Robust Detection of Mosaic Masking Region

Jung-Jae Yu
Human Identification Research Section
ETRI
Deajeon, Republic of Korea
E-mail: jungjae@etri.re.kr

Seung-Wan Han
Human Identification Research Section
ETRI
Deajeon, Republic of Korea
E-mail: hansw@etri.re.kr

***Abstract*— In this paper, a novel method to automatically detect mosaic masking regions in an input image is proposed. Mosaic masking region can be used as an important clue for recognizing commercial pornographic images. The proposed method is composed of three steps. The first step is to extract SRE features using a new cross-shaped feature characteristic, and the second step is to estimate the parameters of a mosaic candidate region, and the final step is SRD verification using the characteristic of luminance distribution in a mosaic masking region. Proposed method is fast and robust to blurring effects caused by image resizing and low quality video compression and it can be used for computer vision application to block adult videos.**

Keywords-component; mosaic; masking; detection; pornographic; adult; feature

I. INTRODUCTION

This work is aimed at the automatic detection of mosaic making regions from input images. A few types of image or video contents contain a mosaic masking region, which is composed of check-board patterns filled with a mixed color in each block. Especially in commercial adult contents made by Japan, mosaic masking is frequently used to hide male and female genitals if the contents are aimed at broadcasting on adult channels. Mosaic masking is also used to hide deadly weapons such as a knife for children protection or to hide faces for personal privacy in terrestrial broadcasting contents. So if the mosaic masking region can be detected automatically, it can be used as an important clue for recognizing especial type of contents. In this paper, we focus on detection of mosaic masking region in commercial pornographic images.

II. RELATED WORKS

Object detection problem has been researched by many computer vision researchers for a long time but most of them are concentrated on particular parts of a human body such as faces[4], hands[5] or pedestrians[6] and on some types of targets for defense industry. Detecting mosaic defections in digital video was tried by a few researchers. In Shui-Fa1 and et al.[1]'s method, even squares are detected firstly and the intersection points of these squares are selected as the features of mosaic macro block (MMB) and the existence of the mosaic region consisted of several MMBs is decided. They also proposed mosaic detection in the edge domain using vertical and horizontal projection of edge intensity but quantitative experimental results are not provided in the paper. More recently, Wei and et al.[2] used canny edge detection and mosaic macro block recognition method. A MMB is regarded as a candidate macro block if it contains solid edges more than 3 or it satisfies a particular condition. Previous trials in [1] and [2] are not suitable for detecting mosaic masking in commercial pornographic images because their purpose is not to detect mosaic masking region but to detect mosaic defection in decompressed video.

III. PROPOSED METHOD

The proposed method is composed of three steps. The first step is to extract cross-shaped feature and the second step is to detect the mosaic candidate region and estimate the parameters of a mosaic candidate region. The final step is the decision about whether the candidate region is a real mosaic masking region or a simple mass of cross-shaped features.

A. SRE feature detection

The cross-shaped features located in each intersection point of mosaic blocks are the foundational materials to find a mosaic candidate. A cross-shaped feature is a point at which a horizontal edge and a vertical edge intersect. To make the feature extraction robust to blurring effect, we do not use the edge intensity itself, but the ratio of averaged edge value in structured sub blocks. The detected features by the proposed method will be called Structured Ratio of Edge (SRE) features. First, the input image is converted to a gray image and horizontal and vertical edge operation is applied to the gray image. The notation of a macro block and sub-blocks is used in the proposed method. Figure 1 shows a notation of a macro block and 9 sub blocks.

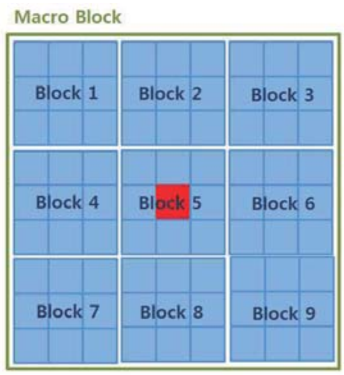

Figure 1. A macro block and 9 sub blocks

In Figure 1, a single macro block is composed of 3*3 sub blocks (Block 1~Block 9) and each sub block is again composed of 3*3 pixels. So a single macro block is composed of total 81 pixels. A macro block is a basic unit detecting a SRE feature. A kernel window for the macro block is moved on the image plane and horizontal and vertical edge intensity are computed.

