Error values analysis for inaccurate projective transformation of a quadrangle

I A Konovalenko¹,²,³ and J A Shemiakina⁴

¹Institute for Information Transmission Problems of Russian Academy of Sciences (Kharkevich Institute), Bolshoy Karetny per. 19, build.1, Moscow, Russia, 127051
²Moscow Institute of Physics and Technology (State University), Institutsky per. 9
³Dolgoprudny, Moscow Region, Russia, 141701
⁴Smart Engines Ltd., pr. 60-letiya Oktyabrya 9, Moscow, Russia, 117312
⁵Institute for System Analysis of Federal Research Center «Computer Science and Control» of Russian Academy of Sciences, pr. 60-letiya Oktyabrya, 9, Moscow, Russia, 117312

e-mail: konovalenko@iitp.ru

Abstract. The paper introduces a quality metric for a projective correction of a document photograph obtained from any viewpoint. For the quantitative quality measurement of the projective transform every document point is assigned a specific shift - the distance between the ideal and real positions of this point on the corrected image. We chosen the maximal shift throughout the document as the overall quality measure. We contradict a hypothesis that the maximal shift is obtained in the document corners. A proof is given that the maximal shift occurs on a document border and provide the analytical solution to the problem of locating it.

1. Introduction

Image projective transformation naturally arises in computer vision when observing planar scenes by a pinhole camera model from arbitrary viewpoints. For example, [1] and [2] suggest algorithms for projective matching of contours and straight segments resp. The paper [3] describes segmentation of projectively corrected vehicle number plates and [4] - passports segmentation. Papers [5] and [6] analyze approaches to solving the problem of projective aiming in the example of insurance documents. Works [7] [8] performs a review of document projective aiming algorithms. Paper [9] describes the “Snapscreen” system which detects and projectively corrects a TV screenshot and further recognizes a TV program. In [10] [11] the problem of projective correction arises from the need of terrain mapping by UAV camera shots. In these works, a camera with a corrected radial distortion [12] is supposed to be used.

There is a need to compare the quality of various projective correction algorithms. Usually, for that purpose the true projective transform is given for some set of samples. After that the task is reduced to computing the similarity between the experimental and true projective transformations. However, in generic case projective transformation has 8 parameters and defining a scalar similarity measure can be performed in a variety of ways.
In published works the Hausdorff metric [13] or its many variations [14] are used for that purpose. It is defined for sets and equals to the largest distance between the points of one set and the corresponding to the nearest points of another set. Usually in computer vision the sets represent distorted contours of the same object. For example, in [15] the Hausdorff metric is used for matching objects and in [16] – for faces detection.

Another widely used in practice instrument is the Jaccard index [17]. It is also defined for sets and equals the intersection area relatively to the union area. This index is a target metric, for example, in the competition «ICDAR» od documents analysis and recognition [18].

An obvious advantage of these error measures is that they give its best possible value to the ideal algorithm answer. The reverse however is not true. Indeed, suppose we need to detect and projectively correct a credit card photograph. Then the algorithm based on the measures above which will find the card rotated by \( \pi \) rads, will also gain the maximal possible score. This is possible due to the fact that these measures have generic purpose and thus do not account for information about pixel pairwise mapping, while for the known projective transformations such information is known.

We, on the contrary, first of all suggest to define the projective correction quality for each document point. The quantitative measure of such quality would naturally be represented by the shift – the distance between the ideal and real positions of a given point on the corrected image. In this case we suggest to use the maximal shift on the document as an overall projective correction measure. The proposed error measure does not have the disadvantage of other measures mentioned above.

2. Transitioning to the coordinates of the transformed image

We first describe the transitioning to the coordinates of the corrected image. Suppose a rectangular document was detected in the photograph plane and it is represented by an ordered set of its corners which we list by homogeneous coordinates (here and further homogeneous coordinates are designated using double square brackets \([\bullet]\)):

\[
P_i = \begin{bmatrix} p_{x,i} \\ p_{y,i} \\ 1 \end{bmatrix}, \quad i \in \{1, 2, 3, 4\}.
\]

True document corners are marked in the same plane:

\[
P_i' = \begin{bmatrix} p'_{x,i} \\ p'_{y,i} \\ 1 \end{bmatrix}, \quad i \in \{1, 2, 3, 4\}.
\]

