ДОКЛАДЫ АКАДЕМИИ НАУК, 2011, том 437, № 3, с. 313–315

ИНФОРМАТИКА

УДК 004.93'11

ОЦЕНКА РАССТОЯНИЯ МЕЖДУ ИЗОБРАЖЕНИЯМИ ПРИ ПАРАЛЛЕЛЬНОМ ПЕРЕНОСЕ

© 2011 г. Член-корреспондент РАН **В. Л. Арлазаров, О. А. Славин, А. Г. Хованский**

Поступило 20.09.2010 г.

О п и с а н и е э ф ф е к т а. Пусть задан конечный набор эталонных изображений (изображений символов) на плоскости, каждое из которых лежит внутри некоторой стандартной выпуклой фигуры (например, внутри заданного прямоугольника). Возьмем еще одно (тестируемое) изображение, которое надо распознать, т.е. отождествить с одним из эталонных изображений. Наложим одно из эталонных изображений на тестируемое.

П р а к т и ч е с к а я з а д а ч а. Найти параллельный перенос эталонного изображения, при котором его совпадение с тестируемым изображением максимально.

Эта задача действительно встречается на практике. Многочисленные эксперименты выявили следующий эффект.

Н а б л ю д а е м ы й э ф ф е к т. Если наугад взятое наложение удачно, т.е. если несовпадение эталонного и тестируемого изображений составляет малую часть эталонного изображения, то параллельный перенос, максимизирующий совпадение изображений, мал (т.е. по длине составляет малую часть диаметра эталонного изображения).

Ниже приводится математическая модель, объясняющая этот эффект на качественном уровне. В модели вместо изображения рассматривается пара, состоящая из множества A (закрашенная часть изображения), и плотности $\phi : A \to R$ (интенсивность закрашивания). Степень совпадения изображений с плотностями ϕ_1 и ϕ_2 моделируется интегралом

$$\int |\phi_1 - \phi_2| d\mu.$$

В разделах 1, 2 мы оцениваем снизу меру несовпадения эталонного изображения и его образа при отображении, сохраняющем меру. Специфика задачи для параллельного переноса обсуждается в разделе 3. В разделе 4 приводятся менее точные, но более обозримые оценки. Обсуждаемый эффект объясняется в разделе 5.

Институт системного анализа Российской Академии наук, Москва

1. А б с т р а к т н а я з а д а ч а. Пусть Ω — пространство с мерой $d\mu$ и X — множество пар $x = (A, \phi)$, где $A \subset \Omega$ — измеримое подмножество и $\phi : A \to R$ — интегрируемая неотрицательная функция, называемая плотностью. На X определено расстояние $p(x_1, x_2) = \int_\Omega |\phi_1 - \phi_2| d\mu$, где

$$x_1 = (A_1, \phi_1), \quad x_2 = (A_2, \phi_2).$$

Пусть $F : \Omega \to \Omega$ — взаимнооднозначное отображение, сохраняющее меру, т.е. $F^* d\mu = d\mu$. Отображение F индуцирует изометрию $\Phi : X \to X$, переводящую точку $x = (A, \phi)$ в точку $\Phi(x) = (F(A), (F^{-1})^*\phi)$. По подмножеству $\Theta \subset \Omega$ и числу K определим подмножество $Y \subset X$, состоящее из точек (A, ϕ), таких что: 1) $A \subset \Theta$, 2) $\int_\Omega \phi d\mu = K$.

З а д а ч а 1. Для функции $S(\Theta, K, F) = \min_{x \in Y} p(x, \Phi(x))$ найти (наилучшую) оценку снизу.

Один из возможных вариантов ответа на задачу 1 изложен в следующем разделе.

Остановимся на одной модификации задачи 1. Расширим множество X и определим множество Z, состоящее из пар $z = (A, \psi)$, в которых плотность ψ может концентрироваться в точках. Точнее, в Z в качестве плотностей рассматриваются обобщенные плотности ψ, являющиеся суммами интегрируемых функций и конечных линейных комбинаций δ-функций в точках $q \in \Omega$ (мы будем обозначать их символом δ_q), т.е. $\psi = \phi + \sum a_k \delta_{q_k}$. На пространство Z автоматически переносятся определения расстояния p и отображения Φ. Пусть даны отображение $F : \Omega \to \Omega$, подмножество $\Theta \subset \Omega$ и число K. Пусть подмножество $W \subset Z$ состоит из точек (A, ψ), где $\psi = \phi + \sum a_k \delta_{q_k}$, таких что: 1) $A \subset \Theta$, 2) $\phi \ge 0$, $a_k \ge 0$, носитель функции ϕ лежит в A и $q_k \in A$, 3) $\int_\Omega \phi d\mu + \sum a_k = K$.

З а д а ч а 1'. Для функции $G(\Theta, K, F) = \min_{z \in W} p(z, \Phi(z))$ найти (наилучшую) оценку снизу.

