# INPAINTING IN OMNIDIRECTIONAL IMAGES FOR PRIVACY PROTECTION

*Evgeniy Upenik, Pinar Akyazi, Mehmet Tuzmen and Touradj Ebrahimi*

Multimedia Signal Processing Group (MMSPG)
Ecole Polytechnique Fédérale de Lausanne (EPFL)
CH-1015 Lausanne, Switzerland
Email: firstname.lastname@epfl.ch

## ABSTRACT

Privacy protection is drawing more attention with the advances in image processing, visual and social media. Photo sharing is a popular activity, which also brings the concern of regulating permissions associated with shared content. This paper presents a method for protecting user privacy in omnidirectional media, by removing parts of the content selected by the user, in a reversible manner. Object removal is carried out using three different state-of-the-art inpainting methods, employed over the mask drawn in the viewport domain so that the geometric distortions are minimized. The perceived quality of the scene is assessed via subjective tests, comparing the proposed method against inpainting employed directly on the equirectangular image. Results on distinct contents indicate our object removal methodology on the viewport enhances perceived quality, thereby improves privacy protection as the user is able to hide objects with less distortion in the overall image.

***Index Terms***— privacy protection, object removal, image inpainting, omnidirectional imaging, virtual reality

## 1. INTRODUCTION

With the advent of smart mobile devices and social networks, photo sharing has become an easy and widespread activity among users. The increasing distribution of images also raises issues on privacy protection and creates the need for adjusting permissions, as the shared content contains sensitive information concerning users. Access control over contents provides exclusive rights to only selected correspondents, thereby enhancing user security and privacy. A widely preferred form of privacy protection is to obfuscate parts of images, instead of encrypting or permuting the whole image [1]. This results in less visual distortions that are confined only to the specific area of interest. The obfuscation can be achieved using a

variety of image processing techniques, such as blurring, mosaicking, masking and object removal. Among these methods, the first three have to introduce a high amount of distortion to hide the underlying content, whereas object removal provides more natural viewing conditions while still being able to protect the content. This process also can be made reversible similar to work [1], therefore access to the original data can be granted to selected users with permissions.

A recently popular multimedia modality is omnidirectional imaging, where the user has an immersive experience. In this paper, we present an object removal methodology via inpainting on omnidirectional images performed on the selected viewport instead of the equirectangular representation, which yields visually plausible results. Given an omnidirectional image, we extract the viewport and apply the mask defining the objects to be removed on the viewport. We remove objects using three distinct state-of-the-art inpainting algorithms [2–4]. Inpainting on the viewport rather than the equirectangular image minimizes the geometric distortions and limits the source region to more relevant components within the content. After removing an object and inpainting the background, we project the viewport back onto the equirectangular image. We, furthermore, assess quality of the protected content by performing subjective evaluations, where we compare the images inpainted using our method and in the equirectangular domain directly, using Absolute Category Rating (ACR) [5].

Image inpainting algorithms can be divided into four general classes: statistical methods, partial differential equation (PDE)-based methods, exemplar-based methods and deep generative models based on convolutional neural networks [4, 6]. Statistical methods make use of parametric models to describe input textures, however they fail in the presence of additional intensity gradients [7]. PDE-based methods propagate information from the known part of the image [8–10] using smoothness priors, which introduces blurring when large and high frequency regions needs to be inpainted. Exemplar-based methods and deep generative models are most widely used, where the former fills the holes in the image using exemplars from local or global search
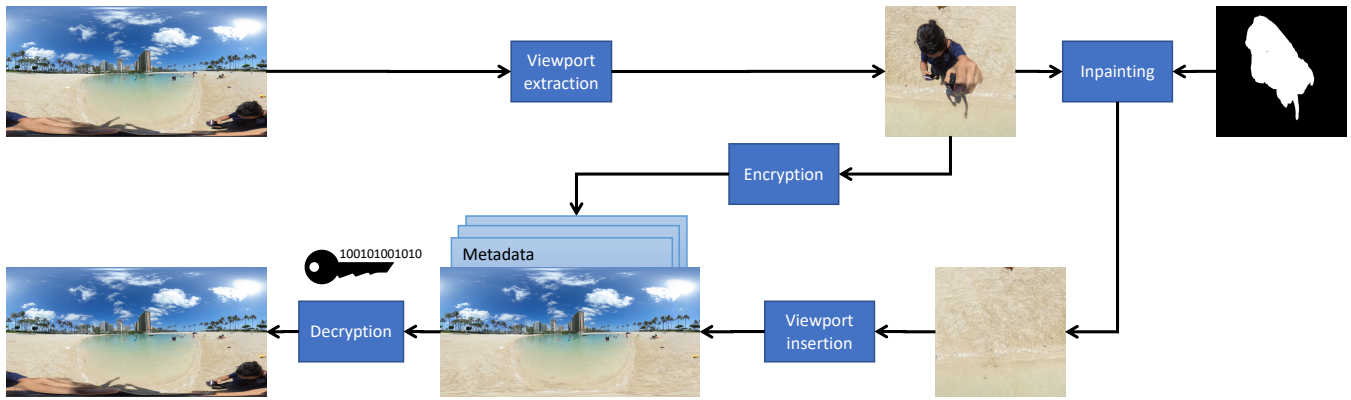
**Fig. 1**. Applying the viewport extraction method for reversible secure object removal using inpainting

