Screen recapture detection based on color-texture analysis of document boundary regions

I.A. Kunina [1,2], A.V. Sher [2,3], D.P. Nikolaev [1,2]
[1] Institute for Information Transmission Problems of RAS (Kharkevich Institute), 127051, Russia, Moscow, Bolshoy Karetny per. 19, build. 1;
[2] Smart Engines Service LLC, pr. 60-letiya Oktyabrya, 9, Moscow, 117312, Russia;
[3] Moscow Institute of Physics and Technology (State University), 141701, Russia, Dolgoprudny, Institutskiy per. 9

Abstract

This paper examines a presentation attack detection when a document recaptured from a screen is presented instead of the original document. We propose an algorithm based on analyzing a moiré pattern within document boundary regions as a distinctive feature of the recaptured image. It is assumed that the pattern overlapping the document boundaries is a recapture artifact, not a match between document and background textures. To detect such a pattern, we propose an algorithm that employs the result of the fast Hough transform of the document boundary regions with enhanced pattern contrast. The algorithm performance was measured for the open dataset DLC-2021, which contains images of mock documents as originals and their screen recaptures. The precision of the proposed solution was evaluated as 95.4 %, and the recall as 90.5 %.

Keywords: document analysis, document liveness detection, screen recapture detection, fast Hough transform.

Citation: Kunina IA, Sher AV, Nikolaev DP. Screen recapture detection based on color-texture analysis of document boundary regions. Computer Optics 2023; 47(4): 650-657. DOI: 10.18287/2412-6179-CO-1237.

Acknowledgements: This work was partially supported by the Russian Foundation for Basic Research (Project No. 18-29-26035).

Introduction

The functionality of modern document recognition systems has long gone beyond data extraction tasks. As optical character recognition (OCR) becomes more prominent in various fields, new requirements are imposed on recognition systems. For example, modern systems are tasked with verifying the authenticity of the presented document. Among different approaches there are checking the presence of the holograms [1], the conformance of the used fonts to the government standards [2], etc.

At the same time, the demand for the recognition of documents from very different sources using mobile phones is ever-growing. A set of measures aimed at detecting an attempt to use a photocopy, mock-up, or an image of a document recaptured from a screen is referred to as document liveness detection, similar to a face liveness detection, which is aimed at presentation attack (presenting 3D masks, photos or video recordings) detection in facial recognition systems.

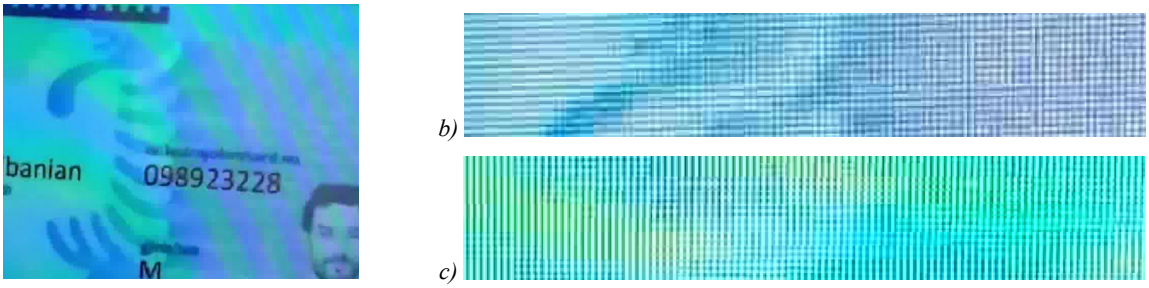
In this paper, we propose a method for the detection of screen recaptured document images based on analyzing a moiré pattern. The moiré pattern, which is caused by overlapping regular structures of the display screen and the camera sensor, shows great diversity in its appearance (see Fig. 1). This makes the modelling of moiré patterns challenging. However, we observed that the patterns, generally, do not change rapidly in local regions, could be divided into a brightness moiré pattern (see examples in Fig. 2*a*, *b*) and a color moiré pattern (see examples in Fig. 2*c* – *e*), and could be approximated by a set of straight, parallel stripes. And even if moiré does not have a pronounced orientation as in Fig. 2*b*, it could be approximated by several overlapping sets with different orientations. To distinguish screen moiré pattern from document own one (compare examples in Figs. 2 and 3), we based our method on analyzing document and background local regions near the boundary. The decision whether the document is authentic or not is based on the assumption that screen recapturing generates a pattern in the image that is similar on both sides of the boundary. The Hough transform is employed to detect and determine the parameters of such a pattern.

The rest of the paper is structured as follows: section 1 discusses existing methods of screen recapture detection; section 2 discusses the details of the used Hough transform; section 3 describes the general scheme of the proposed algorithm for moiré pattern detection; section 4 proposes an additional algorithm for screen recapture detection when screen boundaries are within an input image; finally, section 5 provides results of testing the proposed algorithms on images from the open dataset DLC-2021 [3].

