

RESEARCH

Open Access



Enhanced resampling detection based on image correlation of 3D stereoscopic images

Hak-Yeol Choi, Dai-Kyung Hyun, Sunghee Choi and Heung-Kyu Lee*

Abstract

In this paper, we propose a resampling detection method for stereoscopic images. Although previous resampling techniques can be applied to stereoscopic images, performance improvement is hard to be expected with the two separated results. In this research, we found a strong relationship between the left and right images derived from the characteristics of the stereoscopic images. The proposed technique exploits that relationship of the stereoscopic images as additional information for reliable detection performance. Furthermore, the proposed method includes a preprocessing step to acquire the independent performance from the image's own characteristics. The experimental results exhibit superior performance compared with the existing works.

Keywords: Multimedia forensic, Resampling detection, Stereoscopic images, Visual fatigue

1 Introduction

With recent rapid developments in technology, digital content can now be generated anytime and anywhere by electronic devices such as smartphones, digital cameras, and CCTVs. At the same time, a significant volume of this content is being shared through various media such as social networks and broadcasting channels. Some content is used for important purposes such as evidence in court or as a medium for communicating political issues. If such important digital content is illegally tampered, it may cause severe socio-economic loss.

In the past, suspicious modulation of content was analyzed by human judgment. However, computer graphics technology has been advanced tremendously over the decades, and distinguishing a tampered image from an original image has become a harder job.

Multimedia forensic is a set of technologies to overcome the limitations of human-based forgery analysis. Since multimedia forensics relies on underlying statistical properties to reveal tampering, the forgery can be detected even if the traces are not visible to human eyes. Also, power of computer allows tens of thousands of content files to be analyzed automatically.

There are various types of image tampering processes which have been considered in the multimedia forensic field. Among them, the resampling process is one of the most important issues. When several images are exquisitely spliced together, a resizing is typically applied. The resizing includes a resampling process as a core step. Therefore, it is possible to discriminate the image manipulation by detecting the fingerprint of the resampling process.

Today, stereoscopic content is a rising keyword in the content market. Stereoscopic technology allows the viewer to experience a three-dimensional effect by displaying slightly different views to corresponding left and right eyes. Because new stereoscopic content is constantly being generated, the size of the stereoscopic market is growing exponentially.

Despite the advantages of stereoscopic content, the stereoscopic technology has visual fatigue as a fundamental limitation. To overcome this limitation, a multi-view technology has been developed. In the near future, multi-view content is expected to replace stereoscopic content.

Also, the need for protection of the multi-view content is big. Since the device for generating multi-view content have a number of lens and the processor have high performance, the price for generating multi-view content is more expensive than the monoscopic one. Nevertheless,

*Correspondence: heunglee@kaist.ac.kr
School of Computing, Korea Advanced Institute of Science and Technology,
291, Daehak-Ro, Yuseong-gu, Daejeon 34141, Republic of Korea

forgery protection technology for multi-view content has rarely been developed.

Multi-view content protection technology has advantages over the technology for monoscopic image in that it can use multiple images to detect forgery. But it is hard to realize the benefits without the consideration of the characteristics of multi-view content. If the existing resampling detection technique is repeatedly applied to the multi-view content without consideration of the characteristics of the multi-view content, only the unrelated results will be obtained. However, the method cannot utilize the relation among images in the multi-view content. Therefore, it is hard to expect sufficient improvement in detection performance.

The multi-view image can be considered an extension of stereoscopic image. It means that the forensic work for stereoscopic images can naturally be extended to the multi-view content. Therefore, we conducted this study as a fundamental research for developing the ultimate protection technology for next generation multi-view content.

In this paper, we propose an enhanced resampling detection algorithm for stereoscopic 3D images. In the proposed method, we apply a novel filtering method which separates the resampling detection process from the image's own properties. The conspicuous regions filtering (CRF) process allows the detector to work well with low resampling factors and various types of image sets. Moreover, we developed a correlation-based stereo-signal synthesizing (CSS) scheme to properly use the relation of stereoscopic images as side information.

This paper is organized as follows: in Section 2, we look into the previous resampling detection schemes and main contributions of this paper. In Section 3, we introduce the resampling process as well as the periodicity that exists in the resampled signal. In Section 4, we investigate the characteristics of forgery with stereoscopic images. In Section 5, we propose the resampling detection technique for 3D stereoscopic images, followed by the experimental results in Section 6. Finally, the Conclusion is drawn in Section 7.

2 Related work and main contributions

2.1 Related work

Many resampling detection methods have been proposed in the past. Even though there are various ways to expose the resampling process, most techniques commonly use the relation between pixels. Studies related to resampling detection method using monoscopic image is as follows:

Popescu and Farid developed a resampling process by analyzing invisible correlations from the resampled signals [1]. Using an expectation/maximization (EM) algorithm, they determined that samples are periodically

correlated with their neighbors in the same way. Gallagher exposed the resampling process by investigating the periodicity of the second differentiation of interpolated signals [2]. The method employs discrete Fourier transformation (DFT) to average the covariance from the result signal of second differentiation. Gallagher also estimated the resampling factor by explicating the direct relation between the periodicity and the resampling factor.

Mahdian and Saic improved Gallagher's work [2] by applying Radon transformation [3, 4]. They estimated the factor not only for resizing but also for other geometric transformations, like rotation and skewing. Moreover, they showed the periodic property hidden in the covariance of second derivatives. Kirchner also improved Popescu's work [1] to have lower complexity [5]. To lower the computational complexity, Kirchner replaced the estimation of prediction weight with a linear filter using a fixed-filter coefficient. Kirchner and Gloe proposed the resampling detection method for recompressed images [6]. They used blocking artifacts in recompressed images to improve the performance of resampling detection. Kirchner improved Popescu's work [1] by showing how a specific image structure could be modeled with not only a single predictor but with a series of linear predictors [7]. After that, they revealed the resizing by investigating the periodicity correlation among pixels.

Feng et al. proposed a resampling detection scheme based on examining the normalized energy density present within windows of varying size in the second derivative of an image in the frequency domain [8]. Vazquez-Padin et al. proposed a new type of resampling detector based on set-membership estimation theory [9]. The method estimates the resampling factor based on the effect caused by the quantization of interpolated samples. In addition, Vazquez-Padin et al. proposed a resampling detection algorithm for an upsampled image [10]. They employed SVD to unveil the presence of linear dependencies. And then, they measure the degree of saturated pixels per row/column.

Despite the many advantages, the existing works still have some points to be improved. The first one is that the detection rate dropped seriously when performing with a compressed image, since image compression weakens the resampling periodicity such images have. The second one is that the previous works mentioned above have the tendency to lose performance with specific images that include highly textured region since the traces of resampling are dependent on an image's own properties [11].

Furthermore, these previous works did not consider the stereo environment at all, and so do not accommodate the unique characteristics of stereoscopic 3D images. Nonetheless, the limitations of existing technologies can

be overcome by appropriately utilizing the characteristics of 3D stereoscopic images.

In the multimedia forensics field, only few approaches in multimedia forensics considered correlation among sources. Almost previous forensic works just considered single sources [1–5, 7, 12, 13]. Lately, the forensic works which consider the correlation between sources are proposed [14–17]. The multi-track-based forensic algorithm can be regarded as a new way to overcome the limitations of existing forensic works. By using correlated multi-track sources, the forgery can be revealed based not only on an image's own statistical properties, but also a correlation among several sources. The related work on multi-track-based forensic techniques include the following:

In the prior work, we proposed a resampling detecting method for depth-image-based rendering (DIBR) images [14]. We exposed the resampling process by using the correlation between the center image and the depth map. However, the method did not use the relation of both images efficiently. Moreover, the performance is dependent on the image's own properties, like the previous resampling detectors.

Fouche and Oliver proposed a splicing detection scheme using internal depth of stereoscopic images [15]. In this work, they showed how the relationship between the distance of an object and its internal depth can aid the detection of spliced stereoscopic images. Kirchner et al. proposed the concept of detecting photo tampering using the camera preview or the camera motion [16]. They revealed the forgery by measuring the consistency of the preview and motion data. Comesana and Perez-Gonzalez quantified the advantages of using the joint distribution of composite objects for improving the distinguishability between processing operators [17].

2.2 Main contributions

The proposed method has three main contributions. First, we developed a resampling detection scheme for stereoscopic 3D images. To develop the method, the strict relation between both the left and the right images were investigated in Section 4. Here, we analyzed how much human feels visual fatigue in various inconsistency factors. Based on that fact, the constraints of the image manipulation

that existed only in the stereoscopic images were dealt with.

Second, the previous works have the tendency to lose detection performance by the edge components of the target image. To minimize the influence of the edge component, we designed CRF process. This not only maximizes performance, but also minimizes the dependence between detection performance and the properties of the image itself.

Third, the proposed scheme does not use a single forged image, but uses both the left and the right images of stereoscopic 3D images. The research shows how resampling detection algorithms can be applied to multi-track content. The proposed method introduces a novel CSS process which properly utilizes the relation of both the left and the right images. The CSS process improved the resampling detection performance considerably by maximizing the statistical properties of resampled stereoscopic images.