We can thus find [19] the matrix of the projective transformation \( P \) which maps the coordinates of true and detected document corners:

\[
p_i' = P p_i, \quad i \in \{1, 2, 3, 4\}.
\]

The photography is then projectively corrected so that the detected document quadrangle is mapped to the rectangle with axes parallel to the coordinate axes and a known true ratio (see fig. 1 and fig. 2). Suppose we know that the document height is \( e \) times greater than its width. In this case corners \( p_i \) will be transformed to the points with coordinates:

\[
q_1 = \begin{bmatrix} 0 \\ 0 \\ 2(1 + e) \end{bmatrix}, \quad q_2 = \begin{bmatrix} 1 \\ 0 \\ 2(1 + e) \end{bmatrix}, \quad q_3 = \begin{bmatrix} 1 \\ e \\ 2(1 + e) \end{bmatrix}, \quad q_4 = \begin{bmatrix} 0 \\ e \\ 2(1 + e) \end{bmatrix}.
\]
Figure 1. Ideal projective transformation example. The true quadrangle is marked white, while the detected one – as black. (a) Initial document photograph. (b) Corrected photograph. The detected quadrangle is transformed to the rectangle parallel to the coordinate axes.

Figure 2. Example of bad projective correction. The true quadrangle is marked white, while the detected one – as black. (a) Initial document photograph. (b) Corrected photograph. The detected quadrangle is transformed to the rectangle parallel to the coordinate axes.

This coordinate frame is chosen in such a manner that the unit length is the document perimeter because it guarantees that the error measure is invariant to the image pixels count. We may then find the homography matrix $H$ which maps source image points to the corrected points. Such
matrix should satisfy:
\[ q_i = H p_i, \quad i \in \{1, 2, 3, 4\}. \]
The true document corners on the source photograph, like other points, would map to the transformed image by the same law:
\[ q'_i = H p'_i, \quad i \in \{1, 2, 3, 4\}. \]
Now find the homography matrix
\[ A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}, \]
which transforms the detected document corners to the true ones on the transformed image
\[ q'_i = A q_i, \quad p'_i = P p_i, \quad p_i = H^{-1} q_i, \]
\[ q'_i = H P H^{-1} q_i, \quad A = H P H^{-1}. \]
This transformation is defined for all points of the transformed image
\[ q' = A q \]
and describes the transitioning of document points from their expected positions to the true ones.

Use the Cartesian coordinates:
\[
\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1/q_3 \\ q_2 \\ q_1 \end{bmatrix}, \quad \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1/q_3 \\ q_2 \\ q_1 \end{bmatrix},
\]
the given transformation is then written as:
\[
x'(x, y) = \frac{a_{11}x + a_{12}y + a_{13}}{a_{31}x + a_{32}y + a_{33}},
\]
\[
y'(x, y) = \frac{a_{21}x + a_{22}y + a_{23}}{a_{31}x + a_{32}y + a_{33}}. \tag{1}
\]

3. Problem statement

Vector field [20]
\[ s_A(x, y) = \begin{bmatrix} x'(x, y) \\ y'(x, y) \end{bmatrix} - \begin{bmatrix} x \\ y \end{bmatrix} \]
corresponds to the projective transformation with matrix \( A \).

The shift is now defined as the second norm of the vector field \( s \):
\[ u_A(x, y) = || s_A(x, y) ||_2 = \left\| \begin{bmatrix} x'(x, y) \\ y'(x, y) \end{bmatrix} - \begin{bmatrix} x \\ y \end{bmatrix} \right\|_2. \tag{2} \]
Denote \( r = \begin{bmatrix} x \\ y \end{bmatrix} \) u \( u_A(r) = u_A(x, y) \). The detected document rectangle on the corrected image is designated as \( Q(e) \):
\[ Q(e) = \left\{ r : x \geq 0, \quad y \geq 0, \quad x \leq \frac{1}{2(1+e)}, \quad y \leq \frac{e}{2(1+e)} \right\}. \]
The maximal shift is then defined as
\[ U(A, e) = \max_{r \in Q(e)} u_A(r), \tag{3} \]
and we need to compute it analytically.
4. Contradiction of the corner maximum hypothesis
Intuitively one might deduce a hypothesis that the maximal shift is always obtained on one of the four document corners. The verification of this hypothesis gave a negative result, see fig. 3.