regions [2, 3, 11–13] and the latter exploits semantics learned from large scale datasets [14–16]. We have selected two robust methods for exemplar-based inpainting [2, 3] as well as one semantic learning-based state-of-the-art method [17] in order to reduce the bias of the preferred inpainting technique on our object removal strategy on the viewport.

While most works on object removal and inpainting focus on planar images, panoramic content is considered in [18]. A field-of-view expansion method using retargeting techniques combined with Graphcut Textures is proposed to remove objects near the equator, and extended to farther portions of the sphere by tripod rotations. Although objects can be removed regardless of their locations, rotation of the full equirectangular image is more costly than viewport extraction. Our approach minimizes the geometric distortions within a limited search region, thereby reducing the complexity of inpainting simultaneously.

## 2. VIEWPORT EXTRACTION METHOD

In this section we present a method to perform object removal in omnidirectional images using inpainting.

### 2.1. Omnidirectional images

Omnidirectional content is a type of immersive multimedia which allows interactions such as change of the direction of sight from a fixed point of view. Omnidirectional images typically cover full spherical field of view and can be represented in different formats or projections to a plane, such as equirectangular, cubic, etc. in order to be stored or processed. We elaborate our method considering equirectangular projection, because of its common usage. However, the method can be similarly applied to omnidirectional content represented in other formats.
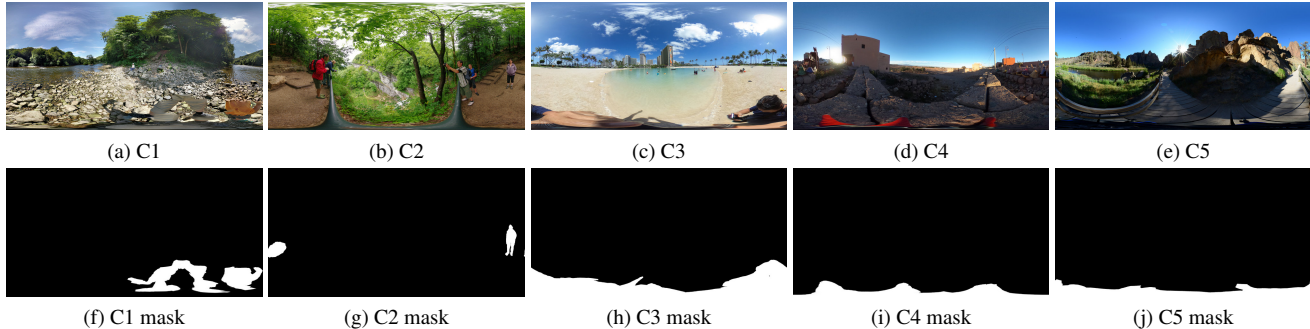
Normal way of consuming omnidirectional content is through a rendering application which displays a current view on a screen of a hand-held device, on a head-mounted display (HMD), or on a regular screen, whilst the equirectangular representation remains undiscovered by viewers. This back-end representation though requires image processing algorithms to comply with it. Equirectangular projection introduces strong geometrical distortions in the near-pole areas, where all objects are stretched horizontally. This fact brings particular difficulties for the state-of-the-art inpainting algorithms.