In Equation (1), *HrzEdge* means the edge image which is filled with absolute value of horizontal edge and *VrtEdge* means the other edge image filled with absolute value of vertical edge. $HrzEdge_{Bi}$ is the averaged value of Block i in *HrzEdge* and $VrtEdge_{Bi}$ is the averaged value of Block i in *VrtEdge*.

$$\begin{aligned} Edge_{up} &= \frac{VrtEdge_{B2}}{\left(VrtEdge_{B1} + VrtEdge_{B3}\right)}, \\ Edge_{down} &= \frac{VrtEdge_{B8}}{\left(VrtEdge_{B7} + VrtEdge_{B9}\right)} \\ Edge_{left} &= \frac{HrzEdge_{B4}}{\left(HrzEdge_{B1} + HrzEdge_{B7}\right)}, \\ Edge_{right} &= \frac{HrzEdge_{B6}}{\left(HrzEdge_{B3} + HrzEdge_{B9}\right)} \end{aligned} \tag{1}$$

The center pixel in each macro block (the red rectangle in Figure 1) is recognized as a cross-shaped feature candidate if all of $Edge_{up}$, $Edge_{down}$, $Edge_{left}$ and $Edge_{right}$ in Equation (1) are higher than a predefined threshold value, $Edge_{th}$ and $CrsEdge_{center}$ in Equation (2) is higher than $CrsEdge_{th}$.

$$CrsEdge_{center} = \frac{Edge_{up} + Edge_{down} + Edge_{left} + Edge_{right}}{4} \tag{2}$$

Then the non-maximum suppression is applied to the feature candidates. And isolated features are also deleted because a mosaic region usually includes many cross-shaped features. If the number of detected SRE features is smaller than a predefined threshold value, $CrsFea_{th}$, the current input image is supposed to include no mosaic masking region and the overall process for the input image is left off.

In our implementation $CrsFea_{th}$, $Edge_{th}$ and $CrsEdge_{th}$ are configured as 3, 2.0, 5.0. Finally selected SRE features are saved in a list-typed data structure, FL (Feature List). At this moment, each data object in the list contains only x, y coordinates of the SRE feature but other parameters for the mosaic region and memory pointers to a neighboring SRE features will be saved in each data object.

In this feature detection process, not the edge intensity value itself, but the ratio of averaged edge value in each sub block is used and the characteristic of structured edge distribution is also used. So the proposed method is available for degraded input images in which features cannot be detected by a traditional feature detection method [3].

B. Parameter estimation

In this step, mosaic candidate regions are detected according to the feature density around the each SRE feature and the parameters for each mosaic candidate region are estimated. It is premised that all mosaic regions in an image have similar size. So a unique radius value for an input image is estimated. First, all of the SRE features are considered as a center point of mosaic candidate regions and the most plausible radius parameter is estimated.

$P(MCR \mid \mathbf{x}_c, r)$ is defined as the probability that the region defined by center position, xc and radius, r is a Mosaic Candidate Region (MCR). Our purpose is to estimate the optimum r when the current SRE feature coordinate, xc is assumed to be the center of a mosaic region. So the r value with the maximum $P(r \mid \mathbf{x}_c, MCR)$ should be estimated. By the Bayesian rule in Equation (3), optimum r is the value with the maximum $P(MCR \mid \mathbf{x}_c, r)$ if the $P(\mathbf{x}_c, r)$ is uniform in $[r_{min}, r_{max}]$.

$$\begin{aligned} P(r \mid \mathbf{x}_c, MCR) &= \frac{P(MCR \mid \mathbf{x}_c, r) P(\mathbf{x}_c, r)}{P(\mathbf{x}_c, MCR)} \\ &\propto P(MCR \mid \mathbf{x}_c, r) \\ &\quad \text{if } P(\mathbf{x}_c, r) \text{ uniform for } \{r \mid r_{\min} < r < r_{\max}\} \end{aligned} \tag{3}$$

To detect mosaic candidate regions, P(MCR | xc, r) can be approximated as a density function of neighboring SRE features as in Equation (4).