![Figure 3](image.png)

**Figure 3.** A counterexample for the hypothesis that the maximal shift always lies on one of the document corners.

5. Maximum on the border
We now prove that the maximal shift always lies on the document border. Expand 2 with respect to 1:

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

(4)

For convenience we will analyze the shift squared:

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

Consider two possible cases:
• Let our projective transformation be affine, i.e. \( a_{31}^2 + a_{32}^2 = 0 \). Then the squared shift is written as:

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

One can see that \( u^2_A(x, y) \) is a positively defined quadratic form and hence does not have local maxima.

• Suppose the projective transformation is not affine, i.e. \( a_{31}^2 + a_{32}^2 > 0 \). Consider then a line with equation \( a_{31}x + a_{32}y + a_{33} = 0 \), which we shall call a horizon. The function \( u^2_A(x, y) \) reaches its global maximum on this line with the value +\( \infty \). Let us prove that the function \( u^2_A(x, y) \) does not have local maxima outside the horizon. Take any point \( r = \left[ \begin{array}{c} x \\ y \end{array} \right] \) outside horizon and consider the line \( a_{31}x + a_{32}y + a_{33} = \text{const} \neq 0 \) parallel to the horizon, which passes through the point \( r \). The squared shift on this line is then written as:

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

From this formula one can see that the shift of the line described above is a positive parabola, hence, it does not have local maxima. In particular, the point \( r \) is not a local maximum on this line. It then follows that the point \( r \) is not a local maximum for \( u^2_A(x, y) \), q.e.d.

We deduced that the function \( u^2_A(x, y) \) has no local maxima outside the horizon and equals +\( \infty \) there. Again consider two cases:

• Let horizon intersect the set \( Q(e) \). Then it intersects its border, hence, the border contains the +\( \infty \) value and the border maximum is reached.

• Suppose the horizon does not intersect the set \( Q(e) \). Then the \( Q(e) \) set has no local maxima and the maximum is reached on the \( Q(e) \) border.

We proved that in any case the maximum of \( u^2_A(x, y) \) is reached on the \( Q(e) \) border. The same proof works for any bound set \( M \).

6. Finding the solution on the border

Suppose the border of our set consists of straight segments. We will search for the maximal \( u^2_A(x, y) \) value on each segment separately and will then choose the biggest one.

In each segment the maximum is reached either on its ends or in the local maximum on the segment. Find the local maxima on the segment. For that purpose we parametrize the segment with ends \( r_1 = \left[ \begin{array}{c} x_1 \\ y_1 \end{array} \right] \) and \( r_2 = \left[ \begin{array}{c} x_2 \\ y_2 \end{array} \right] \) in the following way:

\[
r(t) = \left[ \begin{array}{c} x(t) \\ y(t) \end{array} \right] = (r_2 - r_1)t + r_1, \quad t \in [0, 1].
\]

The shift square on the segment is then represented as:

\[
u^2_A(t) = \left( \frac{a_{11}x(t) + a_{12}y(t) + a_{13}}{a_{31}x(t) + a_{32}y(t) + a_{33}} - x(t) \right)^2 + \left( \frac{a_{21}x(t) + a_{22}y(t) + a_{23}}{a_{31}x(t) + a_{32}y(t) + a_{33}} - y(t) \right)^2 = \left( \frac{a_{11}(x_2 - x_1)t + x_1 + a_{12}(y_2 - y_1)t + y_1 + a_{13}}{a_{31}(x_2 - x_1)t + x_1 + a_{32}(y_2 - y_1)t + y_1 + a_{33}} - (x_2 - x_1)t + x_1 \right)^2 + \]

\[
\left( \frac{a_{21}(x_2 - x_1)t + x_1 + a_{22}(y_2 - y_1)t + y_1 + a_{23}}{a_{31}(x_2 - x_1)t + x_1 + a_{32}(y_2 - y_1)t + y_1 + a_{33}} - (x_2 - x_1)t + x_1 \right)^2 +
\]
\[
\left(\frac{a_{21}((y_2 - y_1) t + y_1) + a_{23}}{a_{31}((x_2 - x_1) t + x_1) + a_{32}((y_2 - y_1) t + y_1) + a_{33}} - ((y_2 - y_1) t + y_1)\right)^2.
\]