1. A review of modern methods for screen recapture detection

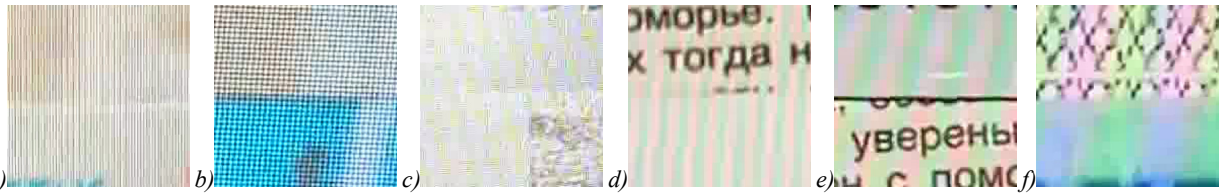
Existing screen recaptured detection methods mainly focus on indoor and outdoor scenes [4 – 7] and biometric (e.g. face) images [8 – 10]. However, document images are very different from natural and face ones. For

example, the real natural images and face images are characterized by the differences in depth between background and foreground, while the recaptured faces and real world are of uniform depth. Such differences are not observed in document images, since both genuine and recaptured document images are of uniform depth.



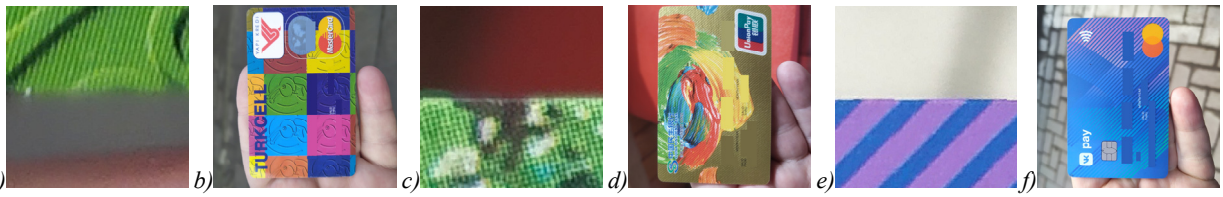
a) b) c)

Fig. 1. Diversity of moiré pattern appearance on documents from the open dataset DLC-2021: (a) color pattern with changing direction and period, (b) brightness pattern with changing orientation severity, (c) combined pattern



a) b) c) d) e) f)

Fig. 2. Examples of the considered moiré pattern withing boundary regions on documents from the open dataset DLC-2021: (a) and (b) show brightness moiré patterns with different orientation severity; (c)-(f) show color moiré patterns with different texture intervals and colors



a) b) c) d) e) f)

Fig. 3. Examples of real documents with their own patterns: (a) shows detailed texture from (b), the same is for (c) and (d), (e) and (f) respectively

Document-specific recapturing attacks are considered in [11, 12], but the authors focused on printed copies of documents, not copies displayed with a screen.

Detection of screen recaptured documents using DenseNet neural network architecture is considered in [13]. To train the neural network, a dataset was constructed using German ID cards and residence permits and consisted of bona fide ID cards, printed ID cards and screen recaptures of ID cards. Unfortunately, neither the neural network nor the test dataset has been published. The images within the dataset are not publicly available due to the personal data they contain, and the authors acknowledge the need for further research of the proposed solution for different types of documents. Also, the authors point out that when a document is not well aligned with the presence of fingertips in the background, the neural network classifies originals as recaptures, and the bright reflection on the face region in a recaptured image makes the neural network classify it as an original one.

Being close to a document topic, paper [14] presents a method to detect moiré patterns in screen recaptured images of receipts using a multi-input deep Convolutional Neural Network trained with Haar Wavelet Decomposition. The images within the dataset are not publicly available too due to the personal data they contain.

Based on the literature review, we decided to employ the DLC-2021 [3] dataset as the only public dataset which is dedicated to the specific problem considered in the paper. The DLC-2021 dataset includes images of mock documents as originals and images of their copies recaptured from different screens. In addition to the data itself, the authors published a basic screen recapture detector based on deep learning. This algorithm can be used as a baseline algorithm for performance comparison.

2. Fast Hough transform

To calculate the Hough transform (HT), we use the Brady-Yong [15] method, also referred to as the fast Hough transform (FHT) [16] because of its lower complexity comparing to the classical method ($\Theta(n^2 \log n)$ vs. $\Theta(n^3)$ operations) [17].

In the Brady-Yong FHT algorithm, all straight lines are divided into two classes: mostly vertical lines ($|\mathrm{tg}(\phi)| \le 1$, where ϕ is the angle between the normal to the line and x-axis) and mostly horizontal lines ($|\mathrm{ctg}(\phi)| \le 1$). Every line is approximated with a discrete dyadic pattern and parametrized by two points on the opposite boundaries of the image: a “vertical” line is parametrized by points $(s, 0)$ and $(s+t, n-1)$ on the horizontal boundaries, and a “horizontal” line is

parametrized by points $(0, s)$ and $(n-1, s+t)$ on the vertical boundaries. Here s is an integer shift of a pattern, and t is an integer slope. Such a line corresponds to a point (s, t) in a Hough image, and its value is a sum of values of image pixels along the pattern.

