

Анализ величин ошибки при неточном проективном преобразовании четырехугольника

И.А. Коноваленко^{1,2,3}, Ю.А. Шемякина^{3,4}

¹Институт проблем передачи информации им. А.А. Харкевича Российской академии наук — ИППИ РАН, Большой Каретный переулок, 19 стр. 1, Москва, Россия, 127051

²Московский физико-технический институт (государственный университет) — МФТИ (ГУ), Институтский переулок, 9, Долгопрудный, Московская область, Россия, 141701

³Smart Engines Ltd., проспект 60-летия Октября, 9, Москва, Россия, 117312

⁴Институт системного анализа Федерального исследовательского центра «Информатика и управление» Российской академии наук — ИСА ФИЦ ИУ РАН, проспект 60-летия Октября, 9, Москва, Россия, 117312

Аннотация. В данной работе предлагается метод оценки качества детектирования прямоугольного документа на фотографии, сделанной с произвольного ракурса. Пусть дан истинный четырёхугольник границ документа на фотографии и соответствующий четырёхугольник, найденный детектором. Тогда эти четырёхугольники связаны известным проективным преобразованием, малость которого характеризует качество работы детектора. При этом величина ошибки для каждой точки документа описывается её смещением, т.е. расстоянием между её истинным и ожидаемым положением в связанных с задетектированным четырёхугольником координатах. В качестве меры ошибки работы детектора предложено брать максимальное смещение точки документа. В работе приведен контрпример для гипотезы о том, что максимальное смещение достигается на углу документа; доказано, что максимальное смещение достигается на границе документа; приведено аналитическое решение задачи поиска максимального смещения на границе документа.

1. Введение

Задача проективного исправления изображений естественным образом повсеместно возникает в компьютерном зрении при наблюдении в модели камеры-обскуры плоских сцен с произвольного ракурса. Например, в статьях [1] и [2] предлагаются алгоритмы для проективного сопоставления контуров и отрезков соответственно. В работе [3] описывается сегментация проективно исправленных номерных знаков автомобилей, а в [4] - паспортов. В статье [5] анализируются подходы к решению задачи проективного наведения на примере страховых свидетельств. В статье [6] описывается система «Snapscreen», которая детектирует и проективно исправляет фотографию экрана телевизора с дальнейшим распознаванием телевизионной передачи.

Существует необходимость сравнивать между собой качество работы алгоритмов проективного исправления. Для этого, как правило, на некоторой выборке примеров задаётся истинное проективное преобразование. После этого задача сводится к вычислению сходства между экспериментальным и истинным проективными преобразованиями. Однако, проективное преобразование в общем случае является 8-параметрическим и

определение скалярной величины сходства между ними является нетривиальным и допускает множество способов.

В литературе для этого чаще всего используется метрика Хаусдорфа [7] и большое количество её вариаций [8]. Она определяется на множествах и равняется наибольшему расстоянию от точек одного множества до соответствующих им ближайших точек второго множества. В компьютерном зрении, как правило, множествами являются искажённые контуры одного и того же объекта. Например, в работе [9] метрика Хаусдорфа используется для сопоставления объектов, а в работе [10] — для детекции лиц. Ещё одним часто используемым на практике инструментом является коэффициент Жаккара [11]. Он также определяется на множествах и равняется площади пересечения множеств, отнесённой к площади их объединения. Данный коэффициент является целевым, например, на конкурсе по анализу и распознаванию документов «ICDAR» [12].

Очевидное положительное свойство данных мер ошибки состоит в том, что идеальному ответу алгоритма они сопоставят своё наилучшее возможное значение. Однако обратное уже является неверным. Действительно, пусть стоит задача детекции и проективного исправления фотографии банковской карты. Тогда по вышеприведённым мерам алгоритм, который найдёт банковскую карту повернутой на π радиан, также получит максимально возможную оценку. Это возможно потому, что данные меры в силу своего изначально общего предназначения никак не учитывают информацию о попарном соответствии пикселей (с точностью до дискретизации) с двух изображений, тогда как для известных проективных преобразований такая информация имеется.