After substitutions \(\Delta x = x_2 - x_1, \quad \Delta y = y_2 - y_1\) we get:

\[
u_A^2(t) = \left(\frac{a_{11}(\Delta x t + x_1) + a_{12}(\Delta y t + y_1) + a_{13}}{a_{31}(\Delta x t + x_1) + a_{32}(\Delta y t + y_1) + a_{33}} - (\Delta x t + x_1)\right)^2 + \]

\[
\left(\frac{a_{21}(\Delta x t + x_1) + a_{22}(\Delta y t + y_1) + a_{23}}{a_{31}(\Delta x t + x_1) + a_{32}(\Delta y t + y_1) + a_{33}} - (\Delta y t + y_1)\right)^2 =
\]

\[
\left(\frac{a_{11}\Delta x + a_{12}\Delta y t + [a_{11}x_1 + a_{12}y_1 + a_{13}]}{a_{31}\Delta x + a_{32}\Delta y t + [a_{31}x_1 + a_{32}y_1 + a_{33}]} - (\Delta x t + x_1)\right)^2 +
\]

\[
\left(\frac{a_{21}\Delta x + a_{22}\Delta y t + [a_{21}x_1 + a_{22}y_1 + a_{23}]}{a_{31}\Delta x + a_{32}\Delta y t + [a_{31}x_1 + a_{32}y_1 + a_{33}]} - (\Delta y t + y_1)\right)^2.
\]

After substitutions

\[
k_x = a_{11}\Delta x + a_{12}\Delta y, \quad b_x = a_{11}x_1 + a_{12}y_1 + a_{13}, \\
k_y = a_{21}\Delta x + a_{22}\Delta y, \quad b_y = a_{21}x_1 + a_{22}y_1 + a_{23}, \\
k_z = a_{31}\Delta x + a_{32}\Delta y, \quad b_z = a_{31}x_1 + a_{32}y_1 + a_{33}
\]

we get:

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

Let the derivative be zero:

\[
\frac{du_A^2(t)}{dt} = 0,
\]

\[
\left(\frac{k_x t + b_x}{k_z t + b_z} - (\Delta x t + x_1)\right) \left(\frac{k_z b_z - k_z b_x}{(k_z t + b_z)^2} - \Delta x\right) +
\]

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

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

\[
\left(k_z \Delta y t^2 + (k_z y_1 + b_z \Delta y - k_y) t + (b_z y_1 - b_y)\right) \left(k_z \Delta x t^2 + 2k_z b_z \Delta y t + (k_z b_y - k_y b_z + b_z^2 \Delta y)\right) = 0.
\]

after substitutions

\[
h_{x1} = k_z \Delta x, \quad h_{x2} = k_z x_1 + b_z \Delta x - k_x, \quad h_{x3} = b_z x_1 - b_x,
\]

\[
h_{y1} = k_z \Delta y, \quad h_{y2} = k_z y_1 + b_z \Delta y - k_y, \quad h_{y3} = b_z y_1 - b_y,
\]

\[
h_{x4} = k_z^2 \Delta x, \quad h_{x5} = 2k_z b_z \Delta x, \quad h_{x6} = k_z b_x - k_x b_z + b_z^2 \Delta x,
\]

\[
h_{y4} = k_z^2 \Delta y, \quad h_{y5} = 2k_z b_z \Delta y, \quad h_{y6} = k_z b_y - k_y b_z + b_z^2 \Delta y.
\]

we get:

\[
\left(h_{x1} t^2 + h_{x2} t + h_{x3}\right) \left(h_{x4} t^2 + h_{x5} t + h_{x6}\right) +
\]
\[
\left( h_y t^2 + h_y t + h_y \right) \left( h_y t^2 + h_y t + h_y \right) = 0,
\]

after substitutions
\[
c_0 = h_x h_x + h_y h_y,
\]
\[
c_1 = h_x h_x + h_y h_y + h_y h_y,
\]
\[
c_2 = h_x h_x + h_y h_y + h_y h_y + h_y h_y,
\]
\[
c_3 = h_x h_x + h_y h_y + h_y h_y + h_y h_y,
\]
\[
c_4 = h_x h_x + h_y h_y.
\]

we get:
\[
c_0 + c_1 t + c_2 t^2 + c_3 t^3 + c_4 t^4 = 0. \tag{6}
\]

Among the real solutions of equation 6 we should consider only those \( t \) which satisfy the above condition \( t \in [0, 1] \). They should further be substituted in 5, and then finally to substitute the result in 4 to compute the shift value.