We will denote the Hough images of vertical and horizontal lines as H_v and H_h respectively. Due to the fact that the Brady-Yong algorithm calculates FHT for lines with slope within 45°, to obtain the Hough image H_v (or H_h) two Hough images are computed independently and combined into one according to the scheme described in [16].

3. Moiré pattern detection along document borders

3.1. General description of the proposed algorithm

The general algorithm of the detection of the moiré pattern overlapped the document boundary is illustrated in Fig. 4*a*. The input of the algorithm includes a frame (an image in question) and the coordinates of the quadrilateral that defines the position of the document in the image. The detector extracts adjacent rectangular regions for each line segment of the document boundary (see section 3.2), splits each region into small windows, and searches for a striped pattern (brightness or color) in the windows that is oriented similarly across the document boundary. Recapture artifacts are considered to be detected if a color or brightness "leaking" pattern has been found for at least one region.

To detect and analyze moiré texture, the detector relies on the Hough image of the preprocessed region. The preprocessing of the region maximizes the contrast of the expected texture against other image details (the contrast enhancement for brightness and color texture is described in sections 3.3 and 3.4 correspondingly). After contrast enhancement, the texture parameters in each window are evaluated based on a Hough image: the main orientation and expression degree (see section 3.5). Then, for each pair of windows adjacent to the document boundary, the similarity of the parameters on different sides of the boundary is checked (see section 3.6). A "leaking" texture is considered to be present if at least T_W pairs of windows with close parameters are found for the same texture type, i.e. color or brightness patterns.

This basic procedure is employed to detect and analyze the brightness pattern (see diagram in Fig. 4*b*). Due to the wide variation of the color texture intervals (as seen when comparing the examples in Fig. 2*c*–*f*) the color moiré pattern search is performed for the boundary region image shrunk horizontally with a factor 2^s, $s = 0, 1\ldots, S-1$, where S is the number of employed scaling factors (see the diagram in Fig. 4*c*).

In the experiments below, the following values of the mentioned parameters were used: $T_W = 2$, $S = 4$.

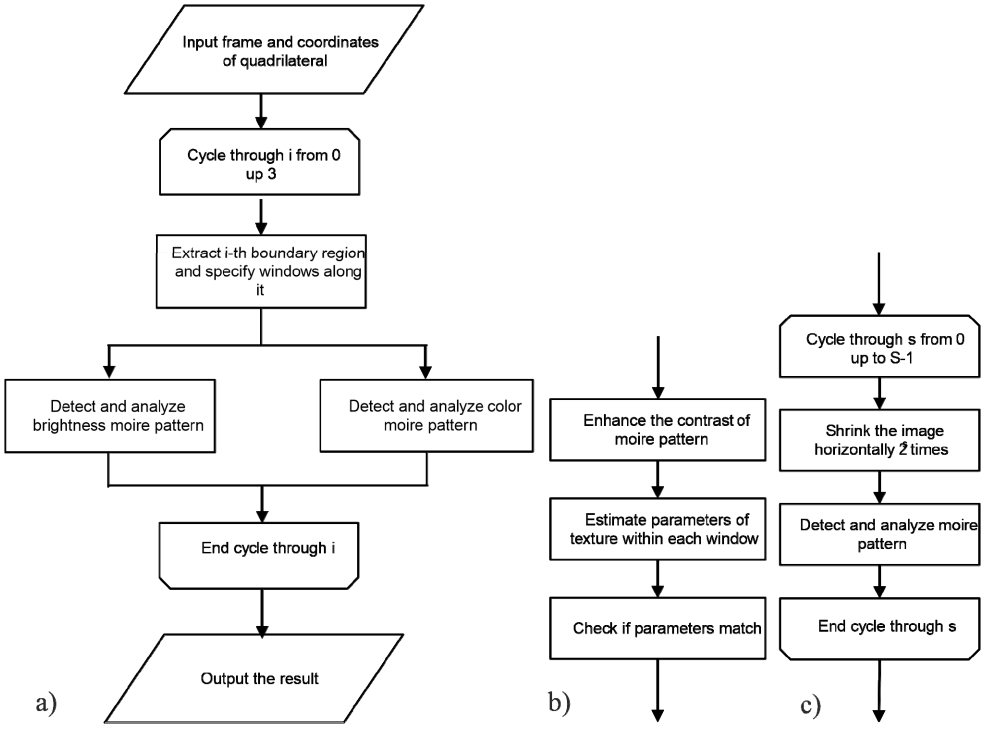


Fig. 4. General diagram of the moiré pattern "leak" detection algorithm: (a) top-level diagram, (b) basic diagram of moiré patterns detection and comparison at the boundary, (c) multi-scale search and comparison for color moiré patterns

3.2. Extraction of the document boundary region in the input image

Let the linear segment of the document boundary be specified by a segment with vertices $\mathbf{p_b} \stackrel{def}{=} (x_b, y_b)$ and $\mathbf{p_e} \stackrel{def}{=} (x_e, y_e)$ for an input frame $\mathrm{I_{in}}$. Denote a boundary as vertical if $|x_b - x_e| < |y_b - y_e|$, and horizontal if otherwise.