3 Preliminary

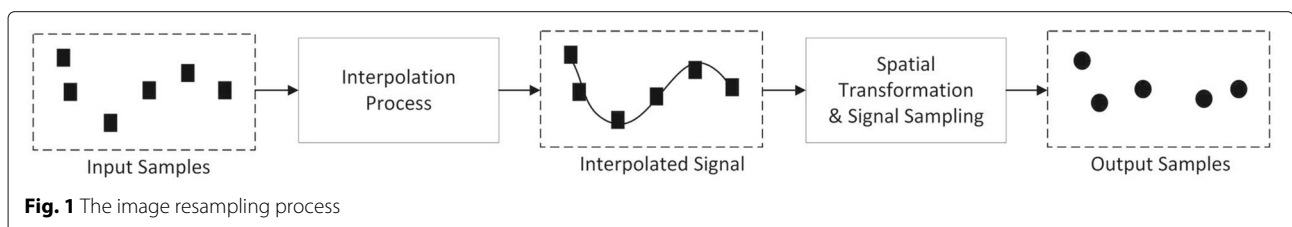
3.1 Resampling process

Image resampling is a process that transforms a coordinate system of a sampled image [18]. Both coordinates are related with a spatial transforming mapping function. The result of transforming the mapping function is a resampled grid. In other words, the input signal is resampled into a new location by the mapping function.

The image resampling process can be divided into two processes: a sampling process and an interpolation process followed by the sampling process. The entire image resampling process is described in Fig. 1 as a one-dimensional case. The process of interpolation is performed by convolving the discrete signal with a continuous interpolation kernel. The value of the interpolated signal is the same as the sum of the values of the input signals scaled by the corresponding values of the interpolation kernel. The continuous interpolation kernel then can be sampled by the resampling grid. In this way, the input image is assigned to the new coordinate.

3.2 Periodic property of resampled signal

In most cases, the resampling process causes a hidden periodicity in an image [2]. Since the periodicity exists



in the resampled region of the image, it can be found in the result of the variance of n -th order differentiation of the image signal. Also, the resampling factor can be estimated based on the periodicity of the signals. Here, we introduce a resampling detection method [4] based on that facts. As we mentioned in Section 3.1, the resampling process is composed of main two functions: image sampling and image reconstruction. Also the image reconstruction is performed with an interpolation kernel. Therefore, we can conclude that the resampling process causes the hidden periodicity only if we prove the periodicity came from the interpolation process.

Let the interpolated signal be i and an original signal be a . Then, the relation of interpolated signal and original signal could be described as follows:

$$i(x) = \sum_{q=-\infty}^{\infty} a(q\Delta_x) h\left(\frac{x}{\Delta_x} - q\right) \quad (1)$$

where h is an interpolation kernel and the resampling step $\Delta_x \in \mathbb{R}$, and a constant $q \in \mathbb{Z}$. In real application, the kernel h will be considered as low-order local polynomial such as linear, cubic. Also, the range of q will be determined according to the size of kernel h .

Let $D^n\{x\}$ as a n -th order partial differentiation of signal x . Then $D^n\{x\}$ will be x when $n = 0$. With that notation, the result of n -th order partial differentiation of interpolated signal i could be represented as follows:

$$D^n\{i\}(x) = D^n \left\{ \sum_{q=-\infty}^{\infty} a(q\Delta_x) h\left(\frac{x}{\Delta_x} - q\right) \right\} \quad (2)$$

Also, Eq. (2) could be reconstructed by convolution of h with a differentiation kernel based on a derivative theorem and a convolution theorem.

$$D^n\{i\}(x) = \sum_{q=-\infty}^{\infty} a(q\Delta_x) D^n\{h\}\left(\frac{x}{\Delta_x} - q\right) \quad (3)$$

When a signal has stationary property with a variance σ^2 , the variance of Eq. (3) at position x could be described as Eq. (4).

$$\text{var}\{D^n\{i\}(x)\} = \sigma^2 \sum_{q=-\infty}^{\infty} D^n\{h\}\left(\frac{x}{\Delta_x} - q\right)^2 \quad (4)$$

It is easy to show that the variance of n -th order differentiation has the periodicity with period $k\Delta_x$. Following Eq. (5) proves that fact.

$$\begin{aligned} \text{var}\{D^n\{i\}(x + k\Delta_x)\} \\ &= \sigma^2 \sum_{q=-\infty}^{\infty} D^n\{h\}\left(\frac{x + k\Delta_x}{\Delta_x} - q\right)^2 \\ &= \sigma^2 \sum_{q=-\infty}^{\infty} D^n\{h\}\left(\frac{x}{\Delta_x} - (q - k)\right)^2 \\ &= \text{var}\{D^n\{i\}(x)\} \end{aligned} \quad (5)$$

Therefore, the periodicity could be revealed by examining the variance.

Also, the extension into the two-dimensional representation is straightforward. Eq. (6) shows the representation.

$$\begin{aligned} \text{var}\{D^n\{i\}(x, y)\} \\ &= \sum_{q_1=-\infty}^{\infty} \sum_{q_2=-\infty}^{\infty} D^n\{h\}\left(\frac{x}{\Delta_x} - q_1, \frac{y}{\Delta_y} - q_2\right)^2 \end{aligned} \quad (6)$$

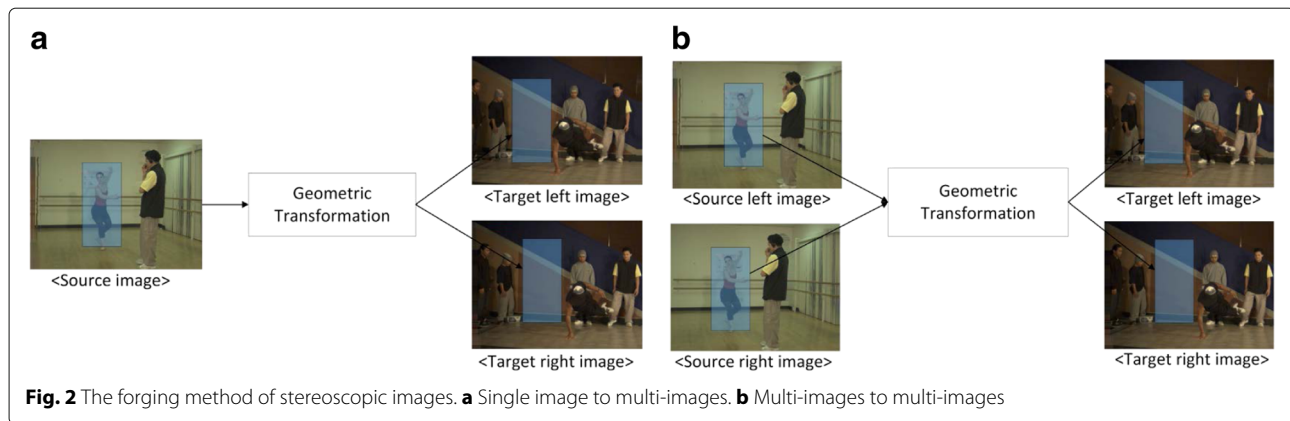
An image is composed of the two-dimensional signals, the periodicity could be appeared in both the row and the column of signals. With an image, the range of summation will not be infinite as Eq. (6) since an image is finite discrete signal. But it does not alter the fact that the image have periodicity even if it has been resampled.

4 Forgery on stereoscopic images

Before we discuss the proposed scheme further, we need to discuss the forgery process of stereoscopic 3D images in terms of visual fatigue. The forgery of stereoscopic images can be conducted in two ways [15]. The first one is a way that uses the sub-block from a monoscopic image. In order to obtain high-quality forged images, the sub-blocks are often geometrically transformed. After that, the sub-block is synthesized to the target images. Figure 2a describes the first forgery method of stereoscopic images. The second one is a way that uses the sub-blocks from both the left and the right images of stereoscopic images. Figure 2b describes the second forgery method of stereoscopic images.

The stereoscopic images have clear differences compared with monoscopic image in terms of the coexistence of both the left and the right images. That characteristic is related to several constraints that have to be followed during the forgery process.

One of the most important issues in stereoscopic images is visual fatigue. Without a precise viewing condition, a viewer may feel discomfort. There are several causal factors of visual fatigue [19, 20]. Among them, excessive binocular disparity, motion characteristics, binocular asymmetry, and depth cue conflict are the most noticeable factors. Also, the binocular asymmetry can be departmentalized as a geometrical asymmetry, luminance asymmetry, chrominance asymmetry, and window edge violation.



Out of those factors, we focused on geometrical asymmetry since the resampling process is related to geometric distortion. Geometrical asymmetry means the desynchronization of both the left and the right images. The normal synchronization of stereoscopic images requires the consistency of both images in terms of geometric shape, such as the size or the angle of the object.

Once the consistency is broken, viewers begin to feel strong visual fatigue [19, 20]. From the viewpoint of manipulation, the faulty forgery will cause serious visual fatigue. Then, the viewers might easily recognize the trace of the forgery. Therefore, the manipulator must satisfy the constraints during forgery.

To prove that the visual fatigue is caused by geometric asymmetry, we adopted a double-stimulus continuous quality-scale (DSCQS) investigation as a subjective analysis. For the experiment, 30 expert test subjects participated. The overall perceived quality was rated based on a quality categorical scale from 1 (bad) to 5 (excellent) using the criteria of mean opinion score (MOS). The test subjects entered scores on the scoring table described in Table 1, and all of the scores were averaged to draw a conclusion. The display device used was LG LED 3D TV (27MT93D). In the experiments, we used static images taken by camera.

The number of types of geometrical distortion that can occur in a resampling process is two. The first one is the difference of vertical scaling between the left and the right

images. The visual fatigue caused by the difference of horizontal scaling is known to be much less than the difference of vertical scaling. The second one is the difference of vertical alignment between the left and the right images. The horizontal alignment cannot be regarded as a causal factor of visual fatigue since it is an essential element for expressing depth information in a stereoscopic image.