The task of analytical search for the maximal shift is hence solved.

7. Conclusion
The paper proposes a quality measure for projective correction of a document photograph created from an arbitrary viewpoint. An effective algorithm for its calculation is suggested and its correctness is proven.

8. References
[1] Orrite C and Herrero J E 2004 Shape matching of partially occluded curves invariant under projective transformation Computer Vision and Image Understanding 93(1)34-64
[2] Zhukovsky A E, Arlazarov V V, Postnikov V V and Krivtsov V E 2015 Segments graph-based approach for smartphone document capture Eighth International Conference on Machine Vision (ICMV) 9875 98750P
[3] Povolotskiy M A, Kuznetsova E G and Khanipov T M 2017 Russian license plate segmentation based on dynamic time warping 31st European Conference on Modelling and Simulation (ECMS) 285-291
[4] Chernov T S, Il'in D A, Bezmaternikh P V, Faradjev I A and Karpenko S M 2016 Researcher methods for image segmentation of text blocks of documents using algorithms of structural analysis and machine learning Bulletin of the Russian Foundation for Basic Research 4 55-71
[5] Shemiakina J A, Cshukovskiy A E and Faradjev I A 2017 Investigation of algorithms for calculating the projective transformation in the problem of guidance to a planar object from singular points Artificial Intelligence and Decision Making 1 43-49
[6] Shemiakina J A 2017 Use points and lines to calculate the projective transformation from two images of a flat object Information technology and computer systems 3 79-91
[7] Bolotova Yu A, Spitsyn V G and Osina P M 2017 A review of algorithms for text detection in images and videos Computer Optics 41(3) 441-452 DOI: 10.18287/2412-6179-2017-41-3-441-452
[8] Skoryukina N, Shemiakina J, Arlazarov V L and Faradjev I 2018 Document localization algorithms based on feature points and straight lines Tenth International Conference on Machine Vision (ICMV) 1096 109691H
[9] Skoryukina N, Chernov T, Bulatov K, Nikolaev D P and Arlazarov V 2017 Snapscreen: Tv-stream frame search with projectively distorted and noisy query Ninth International Conference on Machine Vision (ICMV) 10341 103410
[10] Karpenko S, Konovalenko I, Miller A, Miller B and Nikolaev D 2015 Uav control on the basis of 3d landmark bearing-only observations Sensors 15(12) 29802-29820
[11] Kholopov I S 2017 Projective distortion correction algorithm at low altitude photographing Computer Optics 41(2) 284-290 DOI: 10.18287/0134-2452-2017-41-2-284-290
[12] Kunina I A, Gladilin S A, Nikolaev D P 2016 Blind radial distortion compensation in a single image using fast Hough transform Computer Optics 40(3) 395-403 DOI: 10.18287/2412-6179-2016-40-3-395-403
[13] Dubuisson M-P and Jain A K 1994 A modified hausdorff distance for object matching Pattern Recognition Conference A: Computer Vision & Image Processing 1 566-568
[14] Huttenlocher D P, Klanderman G A and Rucklidge W J 1993 Comparing images using the hausdorff distance IEEE Transactions on pattern analysis and machine intelligence 15(9) 850-863
[15] Sim D-G, Kwon O-K and Park R-H 1999 Object matching algorithms using robust hausdorff distance measures IEEE Transactions on image processing 8(3) 425-429
[16] Jesorsky O, Kirchberg K J and Frischholz R W 2001 Robust face detection using the hausdorff distance International Conference on Audio-and Video-Based Biometric Person Authentication 90-95
[17] Jaccard P 1901 ´Etude comparative de la distribution florale dans une portion des alpes et des jura Bull Soc Vaudoise Sci Nat 37 547-579
[18] ICDAR 2017 POD Competition (Access mode: http://www.icst.pku.edu.cn/cpdp/icdar2017_podcompetition/index.html)
[19] Hartley R and Zisserman A 2003 Multiple view geometry in computer vision (Cambridge university press)
[20] Krasnoselskiy M A, Petrov A I, Pvolockiay A I and Zabreyko P P 1963 Vector fields on the plane (State Publishing House of Physical and Mathematical Literature)

Acknowledgements
The research was conducted with the support of the RFBR grants No 17-29-03514 and 15-29-06086.