Let us construct the $\mathrm{I_{corr}}$ image that has a height of $2n_0$ and contains the document boundary which is oriented horizontally and passes through the center of $\mathrm{I_{corr}}$. At first,

we specify a parallelogram P with vertices $\langle \mathbf{p}_b+\mathbf{p}_s,\ \mathbf{p}_e+\mathbf{p}_s,\ \mathbf{p}_e-\mathbf{p}_s,\ \mathbf{p}_b-\mathbf{p}_s\rangle$ which includes the boundary in question (this parallelogram is shown in green in Fig. 5*a*). Here, $\mathbf{p}_s \underline{\underline{def}} (n_0, 0)$ if the boundary is vertical and $\mathbf{p}_s \underline{\underline{def}} (0, n_0)$ if the boundary is horizontal. At second, we get the minimal orthotropic rectangular region I_{roi} that covers P in I_{in} and, if the boundary is vertical, we transpose I_{roi} (the region I_{roi} is shown in Fig. 5*a* as a yellow rectangle, and the corresponding got region is illustrated in Fig. 5*b*). Finally, we construct the image I_{corr} by affine transformation (shear) of the image I_{roi} (see Fig. 5*c*):

$$\mathrm{I}_{\mathrm{corr}}(x,y) \coloneqq \mathrm{I}_{\mathrm{roi}}\left(x, y+\left[\frac{y_b-y_e}{x_b-x_e}\left(x-\left[\frac{w_{roi}}{2}\right]\right)\right]\right), \quad (1)$$
$$x\in[0, w_{roi})\cap\mathbb{Z},\ \ y\in[0, 2n_0)\cap\mathbb{Z},$$

where w_{roi} is the width of I_{roi}.

Further, to eliminate the possible inaccuracy in given document boundary coordinates, the central boundary region is removed from I_{corr}. For this, I_{out} is constructed via vertical concatenation of regions $(0, 0, w_{roi}, n_{fin})$ and $(0, 2n_0-n_{fin}, w_{roi}, n_{fin})$ of I_{corr}. Here and below, the first two parameters are the x and y coordinates of the starting point, the last two ones are the horizontal and vertical dimensions. The result is shown in Fig. 5*d*.

For further processing, the pairs of square windows on different sides of the boundary in I_{out} are specified as follows: $(in_{fin}, 0, n_{fin}, n_{fin})$ and $(in_{fin}, n_{fin}, n_{fin}, n_{fin})$, where $0 \le i < \lfloor w_{roi}/n_{fin} \rfloor$, $i\in\mathbb{Z}$ (see Fig. 5*e*).

In the experiments below, the following values of the mentioned parameters were used: $n_0=90$, $n_{fin}=60$.

3.3. Enhancing the contrast of a brightness moiré pattern

Let I_{in} be the image of a boundary region in question for the brightness moiré detection. We omit color information by averaging the values throughout the color channels and construct a grayscale image L_{in}. Let us estimate the local average of the signal amplitude L_{in} via a median filter with a window of $w_{med}\times w_{med}$ pixels and denote the obtained image as L_{med}. To enhance the contrast of the brightness moiré pattern, we employ the following transformation:

$$\mathrm{I}_{\mathrm{out}}(x,y) \coloneqq 128 + k_{max}\, k_{low}(x,y)\,\mathrm{sgn}(\mathrm{L}_{\mathrm{in}}(x,y) - \mathrm{L}_{med}(x,y)), \quad (2)$$
$$k_{low}(x,y) = \frac{min\left(\left|L_{in}(x,y)-L_{med}(x,y)\right|, T_L\right)}{T_L}.$$

Here, k_{low} regulates the importance of small ($<T_L$) differences in brightness, and the parameter k_{max} sets the contrast of the final single-channel image I_{out}.

An example of a brightness moiré pattern with an enhanced contrast by the proposed method is shown in Fig. 6*b* for the input image shown in Fig. 6*a*.

In the experiments below, the following values of the mentioned parameters were used: $w_{med}=3$, $k_{max}=28$, $T_L=3$.

Fig. 5. Extraction of the document boundary region: (a) input image from the open dataset DLC-2021 with a known document location specified by the red quadrilateral and a boundary region in question specified by the yellow rectangle; (b) corresponds to the image region within the yellow rectangle; (c) corresponds to the area within the green parallelogram after shear correction of the image (b), (d) the final image of the boundary region after removing the central area, (e) the result of specifying the windows for further moiré detection (shown as white square

3.4. Enhancing the contrast of a color moiré pattern

Let I_{in} be the image of a boundary region in question for the color moiré detection. Let us estimate the local color of I_{in} by a window average filter of $w_{avg}\times w_{avg}$ pixels and denote the obtained image and its linear brightness $(L=(R+G+B)/3)$ as I_{avg} and L_{avg} respectively. Finally, we apply locally color normalization to the original image I_{in} under the "gray world" [18] assumption:

$$\mathrm{I}_{\mathrm{gw}}(x,y) \coloneqq \mathrm{L}_{\mathrm{avg}}(x,y)\,\mathrm{I}_{\mathrm{in}}(x,y) \circ \mathrm{I}_{\mathrm{avg}}{}^{\circ -1}(x,y), \quad (3)$$

where the symbol "$\circ$" denotes Hadamard product, and "$^{\circ -1}$" denotes element-by-element conversion.