The experiments were conducted considering both types. The first experiment was conducted to test visual fatigue caused by a difference in vertical scaling. In the following second experiment, visual fatigue caused by a difference in vertical alignment was tested.

In the first experiment, the test subjects scored the visual fatigue when the vertical scaling factor of the right image changed from 0.85 to 1.15 in comparison with the fixed left image. Figure 3a shows the result obtained in the first experiment. The average score is in the middle between 1 (bad) and 2 (poor), which means that the test subjects felt severe visual fatigue even with a slight difference in the scaling factors of both the left and the right images.

In the second experiment, the test subjects scored the visual fatigue when the vertical position factors of the right image changed from 1 to 5% in comparison with the fixed left image. Figure 3b shows the result obtained in the second experiment. The average score is in the middle between 1 (bad) and 2 (poor), which means that the test subjects felt acute visual fatigue in spite of a slight difference in the vertical position of both images.

Through these experimental results, we deduced that two constraints inevitably occur in the forgery of stereoscopic image. The constraints are applied in both forgery ways explained above. The constraints of the forgery process in stereoscopic images are as follows.

- The first constraint is that the forged regions have to be matched for both images. The position of an object in both the left and the right images must have the same vertical position. Based on depth

Table 1 The evaluation scoring table of DSCQS test

Evaluated quality	Score
Excellent	(5)
Good	(4)
Fair	(3)
Poor	(2)
Bad	(1)

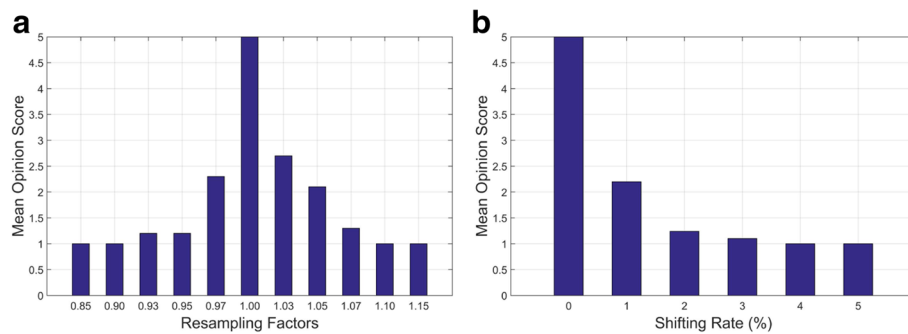


Fig. 3 The results of visual fatigue experiments. **a** The vertical scaling test. **b** The vertical alignment test

information, only slight differences in horizontal position may be allowed.

- The second constraint is that the resampling factors used in forgery have to be the same for both the left and the right images. If different resampling factors are used for both images, it will cause visual fatigue since the size of objects in both images will be different.

Even if only one of the constraints explained above is violated, the quality of the forged image will be excessively damaged. The proposed technique is designed to assume that the constraints are followed, since the constraints are consistent for every forgery situation involving stereoscopic images.

On the other hand, it is hard to detect forged images satisfying above two constraints since they do not cause visual fatigue. Figure 4a, b are forged images maintaining vertical position and scale results by Fig. 2a way. Figure 5a shows the results of visual fatigue viewers feel according to horizontal position change. The results mean that viewers do not feel discomforts when horizontal position is changed. The difference of horizontal position is the reason for a three-dimensional effect of stereoscopic images.

Also, Fig. 4c, d are forged images by Fig. 2b way. Figure 5 shows the results of what visual fatigue viewers feel according to horizontal position change. The result is the same with that of Fig. 5a. The results imply that the forgery of stereoscopic image is easy if some



Fig. 4 Example forged images, (a) and (b) image pair forged by Fig. 2a method, (c) and (d) image pair forged by Fig. 2b method

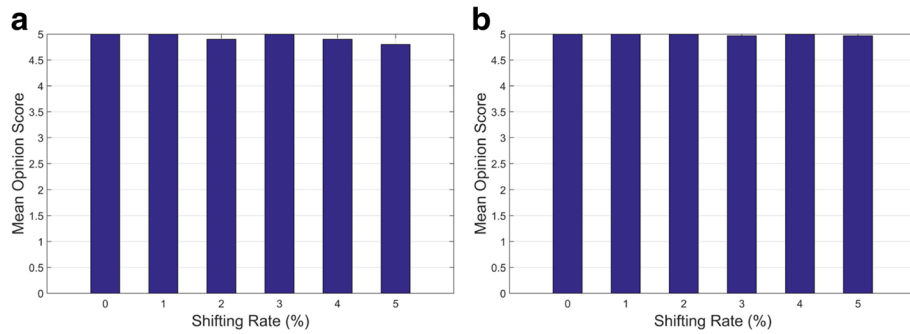


Fig. 5 The results of visual fatigue experiments according to *horizontal* shifting. **a** The result on Fig. 4a, b image pair. **b** The result on Fig. 4c, d image pair

conditions are satisfied. But, it is hard to detect the forged image keeping the constraints by human visual system. Therefore, the proposed technique aims to detect the forgery in a situation where the constraint is satisfied.

5 Proposed method

Figure 6 represents the entire process of the proposed method. The entire process can be largely divided into

three phases. In the periodicity exposing phase, the image blocks from each left and right image are pre-processed to acquire enhanced performance, using the CRF process mentioned before. In the following signal synthesizing phase, two signals are synthesized by the proposed CSS scheme. After the CSS process, the energy of periodicity is quantified by calculating peak to correlation energy (PCE). PCE quantifies an energy of peak

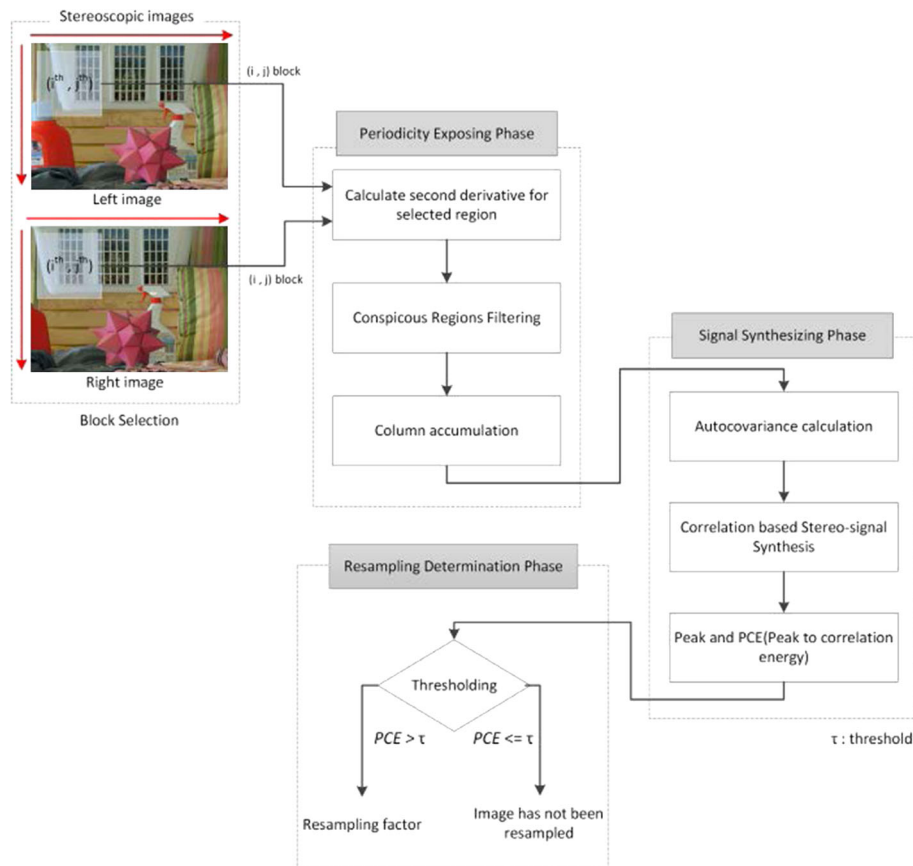


Fig. 6 The entire process of the proposed method

signal comparative to the signal nearby peak. In the last resampling factor determination phase, the resampling process and the factor used are determined. The proposed technique is based on the constraints mentioned in Section 4. Above three phase could be represented in seven steps. Each of the steps will be introduced in detail in the following segments:

1. *Step 1 The second-order differentiation calculation:* for the first process of the proposed method, both blocks from left and right images are extracted. For the suspicious object, the blocks could be selected manually according to the location of object. In this case, a horizontal index could be changed since stereoscopic images have difference of horizontal index of left and right images for same object. After the block selection, as mentioned in Section 3.2, the resampling process can be revealed by examining the periodicity hidden in images. The hidden periodicity can be revealed in the n -th order differentiation of the images. In this paper, we use the second-order differentiation since it could expose the periodicity of not only the linear interpolation but also of the non-linear interpolation [2]. Therefore in the first step, we calculate the second differentiation for each selected region as Eq. 2.

2. *Step 2 Conspicuous region filtering:* there is a hidden periodic property in the second-order differentiation signal calculated in Step 1. However, the strength of the periodicity is not strong enough to be detected. Moreover, the weak periodicity is hidden by conspicuous regions in the second-order differentiation. The conspicuous regions are the regions which have high energy in the second-order differentiation. The conspicuous regions come from the edge component since second-order differentiation filters out the energy from the flat regions of the image.