В данной работе, напротив, в рамках задачи проективного исправления, сфотографированного с произвольного ракурса прямоугольного документа, предлагается для каждой точки документа рассматривать её смещение — расстояние между её истинным и ожидаемым в соответствии с ответом детектора положением. Тогда в качестве меры ошибки работы алгоритма предлагается максимальное смещение по всему документу.

Предложенная мера ошибки, во-первых, лишена вышеописанного недостатка других мер. Во-вторых, она обосновывается своей проблемно-ориентированностью, так как вероятность сбоя дальнейших систем распознавания сильно зависит от того, насколько далеко от ожидаемого положения находится изображение распознаваемых объектов (например, букв или цифр) на самом деле.

2. Переход в координаты исправленного изображения

Все вышеописанные меры ошибки не инварианты к проективному преобразованию, более того — метрика Хаусдорфа инварианта только к группе движений. Таким образом, их значение будет разным в зависимости, например, от ракурса съёмки даже в идеальном случае, если бы алгоритм был к ракурсу съёмки полностью инвариантен. Это приводит к отсутствию осмысленности использования их в чистом виде как мер оценки качества работы алгоритмов проективного исправления.

С другой стороны, выходом алгоритма является проективно исправленное изображение, и именно с ним, как правило, происходит дальнейшая работа (см. рисунок 1 и рисунок 2). Поэтому все меры ошибки принято применять в координатах, связанных с исправленным изображением.

Заметим, кроме того, что в таких координатах любые меры ошибки становятся инвариантными к смене ракурса съёмки.

Теперь опишем аналитически переход в координаты, связанные с исправленным изображением. Пусть на плоскости фотографии детектор обнаружил прямоугольный документ, заданный упорядоченным набором своих углов, которые мы зададим однородными координатами (тут и далее однородные координаты мы будем обозначать двойными прямыми



а)

б)

Рисунок 1. Пример идеального исправления проективного преобразования. Белой рамкой показан истинный четырёхугольник, чёрной - задетектированный. а) Исходная фотография документа. б) Исправленная фотография. Задетектированный четырёхугольник перешёл в прямоугольник, параллельный осям координат.



а)

б)

Рисунок 2. Пример неудачного исправления проективного преобразования. Белой рамкой показан истинный четырёхугольник, чёрной - задетектированный. а) Исходная фотография документа. б) Исправленная фотография. Задетектированный четырёхугольник перешёл в прямоугольник, параллельный осям координат.

скобками $\|\bullet\|$):

$$p_i = \left\| \begin{pmatrix} p_{x,i} \\ p_{y,i} \\ 1 \end{pmatrix} \right\|, \quad i \in \{1, 2, 3, 4\}.$$

На той же плоскости размечены истинные углы документа:

$$\mathbf{p}'_i = \begin{Bmatrix} p'_{x,i} \\ p'_{y,i} \\ 1 \end{Bmatrix}, \quad i \in \{1, 2, 3, 4\}.$$

Тогда можно найти [13] матрицу проективного преобразования P , которая переводит задектированные углы документа в истинные:

$$\mathbf{p}'_i = P\mathbf{p}_i, \quad i \in \{1, 2, 3, 4\}.$$

Далее фотография проективно исправляется таким образом, чтобы перевести задектированный четырёхугольник документа в параллельный координатным осям прямоугольник с истинным соотношением сторон. Более подробно, пусть нам известно, что высота документа в e раз больше его ширины. Тогда углы \mathbf{p}_i перейдут в точки с координатами:

$$\mathbf{q}_1 = \begin{Bmatrix} 0 \\ 0 \\ 2(1+e) \end{Bmatrix}, \quad \mathbf{q}_2 = \begin{Bmatrix} 1 \\ 0 \\ 2(1+e) \end{Bmatrix}, \quad \mathbf{q}_3 = \begin{Bmatrix} 1 \\ e \\ 2(1+e) \end{Bmatrix}, \quad \mathbf{q}_4 = \begin{Bmatrix} 0 \\ e \\ 2(1+e) \end{Bmatrix}.$$

