оператора центрирования.

Всевозможные смещения эталона $B(x, y)$ по двум координатам x и y образуют двухпараметрическую группу. Легко показать, что в этом случае оператор F центрирования зрительной картины $B(x, y)$ по двум координатам будет иметь вид

$$B_0(x, y) = B[x + \Phi_1(B), y + \Phi_2(B)]. \tag{11}$$

Функционалы $\Phi_1(B)$ и $\Phi_2(B)$ должны удовлетворять условиям

$$\Phi_1(B) - \Phi_1(B_0) = l,$$

$$\Phi_2(B) - \Phi_2(B_0) = p, \tag{12}$$

где l и p — произвольные смещения зрительной картины B соответственно по координатам x и y относительно эталонной картины B_0.

Конкретные виды функционалов $\Phi_1(B)$ и $\Phi_2(B)$ могут совпадать с выражениями (9) — (10) для x_c и y_c.

2. *Группа вращений G^2 вокруг начала координат.*

Если зрительная картина $B(x, y)$ подвергается последовательным преобразованиям поворота на углы Θ_1, Θ_2, то результирующее преобразование S_0, определяемое из

$$B \circ S_{\theta_1} \circ S_{\theta_2} = B \circ S_\theta,$$

характеризуется суммой $\Theta = \Theta_1 + \Theta_2$, т. е. группа вращений, как и группа сдвигов, изоморфна группе вещественных чисел по сложению.

Выберем эталонную зрительную картину $B_0(x, y)$, например, такую, что угол наклона некоторой ее оси к оси абсцисс будет равен нулю, т. е. $\Theta_0 = 0$. Тогда для зрительной картины $B(x, y)$, повернутой на угол Θ относительно эталонной, справедливо соотношение $B = B_0 \cdot S_\Theta^{-1}$ или

$$B(x, y) = B_0(x\cos\theta + y\sin\theta, \quad -x\sin\theta + y\cos\theta).$$

Оператор F автоматической «отстройки» от вращения имеет вид

$$B_0(x, y) = B[x\cos\Phi(B) - y\sin\Phi(B), \quad x\sin\Phi(B) + y\cos\Phi(B)], \tag{13}$$

причем $\Phi(B) - \Phi(B_0) = \theta$, в частности, $\Phi(B_0)$ может выбираться равным нулю.

В полярной системе координат оператор (13) имеет более простой вид

$$B_0(\rho, \varphi) = B[\rho, \varphi + \Phi(B)], \tag{14}$$

где $\Phi(B)$ — вещественный функционал, вычисляемый в полярной системе координат и удовлетворяющий условию (6).

В декартовой системе координат в качестве функционалов $\Phi(B)$, удовлетворяющих (13), могут быть использованы, например,

$$\Phi(B) = \frac{\iint\limits_D B(\xi, \eta)\,\mathrm{arctg}\,\frac{\xi}{\eta}\,d\xi d\eta}{\iint\limits_D B(\xi, \eta)\,d\xi\,d\eta} \tag{15}$$

или

$$\Phi(B) = \log_a \iint\limits_D B(\xi, \eta)\,a^{\mathrm{arctg}\,\frac{\xi}{\eta}}\,d\xi\,d\eta. \tag{16}$$

Для функционала в операторе (14) могут быть использованы выражения центра тяжести в полярной системе координат типа (9) или (10).

3. *Группа G^3 изменения масштаба зрительной картины по одной из координат.*

Пусть зрительная картина $B(x, y)$ дважды подвергается преобразованию масштаба по оси x: первый раз с коэффициентом λ_1, второй — с коэффициентом λ_2. Тогда результирующее преобразование S_λ, определяемое как $B \circ S_\lambda = B \circ S_{\lambda_1} \cdot S_{\lambda_2}$, характеризуется произведением $\lambda = \lambda_1\lambda_2$ и, следовательно, группа изменения масштаба по одной из осей изоморфна группе вещественных чисел по умножению.

Пользуясь соотношениями (4) и (5), запишем необходимое и достаточное условие эквивалентности зрительной картины $B(x, y)$ и некоторой эталонной $B_0(x, y)$ такой, что $B(x, y) = B_0(\lambda x, y)$,

$$\Phi(B_0) = \lambda\Phi(B), \tag{17}$$

а также вид оператора нормализации F

$$B_0(x, y) = B[x\Phi(B), y]. \tag{18}$$

Функционал для эталонной картины может быть выбран равным единице $\Phi(B_0) = 1$. В качестве $\Phi(B)$ можно принять обычные линейные функционалы вида

$$\Phi(B) = \iint_D B(\xi, \eta)\psi(\eta)\,d\xi d\eta, \tag{19}$$

где $\psi(\eta)$ — произвольная функция.

Условию (17) удовлетворяет также нелинейный функционал

$$\Phi(B) = \sqrt{\iint_D B(\xi\ \eta)\,\xi\psi(\eta)\,d\xi\,d\eta}. \tag{20}$$

4. *Группа* G^4 *подобия.*

Ясно, что группа преобразований подобия изоморфна группе вещественных чисел по умножению. Условие эквивалентности зрительных картин будет определяться соотношением (17), а оператор F нормализации для картин типа $B(x, y) = B_0(\lambda x, \lambda y)$ имеет вид

$$B_0(x, y) = B[x\Phi(B), y\Phi(B)]. \tag{21}$$

Примером функционала $\Phi(B)$ является известное выражение [2]

$$\Phi(B) = \sqrt{\iint_D B(\xi, \eta)\,d\xi d\eta}. \tag{22}$$

ЛИТЕРАТУРА

1. Минг Куэй Ху. Математическая модель восприятия. Сб. «Проблемы бионики». Изд-во «Мир», 1965.

2. В. А. Махонин, В. П. Савельев. О моделировании константных свойств при опознании. Сб. «Проблемы инженерной психологии», вып. 4. Л., 1966.