2. Частичное решение задач 1 и 1'. Справедлива следующая

Теорема 1. *Если для некоторого целого n множество* $F^n(\Theta)$ *не пересекается с множеством* Θ, *тогда* $S(\Theta, K, F) \ge \frac{2K}{n}$ и $G(\Theta, K, F) \ge \frac{2K}{n}$.

При выполнении дополнительных условий оценку из теоремы 1 нельзя улучшить. Главное из них — следующее

Условие I_n. Существует точка $a \in \Omega$, такая что точки $a, F(a), \dots, F^n(a)$ различны и все они, за исключением точки $F^n(a)$, лежат в множестве Θ.

Предположим дополнительно, что множество Ω снабжено топологией, в которой множество Θ открыто и отображение F непрерывно.

Теорема 2. *Пусть выполнены перечисленные выше условия и* $F^n(\Theta) \cap \Theta = \phi$. *Тогда* $S(\Theta, K, F) = G(\Theta, K, F) = \frac{2K}{n}$.

Теорема 2 для случая задачи 1' допускает усиление: не нужно требовать, чтобы множество Ω было снабжено топологией, в которой множество Θ открыто и отображение F непрерывно.

Теорема 2'. *Пусть выполнено условие* I_n *и* $F^n(\Theta) \cap \Theta = 0$. *Тогда* $G(\Theta, K, F) = \frac{2K}{n}$.

3. Случай параллельного переноса. Пусть пространство Ω с мерой $d\mu$ — это пространство R^n с евклидовой мерой, отображение F — это сдвиг на фиксированный вектор $\mathbf{a} \in R^n$ (т.е. $F(x) = x + \mathbf{a}$), а множество Θ — это любое n-мерное выпуклое подмножество в R^n. Рассмотрим совокупность неотрицательных интегрируемых функций ϕ в R^n, тождественно равных нулю вне выпуклого множества Θ и удовлетворяющих соотношению

$$\int_{R^n} \phi(x) d\mu = K. \qquad (1)$$

Задача 1 для параллельного переноса. Найти (наилучшую) оценку снизу функции $S(\Theta, K, \mathbf{a})$, заданной формулой

$$S = \min_{\phi} \int_{R^n} |\phi(x) - \phi(x - \mathbf{a})| d\mu.$$

Задача 1' в рассматриваемом случае формулируется аналогично.

Вектор $\mathbf{a} \in R^n$ будем записывать в виде $\mathbf{a} = |\mathbf{a}|\mathbf{v}$, где $\mathbf{v}$ — вектор единичной длины и $|\mathbf{a}|$ — длина вектора $\mathbf{a}$. Для вектора $\mathbf{v}$ единичной длины назовем $\mathbf{v}$-диаметром $D_V(\Theta)$ выпуклой фигуры Θ супремум длин отрезков, являющихся сечениями области Θ прямыми, параллельными вектору $\mathbf{v}$. Для ненулевого вектора $\mathbf{a} = |\mathbf{a}|\mathbf{v}$ назовем a-диаметром $D_{\mathbf{a}}(\Theta)$ выпуклой фигуры Θ отношение $\frac{D_{\mathbf{v}}(\Theta)}{|\mathbf{a}|}$ ее v-диаметра к длине $|\mathbf{a}|$ вектора $\mathbf{a}$.

Решение задач 1 и 1' для вектора $\mathbf{a} = |\mathbf{a}|\mathbf{v}$ в случае, когда a-диаметр фигуры Θ не является натуральным числом, доставляет следующая

Теорема 3. *Если для некоторого целого* $n \ge 0$ *справедливы неравенства*

$$n|\mathbf{a}| < D_{\mathbf{v}}(\Theta) < (n+1)|\mathbf{a}|,$$

то $S = G = \frac{2K}{n+1}$.

Решение задачи 1' для вектора $\mathbf{a} = |\mathbf{a}|\mathbf{v}$ в случае, когда a-диаметр фигуры Θ является натуральным числом, доставляет следующая

Теорема 4. *Если для некоторого целого* $n \ge 0$ *справедливо равенство*

$$n|\mathbf{a}| = D_{\mathbf{v}}(\Theta),$$

то $G = \frac{2K}{n+1}$.

Если выполняется равенство $n|\mathbf{a}| = D_{\mathbf{v}}(\Theta)$, то минимум в задаче не достигается на обычных функциях. Однако минимум достигается на линейной комбинации δ-функций и он равен $\frac{2K}{n+1}$.

4. Оценка расстояний между изображениями. В этом пункте для краткости мы будем обсуждать только задачу 1. Формула для функции $S(\Theta, K, \mathbf{a})$ достаточно громоздка. В следствии 1 приводится неточная, но достаточно хорошая оценка этой функции, которая вполне обозрима. В следствии 2 приводится оценка этой функции, которая для фиксированного вектора $\mathbf{v}$ при малых значениях длины $|\mathbf{a}|$ линейна по длине $|\mathbf{a}|$, т.е. равна $C|\mathbf{a}|$ с явно выписанной константой C.

