# Approach for Document Detection by Contours and Contrasts

**Daniil Tropin**[1,2], Sergey Ilyuhin[1,2], Dmitry Nikolaev[1,4], Vladimir V. Arlazarov[1,3]

[1] Smart Engines Service LLC

[2] Moscow Institute of Physics and Technology (NRU)

[3] Federal Research Center “Computer Science and Control” RAS

[4] Institute for Information Transmission Problems (Kharkevich Institute) RAS

# Goal: document detection problem

**Input**: Image

**Output**: Quadrilateral corresponding to the document location

- There is only one document in the image;
- All document borders are visible;
- Document has an unknown internal structure;
- Complex background is possible;
- Priori information about camera intrinsic parameters is NOT available.


Image from MIDV-500 dataset

Example of an image with highlighted quadrilateral of the document border

Contour-based approach

Input image

Contours and Lines

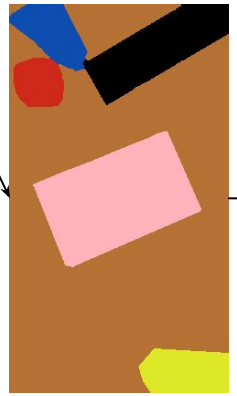
Quads by lines intersection

Quads ranking by contour score

Quad with max contour score

Incorrect match

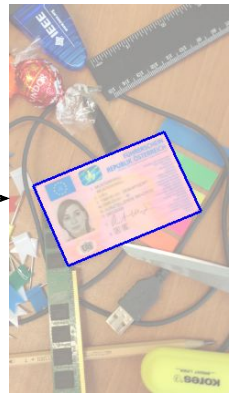
Region-based approach



Regions

Quads by analysis of region

Quads ranking by contrast score



Quad with max contrast score

Correct match

Confidence

# The ranking problem statement

Input:

- Image $\boldsymbol{I}$
- Set of quadrilaterals $\{q\}_{i=0}^{N}$

- Ground truth quadrilateral $\boldsymbol{m}$

It is required to define a function $\boldsymbol{F}$ such that

$$\begin{cases} i^* = arg\,max_{i=0}^{N} F(q_i, I) \\ L(q_{i^*}, m) = 1 \end{cases},$$

where $\boldsymbol{L}$ is a binary quality metric.

$\boldsymbol{F}$

# Proposed function


Image $I$ and quadrilateral $q$

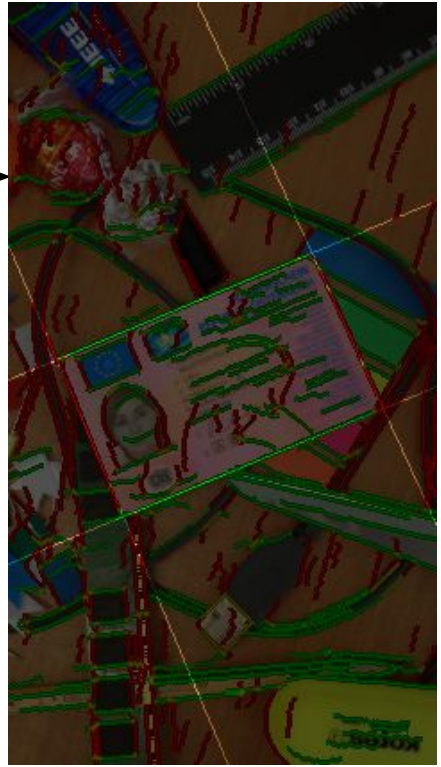
**C**ontour score is based on integral statistics of edge map along borders of $q$

$$F(q, I) = R(q, I) + kC(q, I)$$

**R**egion score is based on $\chi^2$ distance between sets of pixels in $A$ and $B$

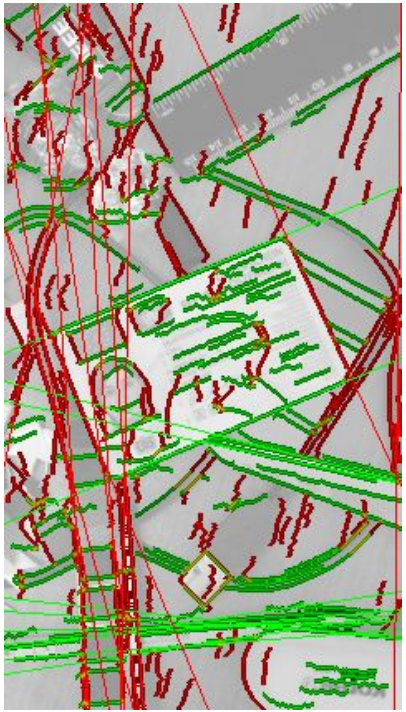
Projectively rectified image ($q \rightarrow t$)