Данная система координат выбрана так, чтобы единицей длины являлся периметр документа, т.к. это гарантирует инвариантность меры ошибки к количеству пикселей изображения. Тогда можно найти матрицу проективного преобразования H отображающего точки из исходного изображения в исправленное. Такая матрица должна удовлетворять условиям:

$$\mathbf{q}_i = H\mathbf{p}_i, \quad i \in \{1, 2, 3, 4\}.$$

При этом истинные углы документа на исходной фотографии, как и все точки, отобразятся в исправленное изображение по такому же закону:

$$\mathbf{q}'_i = H\mathbf{p}'_i, \quad i \in \{1, 2, 3, 4\}.$$

Теперь найдём матрицу проективного преобразования $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, которая переводит задектированные углы документа в истинные на преобразованном изображении $\mathbf{q}'_i = A\mathbf{q}_i$:

$$\begin{aligned} \mathbf{q}'_i &= H\mathbf{p}'_i, & \mathbf{p}'_i &= P\mathbf{p}_i, & \mathbf{p}_i &= H^{-1}\mathbf{q}_i, \\ \mathbf{q}'_i &= HPH^{-1}\mathbf{q}_i, \\ A &= HPH^{-1}. \end{aligned}$$

Это преобразование определено для всех точек исправленного изображения

$$\mathbf{q}' = A\mathbf{q}$$

и описывает переход точек документа из их ожидаемых положений в истинные.

Перейдём в декартовы координаты:

$$\begin{vmatrix} x \\ y \end{vmatrix} = \frac{1}{q_3} \begin{vmatrix} q_1 \\ q_2 \end{vmatrix}, \quad \begin{vmatrix} x' \\ y' \end{vmatrix} = \frac{1}{q'_3} \begin{vmatrix} q'_1 \\ q'_2 \end{vmatrix},$$

тогда данное преобразование запишется следующим образом:

$$\begin{aligned} x'(x, y) &= \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}}, \\ y'(x, y) &= \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}}. \end{aligned} \quad (1)$$

3. Постановка задачи

Векторное поле [14] (см. рисунок 3)

$$\mathbf{s}_A(x, y) = \begin{vmatrix} x'(x, y) \\ y'(x, y) \end{vmatrix} - \begin{vmatrix} x \\ y \end{vmatrix}$$

соответствует проективному преобразованию с матрицей A .

Теперь смещение легко определяется как модуль векторного поля \mathbf{s} :

$$u_A(x, y) = \|\mathbf{s}_A(x, y)\|_2 = \left\| \begin{vmatrix} x'(x, y) \\ y'(x, y) \end{vmatrix} - \begin{vmatrix} x \\ y \end{vmatrix} \right\|_2. \quad (2)$$

Обозначим $\mathbf{r} = \begin{vmatrix} x \\ y \end{vmatrix}$ и $u_A(\mathbf{r}) = u_A(x, y)$. Задетектированный прямоугольник документа на исправленном изображении обозначим как $Q(e)$:

$$Q(e) = \left\{ \mathbf{r} : x \geq 0, \quad y \geq 0, \quad x \leq \frac{1}{2(1+e)}, \quad y \leq \frac{e}{2(1+e)} \right\}.$$

Тогда максимальное смещение определяется как

$$U(A, e) = \max_{\mathbf{r} \in Q(e)} u_A(\mathbf{r}), \quad (3)$$

а задача состоит в его аналитическом вычислении.