$$\begin{aligned} f_{density}\left(\mathbf{x}_c, r, FL\right) &= \frac{\sum_{\mathbf{x} \in FL} \sigma\left(\frac{\left\|\mathbf{x}_c - \mathbf{x}\right\|}{r}\right)}{\pi r^2} \\ \text{when } \sigma(x) &= 1 \text{ if } x < 1 \\ \text{else } \sigma(x) &= 0 \end{aligned} \tag{4}$$

If we find r with the highest density, the selected r tends to be smaller than the desirable value which includes all SRE features in a real mosaic masking region. If there is a mosaic region in the input image, detected SRE features may be concentrated inside the mosaic region and the desirable value for r is the minimum value including most of SRE features of the largest feature mass. This value can be estimated by calculating the rate of change for the SRE feature density while reducing r gradually. While the rate of increase is bigger than a threshold value, $K_{density}$ (>1), reducing r continues.

For each data element in a Feature List *(FL)*, if *FL(i).density* is equal or higher than a threshold, $th_{density}$ then the *i*-th SRE feature data, *FL(i)* is considered as a mosaic candidate region. If *FL(i).density* is lower $th_{density}$ then *FL(i)* is deleted in the list. The value for $th_{density}$ in our implementation is explained in section 4.

Next, the parameters for mosaic block arrangement are estimated. This process is repeated for all mosaic candidate regions in the *FL*. The central point of the mosaic candidate region and the width and the height of the each block are estimated in this process.

First, the offset point marked as a yellow point ($x_{i,offset}$, $y_{i,offset}$) in Figure 2 is estimated. The x coordinates of the neighboring features in *FL(i).list* are clustered and the coordinate with the highest frequency is selected as the $x_{i,offset}$. This process is also applied to y coordinates and $y_{i,offset}$ is determined.

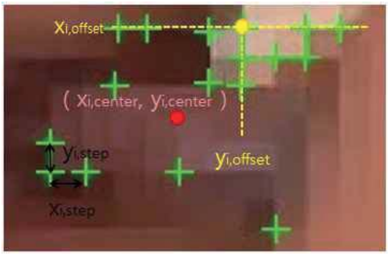

Figure 2. Arrangement parameters for a mosaic block

Next, the width and the height of a single block, $x_{i,step}$ and $y_{i,step}$ are estimated. The optimum $x_{i,step}$, $y_{i,step}$ are the values with the minimum residual error and they are searched by Equation (5). In Equation (5), x_j , y_j are the coordinates of the *j*-th neighboring feature in *FL(i).list* and N_i is the number of features in *FL(i).list*.

$$
\begin{aligned}
&\mathrm{x}_{\mathrm{i,step}} = \arg\min_{x_{step}} \sum_{j=1}^{N_i} \left| x_j - \left(x_{i,offset} + x_{step} * k_{x,i,j} \right) \right| \\
&\mathrm{y}_{\mathrm{i,step}} = \arg\min_{y_{step}} \sum_{i=j}^{N_i} \left| y_j - \left(y_{i,offset} + y_{step} * k_{y,i,j} \right) \right| \\
&k_{x,i,j} = \arg\min_{k} \left| x_j - \left(x_{i,offset} + x_{i,step} * k \right) \right| \\
&k_{y,i,j} = \arg\min_{k} \left| y_j - \left(y_{i,offset} + y_{i,step} * k \right) \right| \\
&\qquad when\ x_{i,step}, y_{i,step}\ is\ given
\end{aligned}
\tag{5}
$$

Finally, the central intersection point ($x_{i,center}$, $y_{i,center}$) of the mosaic candidate region is estimated. The point is the intersection point which is closest to center of mass of neighboring SRE features in *FL(i).list* and it is expressed as a red point in Figure 4. ($x_{i,center}$, $y_{i,center}$) is computed by Equation (6) and it may coincide with a SRE feature in *FL(i).list* or not. In Equation (6), ($x_{i,CM}$, $y_{i,CM}$) is the center of mass about the features in *FL(i).list*. These parameters are necessary for the next verification step. If only mosaic candidate detection based on the feature density is used and the verification is omitted, then this parameter estimation step can be omitted.