Следствие 1. *Справедлива оценка*

$$S(\Theta, K, \mathbf{a}) \ge \frac{2K}{D_{\mathbf{v}}(\Theta) + |\mathbf{a}|}|\mathbf{a}|.$$

Эта оценка точна при $|\mathbf{a}| = D_{\mathbf{v}}(\Theta), \frac{D_{\mathbf{v}}(\Theta)}{2}, \frac{D_{\mathbf{v}}(\Theta)}{3}, \dots$ *и при* $|\mathbf{a}| \to \infty$ (*т.е., образно говоря, точна при* $|\mathbf{a}| = \frac{D_{\mathbf{v}}(\Theta)}{0}$).

Следствие 2. *Для любого* $0 \le y_0$ *при* $0 \le |\mathbf{a}| \le y_0$ *имеем*

$$S(\Theta, K, \mathbf{a}) \ge \frac{2K}{D_{\mathbf{v}}(\Theta) + y_0}|\mathbf{a}|,$$

при $|\mathbf{a}| > y_0$ *имеем*

$$S(\Theta, K, \mathbf{a}) \ge \frac{2Ky_0}{D_{\mathbf{v}}(\Theta) + y_0}.$$

В частности, для $y_0 = D_v(\Theta)$ *получаем следующее: при* $0 \le |\mathbf{a}| \le D_v(\Theta)$ *имеем*

$$S(\Theta, K, \mathbf{a}) \ge \frac{K}{D_v(\Theta)}|\mathbf{a}|,$$

при $|\mathbf{a}| > D_v(\Theta)$ *имеем* $S(\Theta, K, \mathbf{a}) \ge K$ (*отметим, что на самом деле при* $|\mathbf{a}| > D_v(\Theta)$ *справедливо равенство* $S(\Theta, K, \mathbf{a}) = 2K$).

5. Качественное объяснение рассматриваемого эффекта. Пусть ϕ – неотрицательная интегрируемая функции на R^n, равная нулю вне выпуклого множества Θ, для которой выполнено равенство (1). Пусть $g(x)$ – другая неотрицательная интегрируемая функция на R^n. Пусть

$$\int\limits_{R^n}|\phi(x) - g(x)|dx = I.$$

Допустим, что интеграл I значительно меньше, чем K.

Задача 2. Меняя параметр a, добиться, чтобы интеграл

$$\int\limits_{R^n}|\phi(x - a) - g(x)|dx = M(a)$$

принял возможно меньшее значение.

Наша цель – показать, что если $K \gg I$, то для минимизации числа $M(a)$ интересны лишь вектора $\mathbf{a} = |\mathbf{a}|\mathbf{v}$, которые значительно короче, чем v-диаметр $D_v(\Theta)$ области Θ.

Пусть $S(\mathbf{a}) = \min \int\limits_{R^n}|\phi(x) - \phi(x - \mathbf{a})|d\mu$, где минимум берется по всем неотрицательным функциям ϕ, удовлетворяющим соотношению (1) и имеющим носитель в области Θ. Мы можем применять найденные выше оценки функции $S(\mathbf{a})$.

Теорема 5. *Фиксируем вектор* $\mathbf{v}$ *единичной длины. Пусть* $d(\mathbf{v})$ – *минимальное число, такое что для вектора* $\mathbf{a} = |\mathbf{a}|\mathbf{v}$ *при* $|\mathbf{a}| > d(\mathbf{v})$ *выполнено неравенство* $S(\mathbf{a}) > 2I$. *Тогда при* $|\mathbf{a}| > d(\mathbf{v})$ *справедливо неравенство*

$$\int\limits_{R^n}|\phi(x - \mathbf{a}) - g(x)|d\mu > I.$$

В частности, для минимизации $M(a)$ *достаточно рассматривать лишь вектора* $a = |a|v$, *для которых* $|\mathbf{a}| < d(\mathbf{v})$.

Следствие 3. *Пусть для фиксированного единичного вектора* $\mathbf{v}$ *выполнено неравенство* $I < \frac{D_v(\Theta)}{2}$. *Тогда для минимизации интеграла* $M(\mathbf{a})$ *на прямой* $\mathbf{a} = |\mathbf{a}|\mathbf{v}$, *натянутой на вектор* $\mathbf{v}$, *достаточно рассматривать вектора* a, *для которых выполнено неравенство* $|\mathbf{a}| < \frac{2ID_v(\Theta)}{K}$. *При* $|\mathbf{a}| > \frac{2ID_v(\Theta)}{K}$ *справедливо неравенство* $\int\limits_{R^n}|\phi(x - \mathbf{a}) - g(x)|d\mu > I$.

Следствие 3 дает качественное объяснение обсуждаемого эффекта. Действительно, если $I \ll K$, то число $\frac{2I}{K}$ мало. Поэтому минимум $M(\mathbf{a})$ достигается при длине вектора $\mathbf{a}$ значительно меньшей, чем диаметр $D_v(\Theta)$ фигуры Θ в направлении вектора $\mathbf{v}$.