4. Опровержение гипотезы о достижении максимума на углах

Интуитивно возникает гипотеза о том, что максимальное смещение всегда достигается на одном из четырёх углов документа. Проверка этой гипотезы дала отрицательный результат, см. рисунок 3.

5. Достижение максимума на границе

Для решение задачи 3 распишем 2 в соответствии с 1:

$$u_A(x, y) = \sqrt{\left(\frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}} - x \right)^2 + \left(\frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}} - y \right)^2}. \quad (4)$$

Для удобства будем временно вместо смещения рассматривать его квадрат:

$$u_A^2(x, y) = \left(\frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}} - x \right)^2 + \left(\frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}} - y \right)^2.$$

Рассмотрим два возможных случая:

- Пусть наше проективное преобразование — аффинное, т.е. $a_{31}^2 + a_{32}^2 = 0$. Тогда квадрат смещения запишется следующим образом:

$$u_A^2(x, y) = \left(\frac{a_{11}x + a_{12}y + a_{13}}{a_{33}} - x \right)^2 + \left(\frac{a_{21}x + a_{22}y + a_{23}}{a_{33}} - y \right)^2.$$

Тогда он является положительным параболоидом и не имеет локальных максимумов.

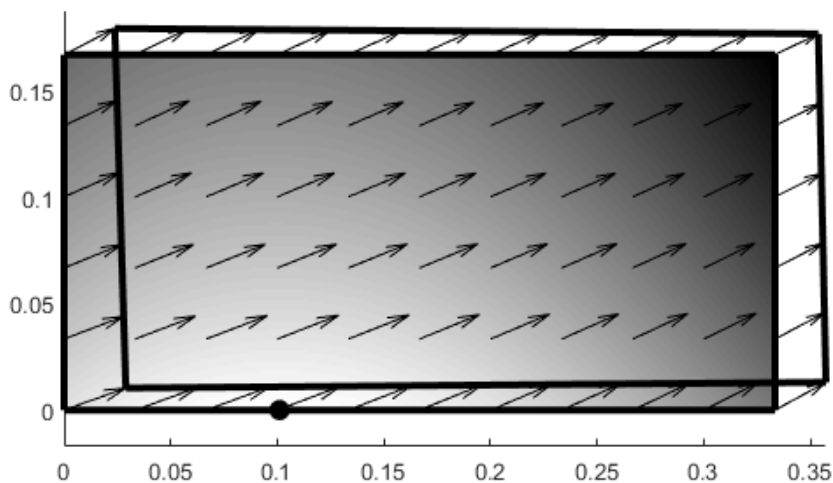


Рисунок 3. Контрпример для гипотезы о том, что максимальное смещение всегда достигается на одном из углов документа. Залитый прямоугольник соответствует ожидаемому положению документа на исправленном изображении ($e = \frac{1}{2}$), четырёхугольник — истинному, векторное поле иллюстрирует проективное преобразование с матрицей A , длина векторов и яркость заливки соответствуют величине смещения, чёрная точка показывает место максимального смещения и не совпадает ни с одним из углов.

- Пусть проективное преобразование не является аффинным, т.е. $a_{31}^2 + a_{32}^2 > 0$. Тогда прямая с уравнением $a_{31}x + a_{32}y + a_{33} = 0$ не бесконечно удалена и на ней функция $u_A^2(x, y)$ достигает своего глобального максимума со значением $+\infty$, либо является неопределённой (в точках своего пересечения с прямыми $a_{11}x + a_{12}y + a_{13} = 0$ и $a_{21}x + a_{22}y + a_{23} = 0$). Доопределим эти точки значением $+\infty$. Докажем, что у функции $u_A^2(x, y)$ вне прямой $a_{31}x + a_{32}y + a_{33} = 0$ отсутствуют локальные максимумы. Возьмём любую точку $\mathbf{r} = \begin{pmatrix} x \\ y \end{pmatrix}$ не на прямой $a_{31}x + a_{32}y + a_{33} = 0$ и рассмотрим прямую вида $a_{31}x + a_{32}y + a_{33} = \text{const}$, которая через точку \mathbf{r} проходит (тогда $\text{const} \neq 0$). Тогда на этой прямой квадрат смещения запишется следующим образом:

$$u_A^2(x, y) = \left(\frac{a_{11}x + a_{12}y + a_{13}}{\text{const}} - x \right)^2 + \left(\frac{a_{21}x + a_{22}y + a_{23}}{\text{const}} - y \right)^2.$$