Because the conspicuous regions have high energy irrespective of the periodicity, conspicuous regions in the second-order differential signal may lower the performance of resampling detector. The phenomenon is particularly noticeable when an image is resampled by low resampling factor, especially downsampling. Moreover, the detection performance is highly dependent on each image because the conspicuous regions are unique properties for each image.

Figure 7b shows the second-order differential result of the selected image block Fig. 7a. The edge components still remain even if the second-order differentiation is calculated. Since the almost

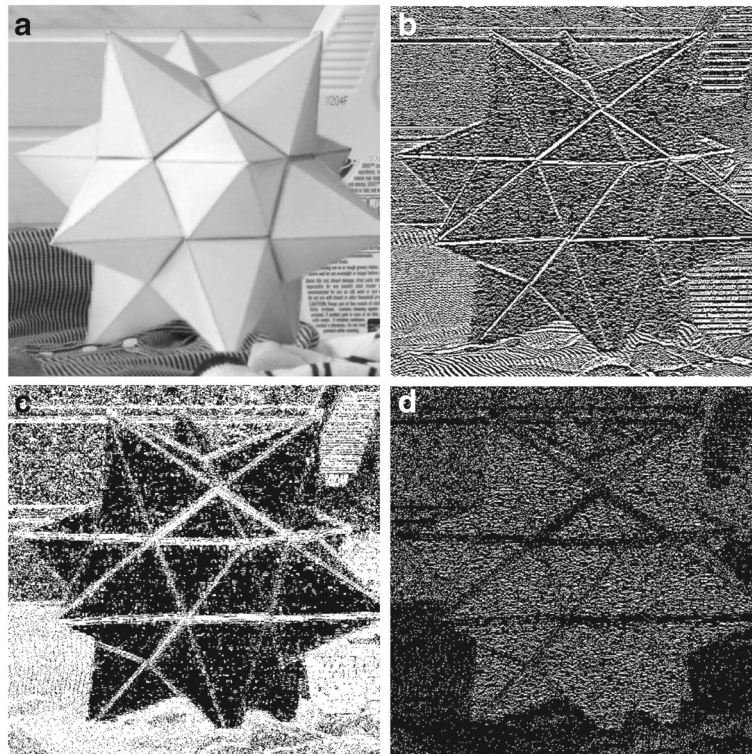


Fig. 7 An example of conspicuous region. **a** A sample image. **b** A second-order differentiation map. **c** A conspicuous region. **d** A map filtered by CRF process

conspicuous regions are not from the periodic pattern but from the edge components in the images, the conspicuous regions disturb the detecting process. Figure 7c shows an example of the conspicuous regions.

To solve those problems, we developed a preprocessing step named CRF to eliminate the conspicuous regions that exist in the second-order differential signals. At first, the distribution of the second-order differentiation map is composed. Then, the threshold value which will be used for filtering is calculated. Finally, the values in the differentiation map over the threshold value are set into 0. Figure 8 shows the flow of the CRF. Figure 9 shows some examples of conspicuous region-filtered images. The filtering of differentiation value in mainly edge region can be observed.

Although CRF process is simple, it provides useful advantages. First, resampling can be exposed even if low resampling factors are used. Likewise, the performance of detection is improved with JPEG-compressed images. Second, with the CRF process, the detection performance is no longer dependent on the image's own properties. That means reliable detection results can be acquired with various types of images. Finally, the CRF process ultimately helps the periodic component in synthesizing process which will be introduced in Step 5. In the synthesizing process, only components that come from the periodicity can be used for synthesizing by the CRF process.

3. *Step 3 Column accumulation:* in the second-order differential signals calculated in Step 1, the periodic property is hidden. Theoretically, the resampling process can be revealed by examining the second-order differential signals. However, in a real image, each periodic pattern is too weak to expose the resampling process for the noise coming from the image itself and other post-processing such as JPEG compression.

To solve the problem, a column accumulation process is applied. Each of the second-order differential

signals that have passed the CRF process in Step 2 is accumulated into a single column signal S . Eventually, column signals with weak periodicity are accumulated to generate a single signal which has a periodicity strong enough to determine the resampling process.

4. *Step 4 Auto-covariance calculation:* as mentioned in Sec. 3.2, the periodic property hidden in resampled images can be revealed by examining the variance of the signal. The variance is calculated among neighboring signals in the single column S . Therefore, the auto-covariance is calculated as below:

$$R_S(\xi) = \sum_i (S(i + \xi) - \bar{S}) (S(i) - \bar{S}) \quad (7)$$

where R_S means the auto-covariance of the signal S .

5. *Step 5 Correlation-based stereo-signal synthesis:* we have two signals from each left and right image calculated in Step 4. In this step, we use the characteristics of the stereoscopic images as mentioned before. The property produce some constraints in the manipulation as described in Sec. 4. To use the relation of both the left and right images, the CSS scheme is applied. We applied the CSS scheme to the signals calculated in Step 4 from both the left and the right images. In the CSS process, both signals are synthesized based on the correlation between the left and the right images.

As mentioned in Sec. 4, the resampling factors used in both left and right images have to be similar to guarantee the consistency of the forged images. Since almost the same resampling factors are used for both left and right images, the normalized positions of the interpolation peak from both images have to be sufficiently close. In experiments, the probability that peak of left and right images have same position was over 96.1%. In the same way, the probability is low for the case that the normalized position is the same in spite of both images are not resampled. The probability that the left and the right images have peak that is same positions without resampling process was 3.2%. The CSS process is conducted

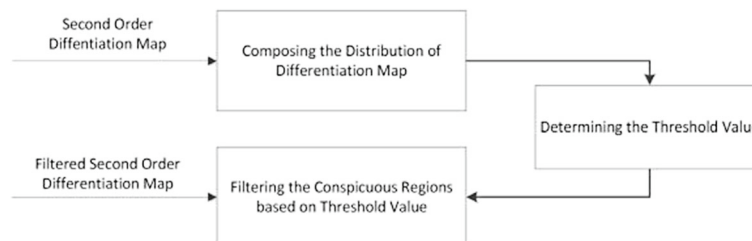


Fig. 8 The flowchart of the CRF process

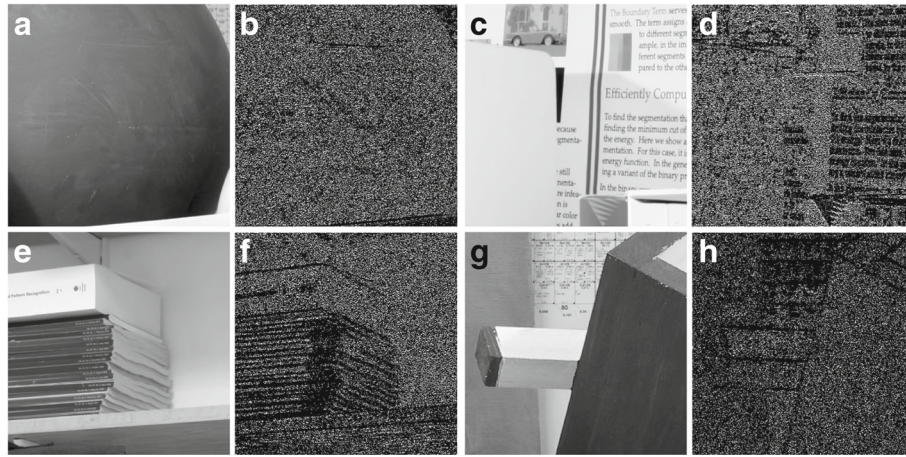


Fig. 9 The examples of conspicuous region-filtered images. Figure (a), (c), (e), and (g) shows original images and Fig. (b), (d), (f), and (h) shows the corresponding conspicuous region-filtered images

based on that fundamental. The entire CSS process is described in Algorithm 1 in detail.

In Algorithm 1, R_l , R_r are the auto-covariance signals from both the left and the right images. F_l , F_r are the Fourier transformed signals from the left to the right, respectively, and F_s is the synthesized signal from the

signals F_l to F_r . w is the weighting kernel and l_w is the size of the partial weighting window. In Algorithm 1, the threshold for checking a \hat{d} value has to be small enough to minimize the false positive error. In this work, the threshold was determined as 10. By using the weighting kernel, the subtle distinction in the

Algorithm 1 Correlation-based stereo-signal synthesizing

```

1: procedure CSS( $R_l, R_r$ )
2:    $F_l \leftarrow FFT(R_l)$ 
3:    $F_r \leftarrow FFT(R_r)$ 
4:
5:    $\hat{d} \leftarrow i_{\text{peak}}(F_l) - i_{\text{peak}}(F_r);$ 
6:   if  $\hat{d} < \tau$  then
7:      $\text{cirshift}(F_r, \hat{d});$ 
8:   end if
9:
10:  for  $n := 1$  to  $l_F$  do
11:     $t_l \leftarrow 0;$ 
12:    for  $k \leftarrow -\lfloor l_w/2 \rfloor$  do  $\lfloor l_w/2 \rfloor$ 
13:       $t_l \leftarrow t_l + F_r(n - k) * w(k);$ 
14:    end for
15:     $\hat{F}_l(n) \leftarrow F_l(n) * t_l;$ 
16:
17:     $t_r \leftarrow 0$ 
18:    for  $k := -\lfloor l_w/2 \rfloor$  to  $\lfloor l_w/2 \rfloor$  do
19:       $t_r \leftarrow t_r + F_l(n - k) * w(k);$ 
20:    end for
21:     $\hat{F}_r(n) \leftarrow F_r(n) * t_r;$ 
22:
23:     $F_s(n) \leftarrow \hat{F}_l(n) + \hat{F}_r(n);$ 
24:  end for
25: end procedure

```

$\triangleright i_{\text{peak}}(F)$ is an index of peak in signal F
 $\triangleright \tau$ is a threshold to check a \hat{d} value
 $\triangleright \text{cirshift}$ is a function for circular shift of a signal
 $\triangleright l_F$ is a length of signal F_l and F_r
 $\triangleright *$ operator represents convolution

interpolation process between both images can be reflected properly. A simple triangular kernel is used as weighting kernel in the experiments.