External region $A$, internal region $B$



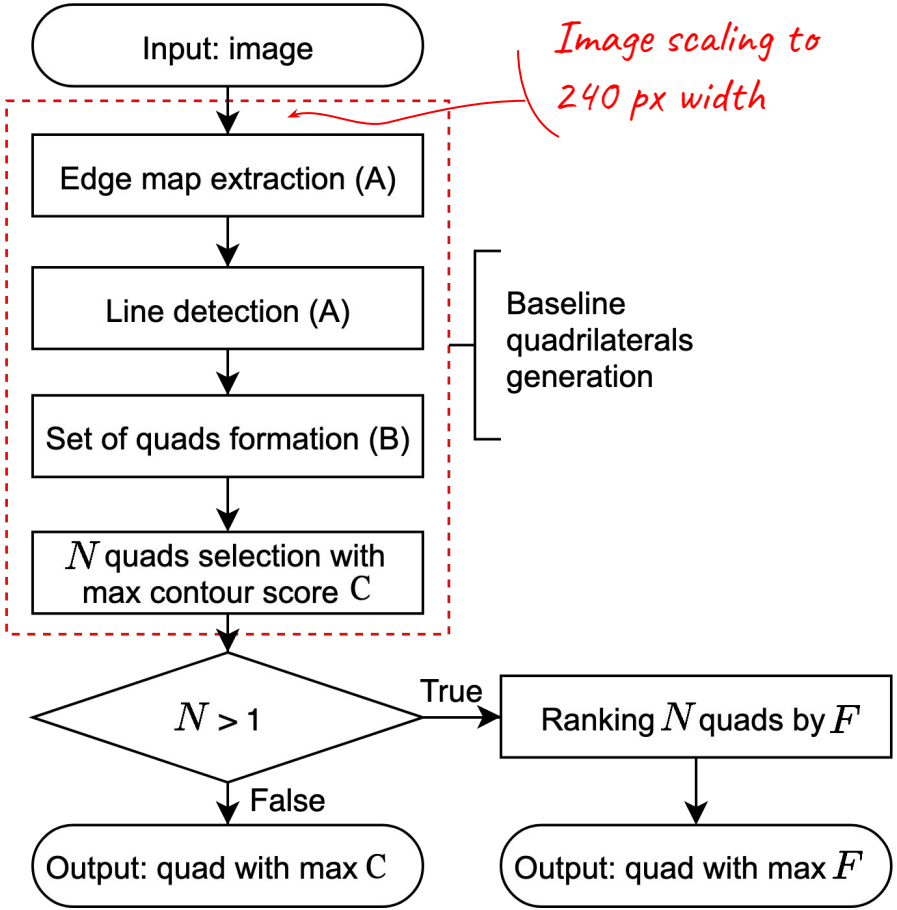
Highlighted edges along lines forming $q$

# Pipeline of algorithm



(A) Horizontal and vertical edges and lines

(B) Quad formation by 2 vertical and 2 horizontal lines



# Error classification on MIDV-500 dataset

15 000 images
50 ID cards
10 backgrounds

Target class of errors

| | (i) Out of frame | (ii) No line | (iii) Ranking error | Total errors | Run-time (ms/frame) |
|---|---|---|---|---|---|
| Proposed algo. (N=1) | 2850 | 660 | 854 | 4366 | 82 |
| Proposed algo. (N=11) | 2803 | 627 | 509 | 3941 | 88 |

| | | | | | |
|---|---|---|---|---|---|
| Improvement value | +47 | +33 | +345 | +425 | -6 |
| Improvement ratio | +1.65% | +5.00% | **+40.40%** | +9.73% | -7.3% |

~ top 1 quad by Contour score only

Jaccard Index < 0.945

iPhone 6 in single thread mode

# Improvement examples of Ranking error





**Red** quadrilateral corresponds to the top **contour** alternative
**Blue** – to the top alternative by **contour and region score**

Figures from left to right
(1), (2), (3), (4)

# Comparison with state-of-the-art

Competitive result on SmartDoc

| System | MIDV-500 | | | SmartDoc | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 4 vertices in | At least 3 in | Full | Bgr 1 | Bgr 2 | Bgr 3 | Bgr 4 | Bgr 5 | Full |
| Proposed algo. (N=1) | 0.968 | 0.955 | 0.861 | 0.98 | 0.974 | 0.982 | 0.966 | 0.294 | 0.906 |
| Proposed algo. (N=11) | **0.972** | **0.961** | **0.87** | 0.983 | 0.974 | 0.983 | 0.97 | 0.319 | 0.91 |
| CS-NUST-2 | 0.739 | 0.705 | 0.626 | 0.988 | 0.976 | 0.984 | 0.974 | 0.948 | 0.978 |
| OctHU-PageScan | 0.403 | 0.374 | 0.319 | | | | | | |
| JCD+CSR | | | | 0.988 | **0.984** | 0.983 | **0.984** | **0.961** | **0.982** |
| GOP | | | | 0.961 | 0.944 | 0.965 | 0.93 | 0.412 | 0.896 |
| LRDE-2 | | | | 0.905 | 0.936 | 0.859 | 0.903 | | |
| LRDE-3 | | | | 0.985 | 0.982 | 0.987 | 0.98 | 0.848 | 0.97 |
| DBSCAN | | | | | | | | | 0.942 |
| SmartEngines | | | | **0.989** | 0.983 | **0.99** | 0.979 | 0.688 | 0.955 |
| SmartDoc (Averaged) | | | | 0.947 | 0.903 | 0.938 | 0.812 | 0.404 | 0.855 |

Top on MIDV-500

Table 2. Mean Jaccard Index

# Conclusion

- The scoring function with contour and contrast features is proposed.
- Proposed modification reduced the number of ranking errors by 40%;
- Runtime is still small even on the mobile phones;
- The highest quality on MIDV-500 dataset;
- The competitive results on 4/5 parts of SmartDoc dataset;

Future work: Geometric properties of the document and Video stream mode.

# Approach for Document Detection by Contours and Contrasts

**Daniil Tropin**[1,2], Sergey Ilyuhin[1,2], Dmitry Nikolaev[1,4], Vladimir V. Arlazarov[1,3]

**daniil_tropin@smartengines.com**

[1] Smart Engines Service LLC
[2] Moscow Institute of Physics and Technology (NRU)
[3] Federal Research Center “Computer Science and Control” RAS
[4] Institute for Information Transmission Problems (Kharkevich Institute) RAS