Из последней формулы видно, что смещение на вышеописанной прямой является параболой с ветвями вверх, следовательно, не имеет локальных максимумов. Отсюда следует, что и точка \mathbf{r} не является локальным максимумом $u_A^2(x, y)$, ч.т.д.

Мы получили, что функция $u_A^2(x, y)$ в любом случае не имеет локальных максимумов вне прямой $a_{31}x + a_{32}y + a_{33} = 0$, а на ней равна $+\infty$. Теперь снова рассмотрим два случая:

- Пусть прямая $a_{31}x + a_{32}y + a_{33} = 0$ пересекает множество $Q(e)$. Тогда она пересекает и его границу, следовательно, на границе есть значение $+\infty$, поэтому максимум на границе достигается.
- Пусть прямая $a_{31}x + a_{32}y + a_{33} = 0$ не пересекает множество $Q(e)$. Тогда в нём нет локальных максимумов, тогда максимум достигается только на границе $Q(e)$.

Мы получили, что в любом случае максимум $u_A^2(x, y)$ на границе множества $Q(e)$ достигается. Это же доказательство применимо для любого ограниченного множества M .

6. Поиск решения на границе

Пусть граница нашего множества состоит из отрезков. Будем искать максимальное значение $u_A^2(x, y)$ на каждом отрезке отдельно, а потом выберем наибольшее.

Внутри каждого отрезка максимум достигается либо на его концах, либо в локальном максимуме на отрезке. Найдём локальные максимумы на отрезке. Для этого параметризуем отрезок с концами $\mathbf{r}_1 = \begin{vmatrix} x_1 \\ y_1 \end{vmatrix}$ и $\mathbf{r}_2 = \begin{vmatrix} x_2 \\ y_2 \end{vmatrix}$ следующим образом:

$$\mathbf{r}(t) = \begin{vmatrix} x(t) \\ y(t) \end{vmatrix} = (\mathbf{r}_2 - \mathbf{r}_1)t + \mathbf{r}_1, \quad t \in [0, 1]. \quad (5)$$

Тогда на отрезке квадрат смещения будет выглядеть следующим образом:

$$\begin{aligned} u_A^2(t) &= \left(\frac{a_{11}x(t) + a_{12}y(t) + a_{13}}{a_{31}x(t) + a_{32}y(t) + a_{33}} - x(t) \right)^2 + \left(\frac{a_{21}x(t) + a_{22}y(t) + a_{23}}{a_{31}x(t) + a_{32}y(t) + a_{33}} - y(t) \right)^2 = \\ &= \left(\frac{a_{11}((x_2 - x_1)t + x_1) + a_{12}((y_2 - y_1)t + y_1) + a_{13}}{a_{31}((x_2 - x_1)t + x_1) + a_{32}((y_2 - y_1)t + y_1) + a_{33}} - ((x_2 - x_1)t + x_1) \right)^2 + \\ &+ \left(\frac{a_{21}((x_2 - x_1)t + x_1) + a_{22}((y_2 - y_1)t + y_1) + a_{23}}{a_{31}((x_2 - x_1)t + x_1) + a_{32}((y_2 - y_1)t + y_1) + a_{33}} - ((y_2 - y_1)t + y_1) \right)^2. \end{aligned}$$