The input of an Algorithm 1 is the auto-covariance signals from the left and the right images. And output is single signal synthesized by CSS process. In the Algorithm 1, the Fourier transformation of both signals are calculated to find out the periodicity of signals. After that, the difference of indices of peaks from the left and the right signals are compared with threshold τ . Only if the difference is under the threshold, the right signal F_r is circular shifted to synchronize with the left signal. Then, both signals are synthesized by applying a partial weighting kernel. Each signal is multiplied with another signal which convoluted with partial weighting kernel with size l_w . The final F_s value is computed by adding the results from left and right signal above. In this step, the signal with same indices are united with the neighboring signals.

With the CSS process, there is an amplifying effect for the intensity of the peak signal. In a later part, the peak in the Fourier transformed signal is amplified if the stereoscopic images have been resampled. On the

other hand, the peak of signals are weakened if the stereoscopic images are not resampled.

In many cases, the images have their own periodic property irrespective of the resampling process. It means that a peak from the Fourier transformed signal may be detected in a non-resampled image. For that reason, almost previous resampling detection methods have a high false positive rate. However, the performance of the detector can be improved with the CSS process for stereoscopic images.

As it is well known, the periodicity in the signal is represented as a peak in the frequency domain. The location of the peak in a Fourier transformed signal is dependent on the resampling factor itself. It means that the resampling factor can be determined by examining the location of the peak. The resampling factor N can be calculated with the information of the location of the peak, as follows.

$$N = \frac{1}{p} \quad \text{or} \quad N = \frac{1}{1-p} \quad (8)$$

where p is the normalized position of the interpolation peaks from the Fourier transformed signal F_s .

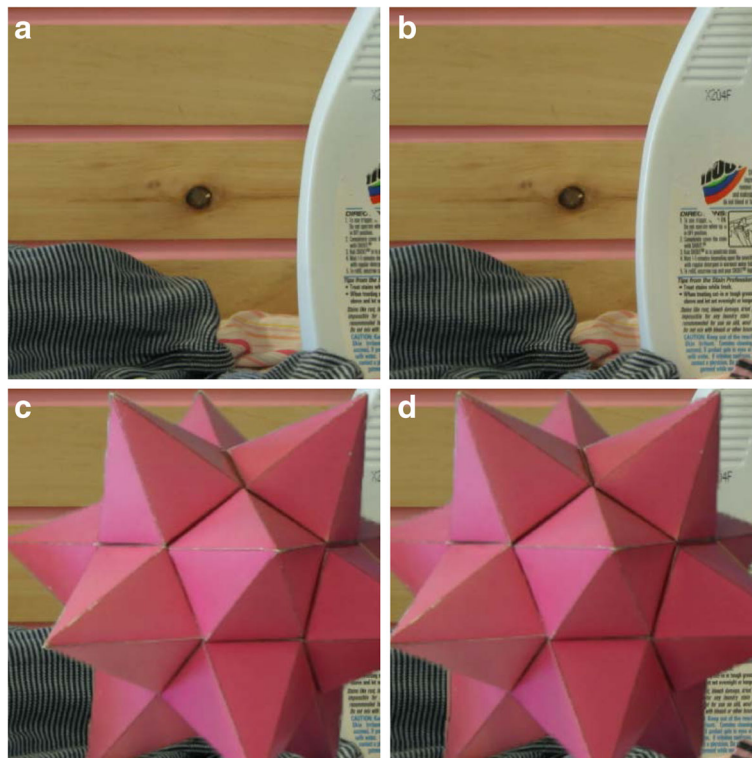


Fig. 10 Two pairs of testing image. **a** Original left image. **b** Original right image. **c** Left image including resampled part. **d** Right image including resampled part

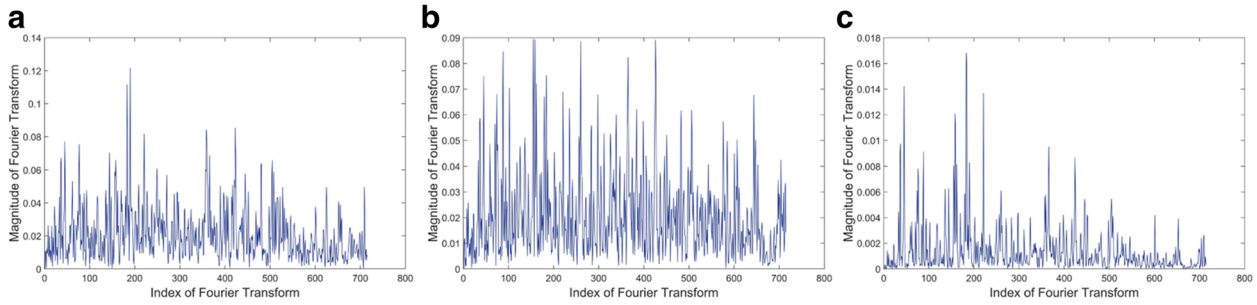


Fig. 11 The FFT result of the covariance signal without resampling process . **a** Left image. **b** Right image. **c** CSS processed image

After Fourier transforming, the effect of the CSS step can be revealed clearly. We tested that effect using image pairs represented in Fig. 10.

Figure 11a, b represents the FFT result of the covariance signal of the non-resampled Fig. 11a, b without the CSS process. Also, Fig. 11c represents the FFT result of the covariance signal of non-resampled image with the CSS process.

Figure 12a, b represents the FFT result of the covariance signal of the Fig. 12a, b images without the CSS process. The Fig. 12a, b includes a resampling region in almost part. Figure 12c represents the FFT result of the covariance signal of a resampled image with the CSS process. The result shows a clear peak when the CSS process is applied. The CSS process makes peak clearer only with resampled image. It makes accurate classification of a non-resampled image and a resampled image possible.

6. **Step 6 Peak to correlation energy calculation:** the result of the signal synthesized by the CSS process includes a single periodic pattern. To determine the resampling process on the synthesized signal, the PCE is calculated. The PCE is a measurement of the relative intensity of the peak against neighboring signals. The PCE is calculated as in Eq. 9.

$$\text{PCE} = \frac{|F_s(m_{\text{peak}})|^2}{\sum_{m=0}^{M-1} |F_s(m)|^2} \quad (9)$$

where m_{peak} is the index of the peak existing in the signal F_s .

7. **Step 7 The resampling factor determination:** as a final step, the resampling process can be determined as below by using a thresholding scheme.

$$\begin{cases} \text{The resampled images} & \text{PCE} > T_F \\ \text{The non-resampled images} & \text{PCE} \leq T_F \end{cases} \quad (10)$$

where T_F means the threshold that separates the resampled images from the non-resampled images. For the experiment, the threshold T_F was set by the corresponding false positive rate.

As can be seen from Eq. 8, the resampling factor can be estimated based on the normalized position of the peak. In this research, we analyzed the true positive rate according to the change of the false positive rate.

6 Experimental results and discussion

6.1 Experimental settings

The experiments were carried out on 1000 pairs (2000 images) of left and right stereoscopic 3D raw images. The partial image pairs of the experimental sets were obtained from Middlebury Stereo Datasets [21, 22]. The other image pairs were captured by two pairs of each 6D and 5D mark2 from Canon Corporation with a dual lens. The dual camera was set on a stereo tripod. Images of various sizes from 1024×768 to 3584×2016 were used in

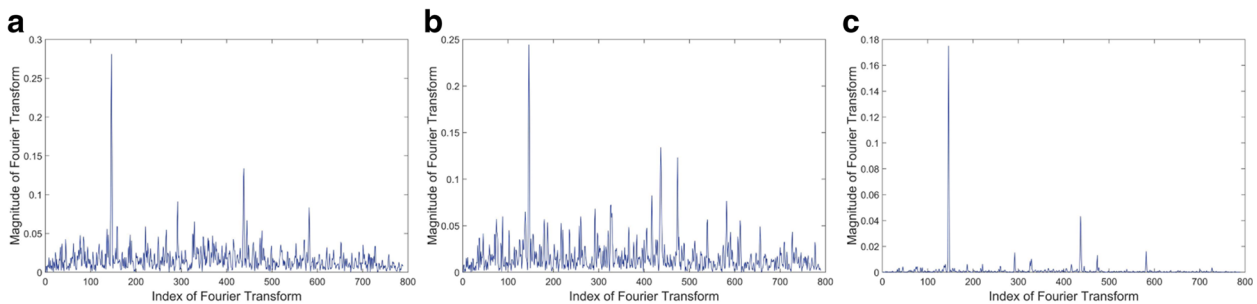


Fig. 12 The FFT result of the covariance signal with resampling process . **a** Left image. **b** Right image. **c** CSS processed image

the experiment. We used 500 image pairs for determining the threshold and 500 pairs for testing to avoid dataset dependency. We used Matlab2014b as a simulation tools. All of the methods in this paper are implemented in the Matlab2014b software.