$$
\begin{aligned}
&\mathrm{x}_{\mathrm{i,center}} = x_{i,offset} + x_{i,step} * k_{i,x} \\
&\mathrm{y}_{\mathrm{i,center}} = y_{i,offset} + y_{i,step} * k_{i,y} \quad ,when \\
&k_{i,x} = \arg\min_{k} \left| x_{i,CM} - \left(x_{i,offset} + x_{i,step} * k \right) \right| \\
&k_{i,y} = \arg\min_{k} \left| y_{i,CM} - \left(y_{i,offset} + y_{i,step} * k \right) \right|
\end{aligned}
\tag{6}
$$

C. SRD verification

Now, mosaic candidate regions in FL and their parameters are prepared. The final step is to decide whether the candidate is a real mosaic region or just a mass of SRE features. The characteristic of Structured Ratio of luminance Deviation (SRD) in mosaic region is used for the verification in this step. Because of the degradation induced from low-quality compression or resizing, sometimes the horizontal and vertical edges in a mosaic region may be blurred and the interior region of each mosaic block (red boxes in Figure 3) also may not be perfectly uniform. But, if the mosaic candidate region is a real mosaic region, the standard deviation of luminance around each intersection point (yellow boxes in Figure 3) is supposed to be higher than the standard deviation around the center of each mosaic block (red boxes in Figure 3). The locations of intersection regions and interior regions can be estimated by parameters, ($x_{i,center}$, $y_{i,center}$) and ($x_{i,step}$, $y_{i,step}$) acquired from parameter estimation step.

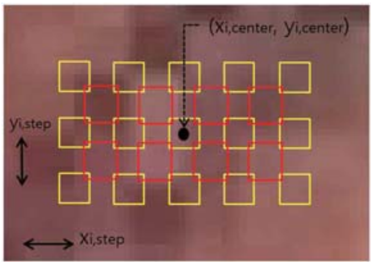

Figure 3. Intersection regions (yellow boxes) and interior regions (red boxes) inside a mosaic region

The SRD verification process is summarized as follows. For each mosaic candidate region in *FL*, the Ratios of averaged Luminance Standard deviation (*RatioLumStd*) in intersection regions and interior regions are calculated. A mosaic candidate region of *FL(i)* is determined as a real mosaic region if *FL(i).density* is higher than a threshold, $th_{density}$, and *FL(i).RatioLumStd* is higher than the other threshold, $th_{RatioLumStd}$. For the final mosaic region, ($x_{i,center}$, $y_{i,center}$) is the center point, and *r* is the radius of the mosaic region.

IV. EXPERIMENTAL RESULTS

As explained in section 2, the previous works [1], [2] are not comparable with the proposed method for mosaic region detection. So three modified versions from the proposed method were tried to analyze the usefulness of the proposed method. In the first method, only SRE feature density is used for mosaic region detection. In other words, only mosaic candidate detection explained in section 3-2 is used and verification step is omitted. The second method is to use only SDR verification and omit the mosaic candidate detection. In other words, all of *FL(i)* are regarded as mosaic candidate regions in the second method. The final method is to use both the mosaic candidate detection and SRD verification. For the experiment, 500 pornographic images including mosaic regions and 1000 non-obscene images without a mosaic region were used. The images are normalized to 640*480 size before SRE feature detection step. Table 1 shows the detection result of first method (OnlyDensity). True Positive Rate (TPR) is the ratio of correctly detected images in the pornographic images and False Positive Rate (FPR) is the ratio of incorrectly detected images in the non-obscene images. Table 2 and Table 3 are the detection results of second (OnlySRD) and the third (DensitySRD) methods. The computation time is 0.01 sec/frame including the image normalization in Intel 3.0 GHz CPU without any parallel processing. Figure 4 shows the ROC curve of those methods. It is remarkable that the accuracy of

OnlyDensity is better than those of other methods in low FPR range. High TPR of OnlyDensity in the range is induced from the fact that the proposed SRE feature is designed for the characteristic of a cross-shape feature in a mosaic region. So mosaic region can be detected simply by detecting SRE features and computing their density when very low FPR (lower than 5%) is needed. But to get high TPR (higher than 95%), SDR verification enhance the accuracy as in Figure 4.

Figure 5 shows the examples of True Positive (correctly detected) and False Negative (incorrectly excluded) mosaic regions with TPR, 96.2%. Only the mosaic regions severely blurred (Figure 5 (c), (d)) were not detected by the proposed method. Figure 6 shows examples of False Positive (incorrectly detected) images with FPR 12.4%. Non-mosaic regions were incorrectly detected in a check-board region or a regularly repetitive edge region around characters. This comes from that the proposed method depends on characteristics of cross-shaped features and the regular arrangement of the features. Incorrectly detected regions in Figure 6 are similar to real mosaic regions in the aspect of these characteristics. When only the half of the real mosaic masking regions are detected, FPR is reduced to the value lower than 1% and most of these clutters are disappeared.