### 2.2. Viewport extraction

A viewport is a part of an omnidirectional image which is observed by a user at one moment. During the process of rendering a viewport is extracted from an equirectangular representation and shown to a user. Unlike in the back-end equirectangular representation, the geometrical distortion in the viewport image is negligible.

A block-diagram in Figure 1 describes the method of viewport extraction for object removal in omnidirectional images. Here we apply inpainting algorithms in the viewport domain such that it is performed on common planar images. Afterwards, the viewport with inpainted area is inserted back to the equirectangular image. In order to make the process of object removal reversible, one can keep the original viewport and store it as metadata in the image file, similarly to how it is done in [19]. The original viewport which contains information critical for privacy protection can be also encrypted. Such transmorphed [1] omnidirectional image can be transmitted and shared securely allowing extraction of hidden data by authorized parties who possess a private decryption key.

(a) C1      (b) C2      (c) C3      (d) C4      (e) C5

(f) C1 mask      (g) C2 mask      (h) C3 mask      (i) C4 mask      (j) C5 mask

**Fig. 2**. Contents selected for experiments (top row) and masks for object removal (bottom). The objects to be removed are depicted with white color on masks.
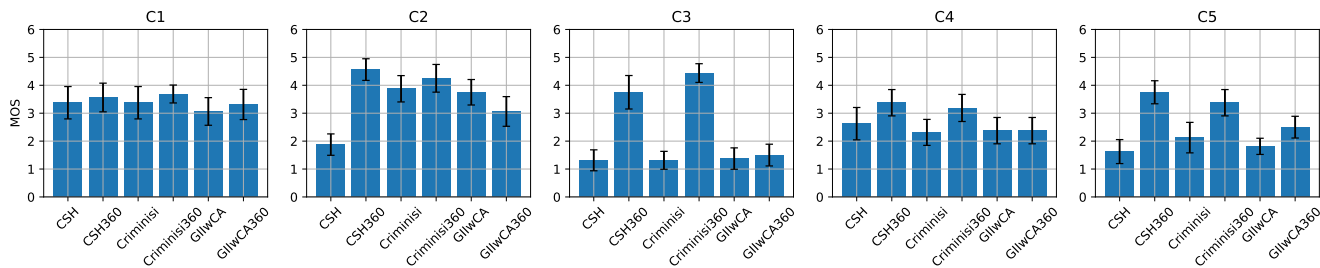


**Fig. 3**. Mean opinion scores with 95% confidence intervals

## 3. EXPERIMENT

This section contains a step-by-step description of removing objects on images in order to further assess the performance of the proposed method.

### 3.1. Dataset

Five distinct omnidirectional contents were selected amongst publicly available photographic works licensed with Creative Commons[1]. Masks for object removal were created manually, as depicted in Figure 2. Original contents and prepared masks can be retrieved from a GIT repository in [2]. Each content was downsampled to 2048x1024 resolution and the resolutions of masks are identical. All contents are natural images with outdoor views, where the location of objects to be removed varies from the equator to the south pole of the scene.

### 3.2. Inpainting algorithms

We use three state-of-the-art inpainting algorithms and their modifications which are deriving from applying the method described in Section 2.2. Table 1 contains the full list of the algorithms which were selected to be enhanced with viewport extraction.

### 3.3. Inpainting procedure

Table 2 contains the viewport parameters selected for each content from the dataset described in Section 3.1. All the viewports have the same size and field of view. The positions were selected in such a way that an object to be removed appears approximately in the center of the viewport and the whole inpainted area fits inside.

For each omnidirectional image represented in equirectangular projection the viewport is extracted according to the parameters specified in Table 2. Then inpainting is applied on the viewport in the areas depicted in Figure 2 using the algorithms CSH, Criminisi, and GIIwCA (Table 1). The inpainted viewport is then inserted back in the equirectangular image. Further in the paper these methods are called CSH360, Criminisi360, and GIIwCA360, respectively.

In order to have a reference to assess viewport extraction enhancement, we also apply inpainting directly in the equirectangular domain for all the images from the dataset using exactly the same masks.