In the experiments, a resizing process was conducted as the resampling process. For the resizing test, resampling factors from 1.01 to 1.30 were used for upsampling tests. Also, 0.70 to 0.99 factors were used for downsampling tests. The size of the blocks used for the experiment was 512×512 . Finally, both bilinear and bicubic interpolation kernels were selected for the resampling process to verify the independency on the interpolation kernel. In particular, anti-aliasing was applied during the resampling process for a realistic experimental scenario.

The JPEG quality factors used in the experiments were from 90 to 100. However, there is a noticeable fact related to the JPEG compression. The JPEG compression was basically conducted on 8×8 blocks and that causes the periodicity to be based on the block artifacts. Separating the periodicity from the target periodicity results from the resampling process is not an easy job because the effect of the periodicity seems to be a nearest-neighbor interpolation. The periodicities coming from the JPEG compression appear in the $n/8$ ($n = 1 \sim 7$) relative location of the signals. We set those range of signals into zeros to eliminate the periodicity from JPEG compression. Therefore, the peak cannot be detected when the peak from the resampling process is overlapped with the peak from the JPEG compression. In our experiments, there is no experimental results for scaling factor 0.80 for such a reason. With our best knowledge, no previous works overcame that problem.

6.2 The effectiveness of CRF process

In this section, we will show the effect of CRF. To show the effect of CRF process, it is better to analyze the peak after the Fourier transformation of the

signal calculated as Eq. 6. After the Fourier transformation, there is strong peak if an image has strong periodicity.

Without the CRF process, it is hard to detect the resampling process with very low factor such as 1.01 or 0.99. On the other hand, it is possible to detect weak periodicity by applying CRF process. Figure 13 describes the example of the CRF effect. Figure 13 shows the results of numerical periodicity of the image of a sample image resampled by the factor 1.01. With low resampling factor, the peak is hard to detect since the strength of signal from the periodicity is weak as seen in Fig. 13a. But when the CRF process is applied, a clear peak can be detected, as in Fig. 13b.

Figure 14 is the graph describing the magnitude of the peak along with the threshold values. The magnitude of the peak could be represented as PCE value. High PCE value means a signal has high periodicity and vice versa. Figure 14a is the graph that represents average PCE value of 1000 images with non-resampled images while Fig. 14b is the graph with resampled images. The threshold values have to be determined considering non-resampled and resampling images together. In this paper, we used the value of 53% of the differentiation map value as the threshold value.

6.3 Performance evaluation

Figure 15 represents the histograms of PCE values of all images. Figure 15a shows the histogram of PCE values of resampled and non-resampled images of Mahdian and Saic method [4]. In Mahdian and Saic method, the PCE histogram from non-resampled images are overlapped with the histogram from resampled images. A false positive errors of the resampling detector are occurred by these overlapped region of histogram. On the other hand, Fig. 15b shows the histogram of PCE values of proposed method. In the proposed technique, the PCE values from resampled images are distributed in a wider and higher range of energy bins. It is easy to recognize that the PCE

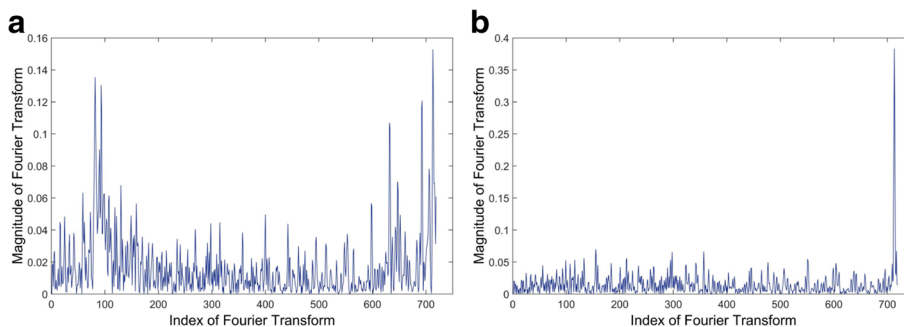


Fig. 13 A periodicity graph of image resampled by 1.01 factor. **a** The periodicity graph without CRF process. **b** The periodicity graph with CRF process

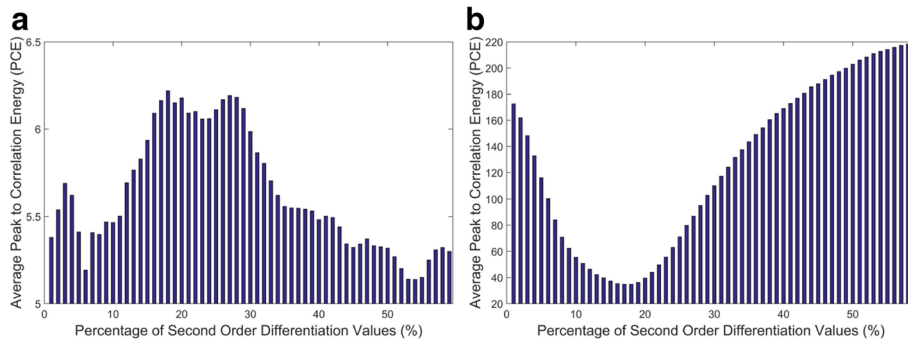


Fig. 14 The PCE values according to various threshold values. **a** Non-resampled images. **b** Resampled images

histograms from non-resampled images and resampled images are relatively separated from Fig. 15b.

We further benchmarked our technique against the three resampling detectors proposed by Mahdian, Kirchner [7], and Choi [14]. Two images were used for the input images because the proposed method targets stereoscopic 3D images, although the Mahdian and Kirchner works used only a single monoscopic image. To achieve fairness in the comparative experiment between the previous works and the proposed method, two PCE results from the left image and the right image were averaged. The decision thresholds were set based on 10^{-2} false positive rate for Mahdian and Kirchner works. Choi's method is a resampling detection technique for DIBR images, nevertheless, we applied it to stereoscopic images to compare with the proposed method. Since Choi's method cannot control the false positive rate through threshold, only Choi's method has a false positive rate of 29.2%.

Table 2 shows the detection results of upsampling with false positive rate 10^{-2} except the Choi's method. With 100 and 90 JPEG quality factors, the proposed method works at almost 100% correction results even with low resampling factors. Otherwise, the comparison works showed lower detection results compared to the proposed method. Table 3 shows the detection results of down-sampling with false positive rate 10^{-2} for Mahdian and Kirchner methods. In this result, wider difference of performance could be observed than upsampling case since CRF process which is introduced in Sec. 5 helps better performance even if a resampled image only has weak periodicity.

To evaluate the detection performance with various thresholds, a receiver operating characteristics (ROC) curve was used. Figure 16 shows the ROC curves with various resampling factors. Figure 16a–d shows the ROC curves with non-compressed images. The ROC curves

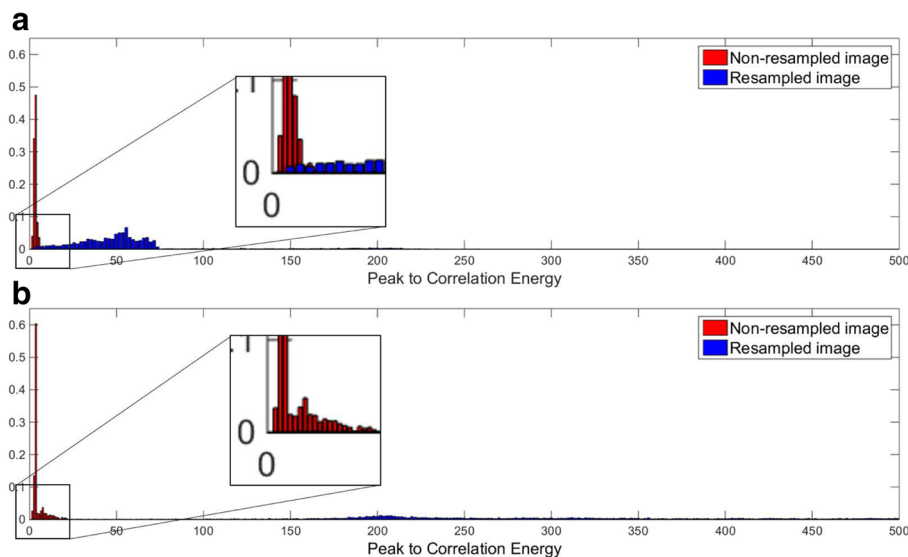


Fig. 15 The histogram of PCE values from non-resampled and resampled images. **a** The Mahdian's method [4]. **b** The proposed method

Table 2 The comparison of the upsampling detection rate results at false positive rate 10^{-2} except Choi's method (the false positive rate of Choi's method 29.2%)

		Scaling factors								
		1.01	1.025	1.05	1.075	1.10	1.15	1.20	1.25	1.30
Proposed	No	90.8	99.4	99.8	99.8	100	100	100	99.8	99.8
	JPEG100	90.0	98.0	99.8	99.8	100	100	100	100	100
	JPEG90	72.2	77.0	84.4	81.2	81.6	96.2	92.6	92.2	87.2
Mahdian [4]	No	84.8	91.2	92.4	95.6	98.0	99.8	100	100	100
	JPEG100	84.4	91.2	92.4	95.2	98.0	99.8	100	100	100
	JPEG90	76.8	85.2	86.2	88.0	88.8	94.8	93.0	91.8	92.8
Kirchner [7]	No	29.0	90.8	91.4	97.0	98.8	89.2	100	99.8	100
	JPEG100	27.4	93.8	94.8	98.2	99.4	95.0	100	100	100
	JPEG90	9.2	75.4	73.4	75.4	69.2	79.6	69.0	75.4	46.4
Choi [14]	No	77.2	80.4	87.9	90.5	90.2	92.0	95.8	96.4	97.5
	JPEG100	76.6	79.8	87.8	90.6	90.2	92.2	95.9	96.8	97.6
	JPEG90	60.8	71.3	82.9	83.8	85.2	86.3	88.4	88.2	86.3

show the detection rate along with the various false positive rates when the bicubic kernel was used as the interpolation kernel. The graph shows that the resampling detection performance is outstanding for the overall resampling factors. Figure 16e–h shows the detection performance on compressed images. With compressed images, the proposed technique worked well for low false positive rate.