После замен $\Delta x = x_2 - x_1$, $\Delta y = y_2 - y_1$ получим:

$$\begin{aligned} u_A^2(t) &= \left(\frac{a_{11}(\Delta x t + x_1) + a_{12}(\Delta y t + y_1) + a_{13}}{a_{31}(\Delta x t + x_1) + a_{32}(\Delta y t + y_1) + a_{33}} - (\Delta x t + x_1) \right)^2 + \\ &+ \left(\frac{a_{21}(\Delta x t + x_1) + a_{22}(\Delta y t + y_1) + a_{23}}{a_{31}(\Delta x t + x_1) + a_{32}(\Delta y t + y_1) + a_{33}} - (\Delta y t + y_1) \right)^2 = \\ &= \left(\frac{[a_{11}\Delta x + a_{12}\Delta y]t + [a_{11}x_1 + a_{12}y_1 + a_{13}]}{[a_{31}\Delta x + a_{32}\Delta y]t + [a_{31}x_1 + a_{32}y_1 + a_{33}]} - (\Delta x t + x_1) \right)^2 + \\ &+ \left(\frac{[a_{21}\Delta x + a_{22}\Delta y]t + [a_{21}x_1 + a_{22}y_1 + a_{23}]}{[a_{31}\Delta x + a_{32}\Delta y]t + [a_{31}x_1 + a_{32}y_1 + a_{33}]} - (\Delta y t + y_1) \right)^2. \end{aligned}$$

После замен

$$\begin{aligned} k_x &= a_{11}\Delta x + a_{12}\Delta y, & b_x &= a_{11}x_1 + a_{12}y_1 + a_{13}, \\ k_y &= a_{21}\Delta x + a_{22}\Delta y, & b_y &= a_{21}x_1 + a_{22}y_1 + a_{23}, \\ k_z &= a_{31}\Delta x + a_{32}\Delta y, & b_z &= a_{31}x_1 + a_{32}y_1 + a_{33} \end{aligned}$$

получим:

$$u_A^2(t) = \left(\frac{k_x t + b_x}{k_z t + b_z} - (\Delta x t + x_1) \right)^2 + \left(\frac{k_y t + b_y}{k_z t + b_z} - (\Delta y t + y_1) \right)^2.$$

Приравняем к нулю производную:

$$\begin{aligned} \frac{du_A^2(t)}{dt} &= 0, \\ \left(\frac{k_x t + b_x}{k_z t + b_z} - (\Delta x t + x_1) \right) &\left(\frac{k_x b_z - k_z b_x}{(k_z t + b_z)^2} - \Delta x \right) + \end{aligned}$$

$$\left(\frac{k_y t + b_y}{k_z t + b_z} - (\Delta y t + y_1)\right) \left(\frac{k_y b_z - k_z b_y}{(k_z t + b_z)^2} - \Delta y\right) = 0,$$

$$\left(k_z \Delta x t^2 + (k_z x_1 + b_z \Delta x - k_x) t + (b_z x_1 - b_x)\right) \left(k_z^2 \Delta x t^2 + 2k_z b_z \Delta x t + (k_z b_x - k_x b_z + b_z^2 \Delta x)\right) +$$

$$\left(k_z \Delta y t^2 + (k_z y_1 + b_z \Delta y - k_y) t + (b_z y_1 - b_y)\right) \left(k_z^2 \Delta y t^2 + 2k_z b_z \Delta y t + (k_z b_y - k_y b_z + b_z^2 \Delta y)\right) = 0,$$

после замен

$$\begin{aligned} h_{x1} &= k_z \Delta x, & h_{x2} &= k_z x_1 + b_z \Delta x - k_x, & h_{x3} &= b_z x_1 - b_x, \\ h_{y1} &= k_z \Delta y, & h_{y2} &= k_z y_1 + b_z \Delta y - k_y, & h_{y3} &= b_z y_1 - b_y, \\ h_{x4} &= k_z^2 \Delta x, & h_{x5} &= 2k_z b_z \Delta x, & h_{x6} &= k_z b_x - k_x b_z + b_z^2 \Delta x, \\ h_{y4} &= k_z^2 \Delta y, & h_{y5} &= 2k_z b_z \Delta y, & h_{y6} &= k_z b_y - k_y b_z + b_z^2 \Delta y. \end{aligned}$$