## 4. SUBJECTIVE EVALUATION

This section describes the subjective evaluations of visual quality including methodology and processing of the collected data.

---

[1] https://creativecommons.org/
[2] https://github.com/mmspg/inpainting360

| Abbreviation | Base algorithm name | Viewport | Source code |
|---|---|---|---|
| CSH | Coherency sensitive hashing | No | http://github.com/PetterS/patch-inpainting [Commit: 03cc575] |
| CSH360 | Coherency sensitive hashing | Yes | - - |
| Criminisi | Exemplar-Based Image Inpainting | No | http://github.com/cheind/inpaint [Commit: 864128c] |
| Criminisi360 | Exemplar-Based Image Inpainting | Yes | - - |
| GIIwCA | Generative Image Inpainting w/ Contextual Attention | No | http://github.com/JiahuiYu/generative_inpainting [Commit: 6bfaa20] |
| GIIwCA360 | Generative Image Inpainting w/ Contextual Attention | Yes | - - |

**Table 1**. Inpainting algorithms used in the paper

| Content | Viewport position | FoV° | Size |
|---|---|---|---|
| C1 | yaw: 30°, pitch: -80° | 90×90 | 1024×1024 px |
| C2 | yaw: 180°, pitch: 0° | 90×90 | 1024×1024 px |
| C3 | yaw: 180°, pitch: -80° | 90×90 | 1024×1024 px |
| C4 | yaw: 0°, pitch: -90° | 90×90 | 1024×1024 px |
| C5 | yaw: 0°, pitch: -90° | 90×90 | 1024×1024 px |

**Table 2**. Viewport positions for inpainted contents

### 4.1. Test methodology

The Absolute Category Rating (ACR) method described in [5] was chosen to assess the effect of the proposed viewport extraction method for object removal on the quality of selected contents. ACR is a single stimulus evaluation where the test stimuli are randomly presented to subjects and voting is performed after each viewing. Overall quality is assessed using a five-grade scale with the following levels: 5 - Excellent, 4 - Good, 3 - Fair, 2 - Poor, and 1 - Bad.

A total of 16 consenting subjects participated in the study, of which 7 were female, with an overall median age of 26.5. All subjects have passed color vision and visual acuity tests prior to experiments, using Ishihara and Snellen charts, respectively. The subjects were placed in an immersive environment; and stimuli were presented to them using a head mounted display (HMD) composed of a VR head-mount with buttons[3] and a mobile device installed inside as a screen. Samsung Galaxy S7 edge SM-G935F with a screen resolution of 2560x1440 pixels was used to display the images. Subjects were sitting on a rotatable chair during the assessment. Immersive textual instructions were provided inside the VR along with a verbal guidance by the experimenter, as described in [20].

Each experiment started with an immersive training, where original contents were presented to the subjects with a red circle indicating the position of the objects to be removed later in the test session. This way, subjects were familiarized with the contents and surroundings of the objects to be removed. The circles were omitted during test session in order not to disturb the natural viewing of the stimuli. During evaluation, subjects assessed the stimuli shown to them consequently without any time restrictions. When ready to rate an image subjects had to activate a 3D immersive voting menu by pressing a button and select the grade

proceeding immediately to the next image. All stimuli were automatically randomized in each session. The experiments were conducted using a testbed for subjective evaluation of omnidirectional visual content proposed in [21]. This software was developed for Android and is publicly available for downloading[4].

### 4.2. Data processing

Outlier detection was performed separately on the raw scores using a method described in [22]. None of the subjects were identified as outliers during our experiments. The mean opinion score (MOS) and 95% confidence intervals (CIs) assuming a Students t-distribution of the scores were computed for each test condition [23].

## 5. RESULTS AND DISCUSSION

Figure 3 presents the results of the subjective evaluation described in Section 4. Bar plots with 95% confidence intervals show how different methods perform on different contents. The plots are grouped by content. As one can see, viewport extraction significantly enhances inpainting. CSH360 performs better than CSH on four contents, Criminisi360 is better on three contents, and GIIwCA360 is superior to GIIwCA on one content.