To enhance the contrast, we convert the image I_{gw} from RGB space to Hexcone HSV space using the formulas from [19], and then set the components S and V of the resulting image to 0.5. Converting the image back to RGB, we obtain I_{out} with enhanced contrast.

An example of color normalization and contrast enhancement are shown in Fig. 6*d* and Fig. 6*e* correspondingly for an input region ($s=0$) shown in Fig. 6*c*.

In the experiments, $w_{avg}=15$ was used.

3.5. Texture parameters estimation

Let I_{in} be one of the windows of the grayscale image with enhanced contrast, whether the image with enhanced brightness contrast or one of the channels of the image with enhanced color contrast. For I_{in} we calculate the Hough images H_h and H_v for mostly horizontal and mostly vertical lines, respectively.

We employ sum of squares of intensity gradients (SSG) [20] method to determine the orientation of the texture by the Hough image:

$$\mathrm{SSG_H}(t_0)\,\mathrm{def}=\ \mathrm{k}^3(t_0)\sum_{s=0}^{w-1}(\mathrm{H}(s+1,t_0)-\mathrm{H}(s,t_0))^2, \quad (4)$$

where

$$\mathrm{k}(t)=\sqrt{1+\frac{(t-\lfloor h/2\rfloor)^2}{\lfloor h/2\rfloor^2}},$$

$w\times h$ is the size of the Hough image H, and the weight function k(t) guarantees the criterion covariance in terms of the input image rotation.

The dominant orientation of the pattern is the direction t_m where the criterion SSG is maximal:

$$\begin{aligned} &h_m := [t_h > t_v],\ \ t_m := \max(t_h, t_v),\\ &t_\chi = \arg\max_t \mathrm{SSG_{H_\chi}}(t),\ \ \chi\in\{h,v\} \end{aligned} \quad (5)$$

where [•] denotes the Iverson bracket, where h_m specifies one of the two angular ranges.

The coefficient of the texture orientation expression c_m is defined as the ratio between the criterion value SSG in the direction perpendicular to the detected one and the maximum:

$$c_m := 1-\left(h_m\frac{SSG_{H_v}(t_m)}{SSG_{H_h}(t_m)}+(1-h_m)\frac{SSG_{H_h}(t_m)}{SSG_{H_v}(t_m)}\right). \quad (6)$$

The value of the coefficient c_m is close to 1 in the case of pronounced texture orientation, and to 0 in the case of isotropic texture.

3.6. Verifying the match of texture parameters

Let the two sets of texture parameters calculated for two opposite boundary adjacent windows be $\langle h_1, t_1, c_1\rangle$ and $\langle h_2, t_2, c_2\rangle$. The parameters characterize the same texture if the following condition is met:

$$h_1=h_2 \wedge c_1>T_c \wedge c_2>T_c \wedge |t_1-t_2|\le\Delta_t \wedge |c_1-c_2|\le\Delta_c. \quad (7)$$

where T_c, Δ_t, Δ_c are configurable thresholds. When analyzing a color texture, the parameters are independently verified in each channel, and finally, the texture is considered "leaking" if the condition (7) is met for at least two of the three channels.

In Fig. 6*b*, 6*f* texture parameters estimation and match verification is shown. The detected texture orientation t_m was visualized by calculating the corresponding coordinates in the space of the corresponding window, and the detected matches were visualized by orthotropic rectangles covered the windows.

In the experiments below, the following values of the mentioned parameters were used: $T_c=0.1$, $\Delta_t=5$ when analyzing the brightness texture, $T_c=0.5$, $\Delta_t=10$ when analyzing the color texture, $\Delta_c=0.2$.

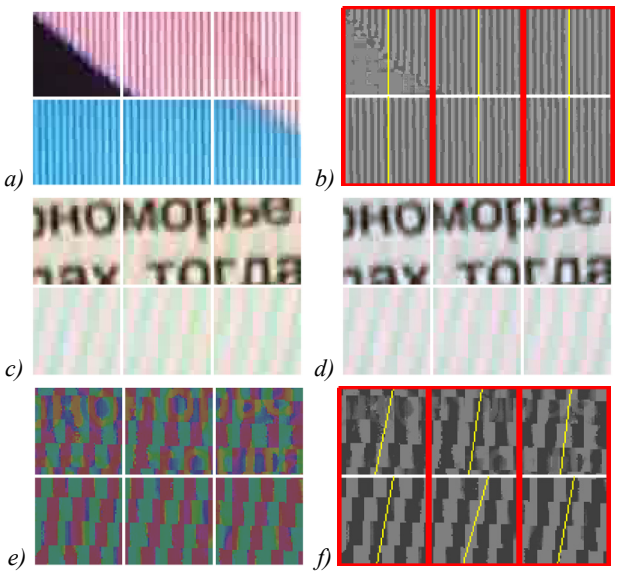


Fig. 6. Detection of brightness and color moiré patterns: (a) input bound- ary image with the brightness pattern; (b) result of matching the brightness pattern within the image with enhanced contrast; (c) input boundary image with color moiré pattern (s = 0); (d) result of "gray world" local normaliza- tion; (e) result of the contrast enhancement of the color pattern; (f) result of matching the color pattern in the green channel of the image with enhanced contrast. Every yellow segment marks the detected orientation t_m *of the pattern in the corresponding window. The red rectangles covered the vertically adjacent texture pairs* t_1 *and* t_2 *mark the detected texture matches*