V. Conclusion

The main contribution of this paper is that a novel feature detection method (SRE) optimized for cross-shaped intersection points in a mosaic masking region, and a new verification method using the characteristic of luminance deviation in structured sub regions (SRD) are proposed. Mosaic candidate regions are generated by computing the density of SRE features and the verification enhance the accuracy in high TPR range of ROC curve. By using only mosaic candidate detection process computing only the density of SRE features, practical detection result can be acquired because the SRE feature itself is optimized for the cross-shaped features which can be observed in a real mosaic region.

Table 1: Result with only candidate detection.

$th_{density}$	True Positive Rate	False Positive Rate
0.00005	94.8	11.7
0.00004	95.8	14.6
0.00003	96.6	16.2
0.00002	97.2	17.2

Table 2: Result with only SRD verification.

$th_{RatioLumStd}$	True Positive Rate	False Positive Rate
10	4.8	0.2
2	55	1.8
1	96.6	13.2

Table 3: Result with candidate detection and SRD verification

$th_{RatioLumStd}$	$th_{density}$	TruePositiveRate	FalsePositiveRate
10	0.00005	4.8	0
2	0.00005	54	1.4
1	0.00005	93.8	9.6
1	0.00003	95.6	11.4
1	0.00002	96.2	12.4
1	0.00001	96.6	13.2

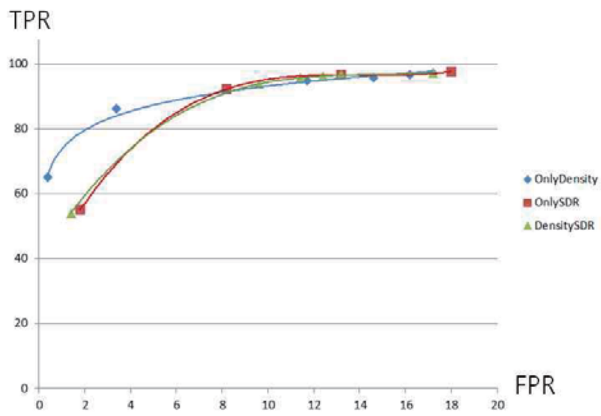

Figure 4. ROC curve of various mosaic detection modes

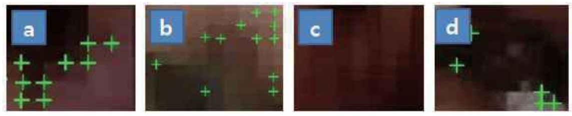

Figure 5: The examples of (a), (b) True Positive and (c), (d) False Negative mosaic masking region.

Figure 6: The examples of False Positive images from non-obscene images

Acknowledgment

This work was supported by the ICT R&D program of MSIP/IITP. [13912-06006, Development of the Filtering Technology for Objectionable Streaming Contents on Smart Platform]

References

[1] S. Shui-Fa1, H. Shu-heng, W. Gang, X. Yi-chun, L. Bang-Jun. “Mosaic Defect Detection in Digital Video,” Chinese Conference on Pattern Recognition, October, 2010.

[2] Z. Wei, J. Lin, L. Zhang and S. Song. “Mosaic Defect Detection Based on Macro Block Solid Edge Detection,” Research Journal of Applied Sciences, Engineering and Technology, 5(1): 3549-3553, 2013.

[3] J. Shi. “Good Features to track,” IEEE Conference on Computer Vision and Pattern Recognition, 1994

[4] P. Viola, M. Jones. “Rapid Object Detection using a Boosted Cascade of Simple Features,” IEEE Conference on Computer Vision and Pattern Recognition, 2001.

[5] A. Mittal, A. Zisserman. P. H. S. Torr. “Hand detection using multiple proposals,” British Machine Vision Conference, 2011.

[6] P. Dollar, C. Wojek, B. Schiele, P. Perona. “Pedestrian Detection: A Benchmark,” IEEE Conference on Computer Vision and Pattern Recognition 2009.