Figure 17 shows the ROC curves when the bilinear kernel is used. The proposed method shows better

performance than the others, just as the bicubic kernel case. The graph shows the same tendency even when the JPEG compression has been applied. Although Mahdian's method showed better performance than Kirchner's method, the proposed scheme outperformed Mahdian's work overall.

Moreover, we used the area under the ROC curve (AUC) to summarize the performance with a single resampling factor. An AUC value of 0 means all of the detection results are false, whereas an AUC value of 1 means a

Table 3 The comparison of the downsampling detection rate results at false positive rate 10^{-2} except Choi's method (the false positive rate of Choi's method 29.2%)

		Scaling factors								
		0.70	0.75	0.80	0.85	0.90	0.925	0.95	0.975	0.99
Proposed	No	73.0	85.8	0	98.8	98.2	99.8	99.6	98.6	90.4
	JPEG100	72.0	84.6	0	98.8	98.2	99.6	99.2	98.2	91.4
	JPEG90	25.2	26.4	0	55.6	58.0	66.4	64.8	60.8	72.4
Mahdian [4]	No	72.6	81.2	0	91.0	91.2	93.0	91.6	90.8	83.4
	JPEG100	72.0	81.0	0	91.0	91.4	93.2	91.2	90.6	82.6
	JPEG90	41.0	45.6	0	70.6	72.6	74.8	76.2	77.2	78.4
Kirchner [7]	No	19.4	40.8	0	66.8	91.6	86.4	77.8	91.0	2.8
	JPEG100	21.2	38.6	0	72.6	94.4	90.4	84.0	80.0	1.6
	JPEG90	2.6	1.6	0	13.2	40.0	36.0	46.0	39.8	2.0
Choi [14]	No	54.7	73.9	0	80.9	82.4	84.3	79.7	78.7	74.8
	JPEG100	54.8	73.7	0	80.9	82.6	84.4	79.7	78.8	75.1
	JPEG90	35.2	42.4	0	63.4	73.2	76.1	74.8	68.5	60.9

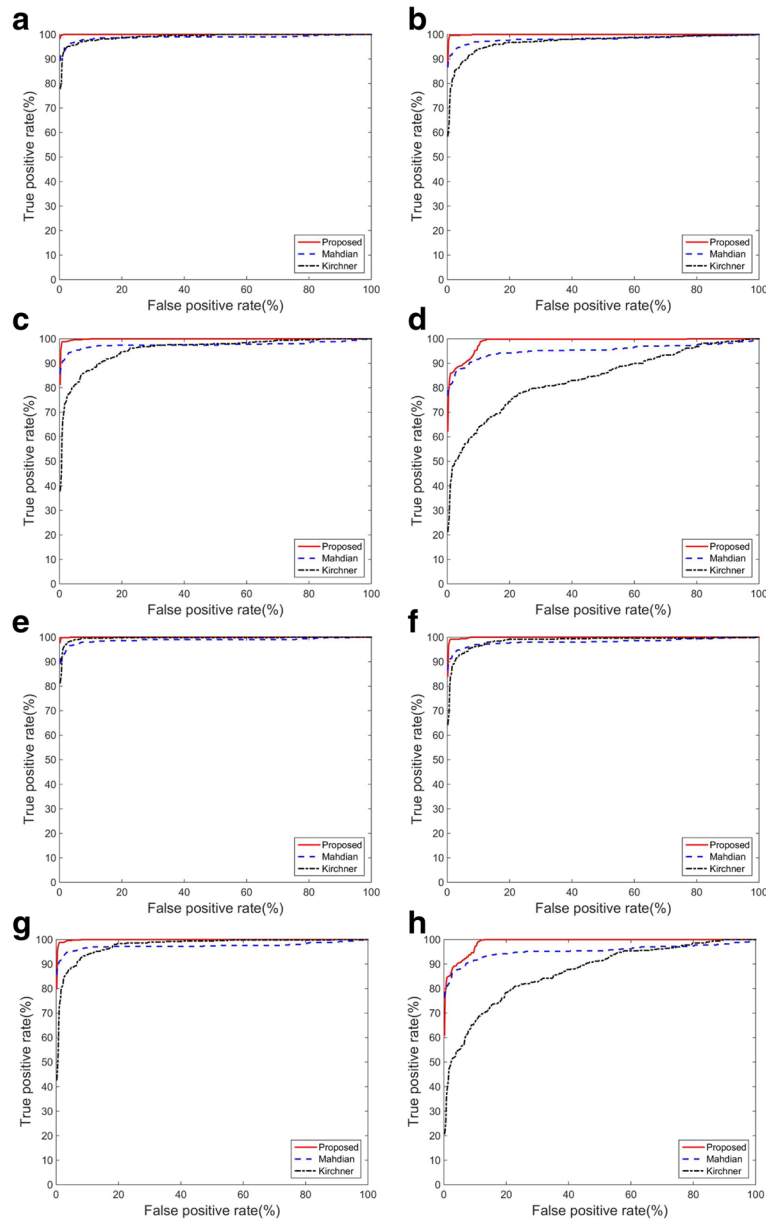


Fig. 16 ROC curves of each methods with bicubic kernel for various resampling factors. **a–d** No compression. **e–h** JPEG quality factor 100

perfect detection. Likewise, an AUC value of 0.5 means random supposition. Figure 18 shows the AUC when the bicubic kernel. The graph shows the proposed method has a higher AUC value than the comparison works for almost all resampling factors, specially low resampling factors.

7 Conclusions

Stereoscopic 3D content became popular as a successive series of movies. Nevertheless, forensic research to develop protections for stereoscopic 3D content has rarely been carried out. Therefore, development of a resampling

detector for stereoscopic 3D images is essential. The forensic research for stereoscopic 3D images can be naturally extended to the multi-view images.

Since previous works considered only monoscopic images, those schemes cannot use the unique characteristics of the stereoscopic 3D images. Prior methods can only generate two independent results from stereoscopic 3D images, and consequently advanced detection performance cannot be expected with those results. Compared to the previous works, better results can be expected by properly using the

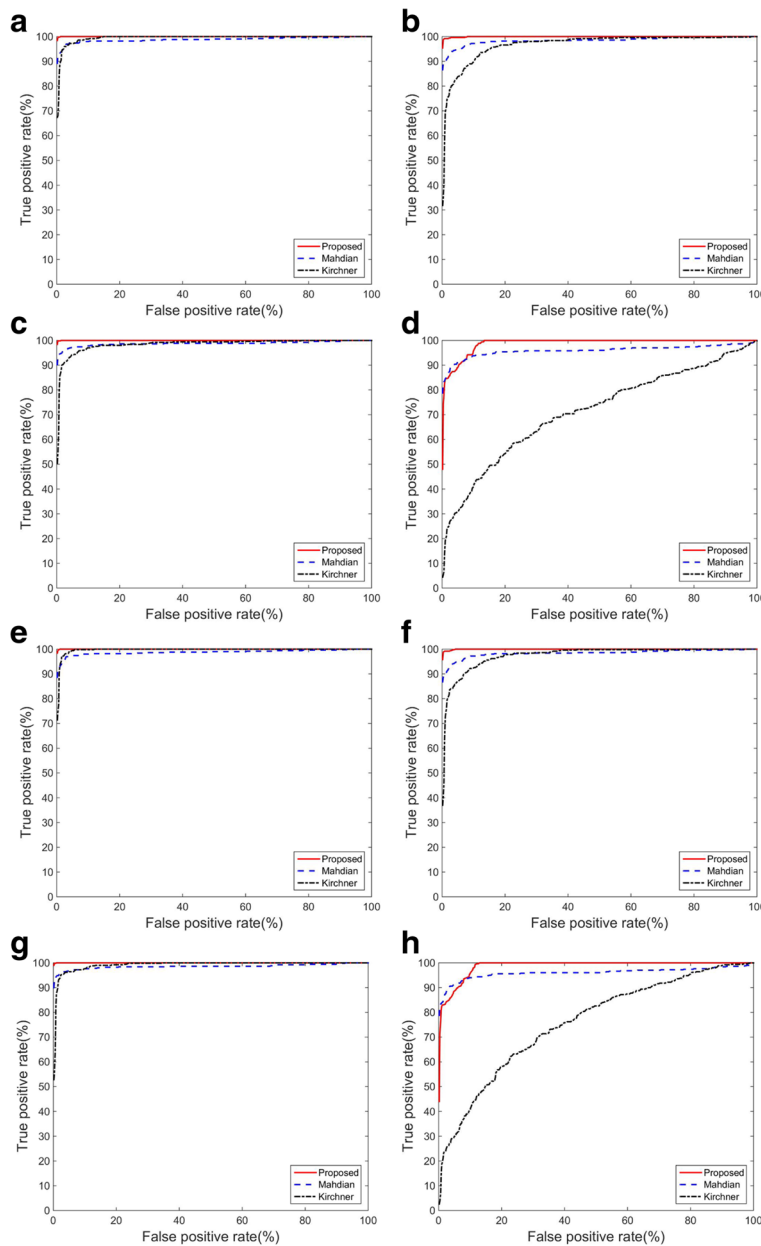


Fig. 17 ROC curves of each methods with bilinear kernel for various resampling factors. **a–d** No compression. **e–h** JPEG quality factor 100

relation of the left and the right images as a side information.