4. Detection of display frames in the image

As a camera moves away from a screen, the moiré pattern caused by interference between the grids of the camera sensor and the screen becomes less and less apparent. At the same time, the frame of the display inevitably shows up in the captured image. Therefore, a reasonable addition to the basic algorithm is to detect preliminarily the presence of the display frame that might surround the document.

The input of the algorithm includes a frame I_{in} and the coordinates of the quadrilateral that defines the position of the document boundaries. The image I_{in} is anisotropically scaled to a square image I_{small} with the side of n_s. Since both the screen and the image have an aspect ratio significantly different from 1, it is possible that only two sides of the frame are present in the image even at a great distance (see Fig. 7*a*). Thus, only two parallel sides of a display are enough to detect the display. To avoid false positives, it is assumed that each

side of the screen frame should be presented by two parallel close lines.

At each point of I_{small} a morphological gradient with a window of $w_s \times w_s$ is calculated to construct an image I_{grad}. Let us assume that the orthotropic rectangle framing the document image has coordinates $(l_{roi}, t_{roi}, w_{roi}, h_{roi})$ within I_{grad}, where (l_{roi}, t_{roi}) are the coordinates of the upper left corner, and w_{roi} and h_{roi} are the horizontal and vertical dimensions correspondingly. Denote a pair of orthotropic rectangular regions: an upper I_{top} with coordinates $(l_{roi}, 0, w_{roi}, c_y)$ and a lower I_{bot} with coordinates $(l_{roi}, c_y, w_{roi}, n_s - c_y)$, where (c_x, c_y) is the center of the input quadrilateral in coordinates of I_{grad}.

For each of these regions, a Hough image for mostly horizontal lines is calculated. Further processing of I_{top} and I_{bot} is identical, so we denote the Hough image as H_h without specifying the region to be processed.

To loosen sharp peaks produced by textured areas, we subtract the smoothed along axis t Hough image H_h from the original one, keeping only positive values:

$$\begin{aligned} &H'_h(s,\bullet) := \\ &= \max(0, H_h(s,\bullet) - G_\sigma(s) * H_h(s,\bullet)) \circ W(s,\bullet), \end{aligned} \tag{8}$$

where $G_t(\sigma)$ is the Gaussian filter kernel with the standard deviation of σ, the symbol "*" denotes convolution, the indicator function W equals 0 if the line with the corresponding parameters intersects the document quadrilateral, and 1 otherwise, the symbol "$\circ$" denotes Hadamard product.

The slope of the estimated monitor edge t_m is selected according to the SSG criterion (see section 3.5), and, to find a pair of edges in the corresponding line of the Hough image, we detect pairs of close peaks characterized by the amplitude sum:

$$\begin{aligned} &t_m := \arg\max_t \mathrm{SSG}_{H'_h}(t), \\ &S_m(s_r) := H'_h(s_r, t_m) + \max_{|s_r - s_l - d| \le \Delta_d} H'_h(s_l, t_m), \end{aligned} \tag{9}$$

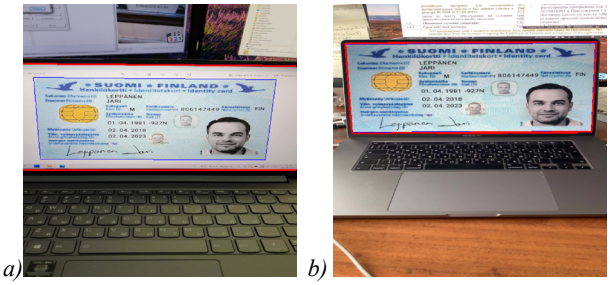
where d is the expected width of the frame, and Δ_d is the allowed deviation from it. If $\max S_m < T_s w_{roi}$, then a side is not found. Otherwise, a side is detected as a line with coordinates $(\arg\max S_m(s), t_m)$.

The display is considered detected if sides are found in both I_{top} and I_{bot}.

The detection of the left and right sides of the screen is performed in exactly the same way, but with the image I_{grad} transposed. The result of the display frames detection when only the top and bottom sides of the screen are visible is shown in Fig. 7*a*, and the result when all sides are visible is illustrated in Fig. 7*b*.

In the experiments below, the following values of the mentioned parameters were used: $n_s = 480$, $w_s = 3$, $\sigma = 2$, $d = 30$, $\Delta_d = 25$, $T_s = 40$.