получим:

$$\left(h_{x1} t^2 + h_{x2} t + h_{x3}\right) \left(h_{x4} t^2 + h_{x5} t + h_{x6}\right) +$$

$$\left(h_{y1} t^2 + h_{y2} t + h_{y3}\right) \left(h_{y4} t^2 + h_{y5} t + h_{y6}\right) = 0,$$

после замен

$$\begin{aligned} c_0 &= h_{x3} h_{x6} + h_{y3} h_{y6}, \\ c_1 &= h_{x2} h_{x6} + h_{x3} h_{x5} + h_{y2} h_{y6} + h_{y3} h_{y5}, \\ c_2 &= h_{x1} h_{x6} + h_{x2} h_{x5} + h_{x3} h_{x4} + h_{y1} h_{y6} + h_{y2} h_{y5} + h_{y3} h_{y4}, \\ c_3 &= h_{x1} h_{x5} + h_{x2} h_{x4} + h_{y1} h_{y5} + h_{y2} h_{y4}, \\ c_4 &= h_{x1} h_{x4} + h_{y1} h_{y4}. \end{aligned}$$

получим:

$$c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 = 0. \quad (6)$$

Из вещественных корней уравнения 6 следует рассмотреть только те t , которые удовлетворяют вышеприведённому ограничению: $t \in [0, 1]$. Далее следует подставить их в формулу 5, после чего результат окончательно подставить в формулу 4 для определения величины смещения.

Таким образом, решена задача аналитического поиска максимального смещения.

7. Заключение

В работе в качестве меры ошибки детектирования прямоугольного документа на сделанной с произвольного ракурса фотографии предложено использовать максимальное смещение среди точек документа. Опровергнута гипотеза о том, что максимальное смещение всегда достигается на углах документа. Доказано, что максимальное смещение всегда достигается на границе документа. Приведён аналитический способ нахождения максимального смещения на границе документа.

8. Благодарности

Исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов №17-29-03514 и №15-29-06086.