In this paper, we proposed a blind and efficient method for detecting the resampling of stereoscopic 3D images. We applied the CRF process to obtain independence from the characteristics of the image itself.

We conducted a subjective evaluation to analyze the visual fatigue experienced by human during the operation of 3D stereoscopic images, and consequently found constraints to be followed during the forgery process.

Moreover, we developed the CSS process, based on the constraints that occur during the forgery of stereoscopic 3D images. We obtained clearer periodicity of stereoscopic images, since the CSS process amplified the periodic property.

The experimental results in this study demonstrated that the proposed scheme has superior detection performance compared with those of the previous methods. The results showed a higher detection rate for JPEG-compressed images.

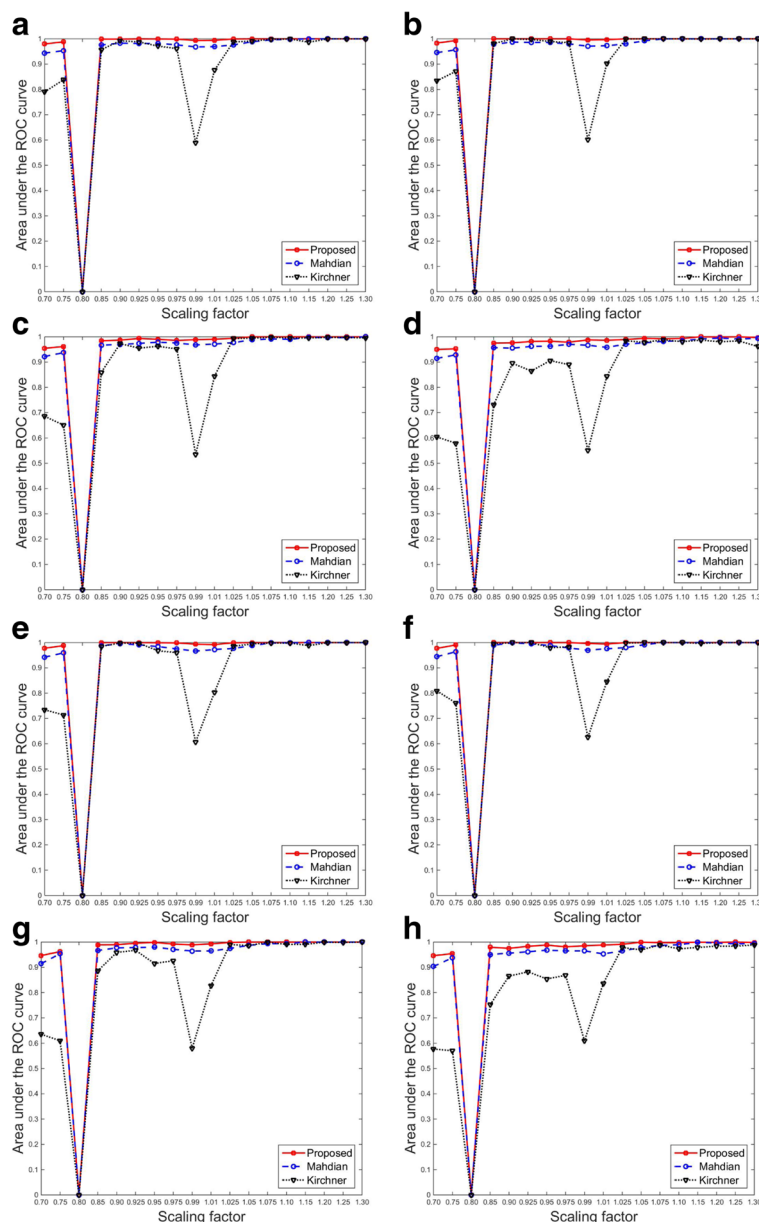


Fig. 18 AUC graphs of each method for various JPEG quality factors, bicubic kernel for various JPEG quality factors. **a–d** Bicubic kernel. **e–h** Bilinear kernel

In future research, the target could be extended to video, which are the other multi-track-based content formats. This research was conducted as a middle-stage solution for tamper protection technology of next generation content such as multi-view images. The research for multi-view images will be an important work.

Acknowledgements

This research project was supported by Ministry of Culture, Sports and Tourism(MCST), and from Korea Copyright Commission in 2016.

Funding

None.

Authors' contributions

H-YC proposed the main idea and developed the details. He carried out almost all the experiments. D-KH decisively inspired to make the mail idea. He constitutes the parts of the discussion. Also, he conducted experiments of the comparison methods. SC has influenced the whole idea, the structure of the paper, and the revision of the paper. H-KL organized the structure of manuscript. Also, he advised about conducting experiments and developing main idea. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 29 August 2016 Accepted: 15 February 2017

Published online: 03 March 2017

References

1. A Popescu, H Farid, in *IEEE Trans. on Signal Process.* Exposing digital forgeries by detecting traces of re-sampling, vol. 53, (2005), pp. 758–767
2. A Gallagher, in *Proc. of the Second Canadian Conf. on Computer and Robot Vision*. Detection of linear and cubic interpolation in JPEG compressed images, (2005), pp. 65–72
3. B Mahdian, S Saic, in *Third International Symposium on Information Assurance and Security*. On periodic properties of interpolation and their application to image authentication, (2007), pp. 439–446
4. B Mahdian, S Saic, in *IEEE Trans. on Information Forensics and Security*. Blind authentication using periodic properties of interpolation, vol. 3, (2008), pp. 529–538
5. M Kirchner, in *Proc. of the 10th ACM workshop on Multimedia and security (MM&Sec2008)*. Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue, vol. 47, (2008), pp. 11–20
6. M Kirchner, T Gloe, in *IEEE Workshop on Information Forensics and Security (WIFS2009)*. On resampling detection in re-compressed images, (2009), pp. 21–25
7. M Kirchner, in *Proc. of the ACM workshop on Multimedia and Security (MM&Sec2010)*. Linear row and column predictors for the analysis of resized images, (2010), pp. 13–18
8. X Feng, I Cox, G Doerr, Normalized energy density-based forensic detection of resampled images. *IEEE Trans. Multimed.* **14**(3), 536–545 (2012)
9. D Vazquez-Padin, P Comesana, F Perez-Gonzalez, in *IEEE Workshop on Information Forensics and Security (WIFS2013)*. Set-membership identification of resampled signals, (2013), pp. 150–155
10. D Vazquez-Padin, P Comesana, F Perez-Gonzalez, in *2015 23rd European Signal Processing Conference (EUSIPCO2015)*. An SVD approach to forensic image resampling detection, (2015), pp. 2067–2071
11. H Nguyen, S Katzenbeisser, Robust resampling detection in digital images. *IFIP Int. Fed. Inf. Process.* **7394 LNCS**, 3–15 (2012)
12. D Vazquez-Padin, C Mosquera, F Perez-Gonzalez, in *IEEE 17th Int. Conf. on Image Processing (ICIP2010)*. Two-dimensional statistical test for the presence of almost cyclostationarity images, (2010)
13. D Vazquez-Padin, F Perez-Gonzalez, in *IEEE Workshop on Information Forensics and Security (WIFS2011)*. Prefilter design for forensic resampling estimation, (2011)
14. H Choi, D Hyun, H Lee, in *The 4th Int. Conf. on 3D Systems and Applications (3DSA2012)*. Enhanced resampling detection for DIBR stereoscopic image, (2012)
15. M Fouche, M Olivier, Using internal depth to aid stereoscopic image splicing detection. *Adv. Digit. Forensic.* **Viii. 383**, 319–333 (2012)
16. M Kirchner, P Winkler, H Farid, in *Proc. SPIE Conf. Media Watermarking, Security, and Forensics 2013*. Impeding forgers at photo inception, (2013)
17. P Comesana, F Perez-Gonzalez, in *2013 IEEE 15th International Workshop on Multimedia Signal Processing (MMSP2013)*. Taking advantage of source correlation in forensic analysis, (2013)
18. J Parker, R Kenyon, D Troxel, Comparison of interpolating methods for image resampling. *IEEE Trans. Med. Imaging.* **2**(1), 31–39 (1983)
19. M Lambooij, W Ijsselstein, in *Journal of Image Science and Technology*. Visual discomfort and visual fatigue of stereoscopic displays: a review, (2009)
20. 9241, I.: Ergonomics of human-system interaction - Part 392: ergonomic recommendations for the reduction of visual fatigue from stereoscopic images. International Standard Organization
21. D Scharstein, R Szeliski, in *CVPR2003*. High-accuracy stereo depth maps using structured light, vol. 1 (IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2003), pp. 195–208
22. D Scharstein, C Pal, in *CVPR2007*. Learning conditional random fields for stereo (IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2007)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com