It is worth noticing that direct using of the proposed algorithm for display frames detection could take a cover of a document as a frame. However, since our method is proposed to be used during remote identity document verification, to avoid the false alarm the recognition system must allow a user to take off the cover and try to pass the verification again. In case when the document is displayed on a screen, removing the cover is possible only with capturing at close distance, where moiré is appeared again.



a) *b)*

Fig. 7. Display frames detection when (a) two screen sides shown, (b) all screen sides shown. Given document boundaries are marked in blue, and detected display boundaries are marked in red. The images were taken from the open dataset DLC-2021

5. Experimental evaluation of the proposed algorithm

The algorithm was tested on subsets of the open dataset DLC-2021 [3] contained images of "real" identity documents of 10 different types and their screen recaptures (denoted by abbreviations *or* and *re* respectively). The coordinates of the document boundaries were known from the given annotations of the corresponding images in the dataset.

A convolutional neural network (CNN), provided by the authors of the DLC-2021 dataset, which detects screen recaptures, was adopted as the baseline algorithm. The quantitative comparison with the baseline was performed on the images from *screen_negative_test.lst* and *screen_positive_test.lst* used for evaluating the quality of CNN. Additionally, the performance of the proposed algorithm was evaluated on all images from the *or* (16264 images) and *re* (21896 images) subsets of the DLC-2021 dataset. The results of the comparison illustrated in Table 1 show that the best modification of the proposed algorithm is its full version ("boundaries analysis + display edges detection"), which is not only superior to the CNN in terms of accuracy, but also allows for the best accuracy among all modifications.

In addition to the accuracy analysis, the computational complexity of the proposed method and the baseline CNN were evaluated. The comparison was performed for ARM Cortex-A73 (arm-v8) and Intel Core i7-3770K (×86) processors using image datasets *01.or0004* and *01.re0001* of Greek passports *grc_passport*. The TensorFlow Lite platform was used to employ CNN in the experiments. The comparison results illustrated in Table 2 show that the proposed algorithm outperforms the CNN by 1.5–2 times.

Conclusion

In this paper, we propose a new method of screen recapture detection for documents presented during

remote user identification. The main idea behind the proposed method is the detection and analysis of the specific moiré pattern within the document and background local regions near the boundary in order to distinguish recapture artifacts and own texture of the documents. The algorithm was evaluated using the DLC-2021 dataset, which contains more than 35 thousand images of original documents and their recaptures. The evaluation showed that the proposed algorithm outperforms the baseline neural network approach in terms of accuracy and computational complexity. Future work includes testing the algorithm stability on images obtained with different registration devices, screen resolutions and types of frames due to the insufficient number of types of devices and screens used to create the DLC-2021 dataset.

Tab. 1. Performance comparison of various modifications of the proposed method and the baseline CNN provided by authors of DLC-2021. The method which demonstrates the maximum accuracy is marked in bold

algorithm version	CNN test set from *or* and *re*			all images from *or* and *re*		
	recall,%	precision,%	accuracy,%	recall,%	precision,%	accuracy,%
the baseline algorithm [3]	89.0	85.9	89.7	-	-	-
boundaries analysis + display edges detection	90.2	93.5	**92.1**	90.5	95.4	**92.0**
only boundaries analysis	82.6	93.2	88.6	83.1	95.2	87.9
only brightness moiré pattern detection	79.4	98.7	89.5	78.6	99.1	87.2
only color moiré pattern detection	53.6	91.5	74.9	55.0	94.0	72.0
color moiré pattern detection without scaling	47.2	97.5	73.6	47.4	98.3	69.2

Table 2. Computational complexity (measured in sec / 1 img) comparison of the proposed method and the baseline CNN provided by authors of DLC-2021

(sec/1 img)	proposed (x86)	CNN (x86)	proposed (arm-v8)	CNN (arm-v8)
grc_passport/01.or0004	0.155	0.224	0.433	0.828
grc_passport/01.re0001	0.128	0.229	0.376	0.828

References

[1] Kada O., Kurtz C., van Kieu C., Vincent N. Hologram detection for identity document authentication. In Book: Yacoubi ME, Granger E, Yuen PC, Pal U, Vincent N, eds. Pattern Recognition and Artificial Intelligence. Pt I. Cham: Springer International Publishing; 2022: 346-357. DOI: 10.1007/978-3-031-09037-0_29.

[2] Chernyshova YS, Aliev MA, Gushchanskaia ES, Sheshkus AV. Optical font recognition in smartphone-captured images and its applicability for ID forgery detection. Proc SPIE 2019; 11041: 110411J. DOI:10.1117/12.2522955.

[3] Polevoy DV, Sigareva IV, Ershova DM, Arlazarov VV, Nikolaev DP, Ming Z, Luqman MM, Burie J-C. Document Liveness Challenge Dataset (DLC-2021). J Imaging 2022; 8(7): 181. DOI: 10.3390/jimaging8070181.

[4] Mahdian B, Novozamsky A, Saic S. Identification of aliasing-based patterns in re-captured LCD screens. 2015 IEEE Int Conf on Image Processing (ICIP) 2015: 616-620. DOI: 10.1109/ICIP.2015.7350872.