9. Литература

- [1] Orrite, C. Shape matching of partially occluded curves invariant under projective transformation / C. Orrite, J.E. Herrero // *Computer Vision and Image Understanding*. — 2004. — Vol. 93(1). — P. 34–64.
- [2] Zhukovsky, A.E. Segments graph-based approach for smartphone document capture / A.E. Zhukovsky, V.V. Arlazarov, V.V. Postnikov, V.E. Krivtsov // *Eighth International Conference on Machine Vision (ICMV 2015)*. — International Society for Optics and Photonics. — 2015. — Vol. 9875. — P. 98750P.
- [3] Povolotskiy, M. Russian license plate segmentation based on dynamic time warping / M. Povolotskiy, E. Kuznetsova, T. Khanipov // *Proceedings - 31st European Conference on Modelling and Simulation (ECMS)*. — 2017. — P. 285–291.
- [4] Чернов, Т. Исследование методов сегментации изображений текстовых блоков документов с помощью алгоритмов структурного анализа и машинного обучения / Т. Чернов, Д. Ильин, П. Безматерных, И. Фараджев, С. Карпенко // *Вестник РФФИ*. — 2016. — Т. 92, № 4. — С. 55–71.
- [5] Шемякина Ю. Исследование алгоритмов вычисления проективного преобразования в задаче наведения на планарный объект по особым точкам / Ю. Шемякина, А. Жуковский, И. Фараджев // *Искусственный интеллект и принятие решений*. — 2017. — № 1. — С. 43–49.
- [6] Skoryukina, N. Screenshot: TV-stream frame search with projectively distorted and noisy query / N. Skoryukina, T. Chernov, K. Bulatov, D. Nikolaev, V.L. Arlazarov // *Proc. SPIE 10341, Ninth International Conference on Machine Vision (ICMV), 103410Y* — 2016. — P. 1–5.
- [7] Dubuisson, M. P. A modified Hausdorff distance for object matching / M.P. Dubuisson, A.K. Jain // *Pattern Recognition, 1994. Vol. 1-Conferece A: Computer Vision & Image Processing., Proceedings of the 12-th IAPR International Conference on* — 1994. — Vol. 1. — P. 566–568.
- [8] Huttenlocher, D. P. Comparing images using the Hausdorff distance / D.P. Huttenlocher, G.A. Klanderman, W.J. Rucklidge // *IEEE Transactions on pattern analysis and machine intelligence*. — 1993. — Vol. 15(9). - P. 850–863.
- [9] Sim, D.G. Object matching algorithms using robust Hausdorff distance measures / D.G. Sim, O.K. Kwon, R.H. Park // *IEEE Transactions on image processing*. — 1999. — Vol. 8(3). — P. 425–429.
- [10] Jesorsky, O. Robust face detection using the hausdorff distance / O. Jesorsky, K.J. Kirchberg, R.W. Frischholz // *International Conference on Audio-and Video-Based Biometric Person Authentication*. — Springer, Berlin, Heidelberg. — 2001. — P. 90–95.
- [11] Jaccard, P. Étude comparative de la distribution florale dans une portion des Alpes et des Jura / P. Jaccard // *Bull Soc Vaudoise Sci Nat*. — 1901. — Vol. 37. — P. 547–579.
- [12] ICDAR 2017 POD Competition Web Site [Electronic resource]. — Access mode: http://www.icst.pku.edu.cn/cdp/ICDAR2017_PODCompetition/index.html (19.11.2017).
- [13] Hartley, R. Multiple view geometry in computer vision / R. Hartley, A. Zisserman — Cambridge university press, 2003. — 655 p.
- [14] Красносельский М.А. Векторные поля на плоскости / М.А. Красносельский, А.И. Перов, А.И. Поволоцкий, П. П. Забрейко — Москва: Государственное издательство физико-математической литературы, 1963. — 248 с.

Error values analysis for inaccurate projective transformation of a quadrangle

I.A. Konovalenko^{1,2,3}, J.A. Shemiakina^{3,4}

¹Institute for Information Transmission Problems of Russian Academy of Sciences (Kharkevich Institute) — IITP RAS, Bolshoy Karetny per. 19, build.1, Moscow, Russia, 127051

²Moscow Institute of Physics and Technology (State University) — MIPT, Institutsky pereulok, 9, Dolgoprudny, Moscow Region, Russia, 141701

³Smart Engines Ltd., pr. 60-letiya Oktyabrya, 9, Moscow, Russia, 117312

⁴Institute for System Analysis of Federal Research Center «Computer Science and Control» of Russian Academy of Sciences — ISA FRC CSC RAS, pr. 60-letiya Oktyabrya, 9, Moscow, Russia, 117312

Abstract. This paper describes method for measure of quality of detections rectangular document in the photo taken from an arbitrary angle. Let there are document's boundaries which are a true quadrangle and a corresponding quadrangle which was detected in the photo. In this case these quadrangles are connected by a known projective transformation, the smallness of which characterizes the quality of the detector operation. The error value for each point of the document is described by its bias, i.e. the distance between its true and expected position in the coordinates associated with the detected quadrilateral. It is suggested to take the maximum bias of a point of the document as a measure of the error of the detection. In this paper, a counterexample is given for the hypothesis that the maximum bias is achieved at the corner of the document; it is proved that the maximum bias is achieved at the boundary of the document; an analytical solution is given to the problem of finding the maximum bias at the boundary of the document.

Keywords: projective transformation, document recognition, rectangular detection, detection quality measure, maximum bias.