[5] Thongkamwitoon T, Muammar H, Dragotti PL. An image recapture detection algorithm based on learning dictionaries of edge profiles. IEEE Trans Inf Forensics Secur 2015; 10(5): 953-968. DOI: 10.1109/TIFS.2015.2392566.

[6] Zhu N, Li Z. Recaptured image detection through enhanced residual-based correlation coefficients. In Book: Sun X, Pan Z, Bertino E, eds. ICCCS 2018: Cloud computing and security. Cham: Springer; 2018: 624-634. DOI: 10.1007/978-3-030-00021-9_55.

[7] Yang C, Yang Z, Ke Y, Chen T, Grzegorzek M, See J. Doing more with Moiré pattern detection in digital photos. IEEE Trans Image Process 2023; 32: 694-708. DOI: 10.1109/TIP.2022.3232232.

[8] Garcia DC, de Queiroz RL. Face-spoofing 2D-detection based on Moiré-pattern analysis. IEEE Trans Inf Forensics Secur 2015; 10(4): 778-786. DOI: 10.1109/TIFS.2015.2411394.

[9] Garcia DC, de Queiroz RL. Evaluating the effects of image compression in Moiré-pattern-based face-spoofing detection. 2015 IEEE Int Conf on Image Processing (ICIP) 2015: 4843-4847. DOI: 10.1109/ICIP.2015.7351727.

[10] Benlamoudi A, Bekhouche SE, Korichi M, Bensid K, Ouahabi A, Hadid A, Taleb-Ahmed A. Face presentation attack detection using deep background subtraction. Sensors 2022; 22(10): 3760. DOI: 10.3390/s22103760.

[11] Yan J, Chen C. Cross-domain recaptured document detection with texture and reflectance characteristics.2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) 2021: 1708-1715.

[12] Chen C, Zhao L, Yan J, Li H. A distortion model-based pre-screening method for document image tampering localization under recapturing attack. Signal Process 2022; 200: 108666. DOI: 10.1016/j.sigpro.2022.108666.

[13] Mudgalgundurao R, Schuch P, Raja K, Ramachandra R, Damer N. Pixel-wise supervision for presentation attack detection on identity document cards. IET Biometrics 2022; 11(5): 383-395. DOI: 10.1049/bme2.12088.

[14] Abraham E. Moiré pattern detection using wavelet decomposition and convolutional neural network. 2018 IEEE Symposium Series on Computational Intelligence (SSCI) 2018: 1275-1279. DOI: 10.1109/SSCI.2018.8628746.

[15] Brady ML, Yong W. Fast parallel discrete approximation algorithms for the Radon transform. SPAA '92: Proceedings of the Fourth Annual ACM Symposium on Parallel Algorithms and Architectures 1992: 91-99. DOI: 10.1145/140901.140911.

[16] Aliev M, Ershov EI, Nikolaev DP. On the use of FHT, its modification for practical applications and the structure of Hough image. Proc SPIE 2019; 11041: 1104119. DOI: 10.1117/12.2522803.

[17] Hough PVC. Machine analysis of bubble chamber pictures. Int Conf on High Energy Accelerators and Instrumentation, CERN 1959: 554-556.

[18] Finlayson GD, Schiele B, Crowley JL. Comprehensive colour image normalization. In Book: Burkhardt H, Neumann B, eds. Computer Vision - ECCV'98. Berlin, Heidelberg: Springer; 1998. DOI: 10.1007/BFb0055685.

[19] Smith AR. Color gamut transform pairs. ACM Siggraph Computer Graphics 1978; 12(3): 12-19.

[20] Bezmaternykh PV, Nikolaev DP. A document skew detection method using fast Hough transform. Proc SPIE 2020; 11433: 114330J. DOI: 10.1117/12.2559069.

Authors' information

Kunina Irina Andreevna, (b. 1992), graduated from Moscow Institute of Physics and Technology (State University) in 2015. Ph.D in Technical Science, is a researcher in the vision systems laboratory at the Institute for Information Transmission Problems. Research interests are image analysis, computer vision, camera calibration. E-mail: *kunina@iitp.ru* .

Sher Artem Vladimirovich, (b. 2000), obtained his bachelor's degree from Lomonosov Moscow State University (MSU) in 2021, since 2021 he has been a Master's student at the Moscow Institute of Physics and Technology (National Research University). Since 2020, Artem has worked as a programmer and technician at Smart Engines Service LLC. Research interests are machine vision, device-efficient neural network implementations, algorithms for fast image processing. E-mail: *sher.av@phystech.edu* .

Dmitry Petrovich Nikolaev, (b. 1978), graduated from Lomonosov Moscow State University (MSU) in 2000, majoring in Physic. Ph.D. in Physics and Mathematics, is a head of the vision systems laboratory at the Institute for Information Transmission Problems. Research interests are machine vision, algorithms for fast image processing, pattern recognition. E-mail: *dimonstr@iitp.ru* .

Code of State Categories Scientific and Technical Information (in Russian – GRNTI)): 28.23.15
Received: October 17, 2022. Final version: May 23, 2023.