For the contents C3-C5 viewport extraction brings higher quality gain than for C1-C2. The fact that on C2 the quality does not improve in two cases out of three can be possibly explained by the position of inpainted area lying near equator where geometrical distortion is minimal hence its compensation is not needed.

## 6. CONCLUSION

In this paper, we presented a method for reversible object removal in omnidirectional images, which is targeted for privacy protection in immersive media. We show by performing subjective quality evaluation involving 16 naive subjects that viewport extraction can enhance the performance of state-of-the-art inpainting algorithms in omnidirectional images.

---

[3]https://mergevr.com

[4]https://github.com/mmspg/testbed360-android

# 7. REFERENCES

[1] L. Yuan and T. Ebrahimi, "Image transmorphing with JPEG," in *2015 IEEE International Conference on Image Processing (ICIP)*, Sept. 2015, pp. 3956–3960.

[2] Antonio Criminisi, Patrick Pérez, and Kentaro Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on image processing*, vol. 13, no. 9, pp. 1200–1212, 2004.

[3] Simon Korman and Shai Avidan, "Coherency sensitive hashing," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 6, pp. 1099–1112, 2016.

[4] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang, "Generative image inpainting with contextual attention," in *CVPR*, 2018.

[5] "ITU-T Rec. P.910 Subjective video quality assessment methods for multimedia applications," 2008.

[6] Pinar Akyazi and Pascal Frossard, "Graph-based inpainting of disocclusion holes for zooming in 3d scenes," in *Proceedings of EUSIPCO*, 2018.

[7] Anat Levin, Assaf Zomet, and Yair Weiss, "Learning how to inpaint from global image statistics," in *null*. IEEE, 2003, p. 305.

[8] Marcelo Bertalmio, Andrea L Bertozzi, and Guillermo Sapiro, "Navier-stokes, fluid dynamics, and image and video inpainting," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. IEEE, 2001, vol. 1, pp. I–I.

[9] David Tschumperle and Rachid Deriche, "Vector-valued image regularization with pdes: A common framework for different applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 506–517, 2005.

[10] Mahmoud Ghoniem, Youssef Chahir, and Abderrahim Elmoataz, "Geometric and texture inpainting based on discrete regularization on graphs," in *Image Processing (ICIP), 2009 16th IEEE International Conference on*. IEEE, 2009, pp. 1349–1352.

[11] Olivier Le Meur, Josselin Gautier, and Christine Guillemot, "Examplar-based inpainting based on local geometry," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*. IEEE, 2011, pp. 3401–3404.

[12] Kaiming He and Jian Sun, "Statistics of patch offsets for image completion," in *Computer Vision–ECCV 2012*, pp. 16–29. Springer, 2012.

[13] Pierre Buyssens, Maxime Daisy, David Tschumperlé, and Olivier Lézoray, "Exemplar-based inpainting: Technical review and new heuristics for better geometric reconstructions," *IEEE transactions on image processing*, vol. 24, no. 6, pp. 1809–1824, 2015.

[14] Junyuan Xie, Linli Xu, and Enhong Chen, "Image denoising and inpainting with deep neural networks," in *Advances in neural information processing systems*, 2012, pp. 341–349.

[15] Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa, "Globally and locally consistent image completion," *ACM Transactions on Graphics (TOG)*, vol. 36, no. 4, pp. 107, 2017.

[16] Raymond A Yeh, Chen Chen, Teck-Yian Lim, Alexander G Schwing, Mark Hasegawa-Johnson, and Minh N Do, "Semantic image inpainting with deep generative models.," in *CVPR*, 2017.

[17] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang, "Free-form image inpainting with gated convolution," *arXiv preprint arXiv:1806.03589*, 2018.

[18] Andrew MacQuarrie and Anthony Steed, "Object removal in panoramic media," in *Proceedings of the 12th European Conference on Visual Media Production*. ACM, 2015, p. 2.

[19] Martin Rerabek, Evgeniy Upenik, and Touradj Ebrahimi, "JPEG backward compatible coding of omnidirectional images," in *Applications of Digital Image Processing XXXIX*, 2016, vol. 9971.

[20] Evgeniy Upenik, Martin Rerabek, and Touradj Ebrahimi, "On the performance of objective metrics for omnidirectional visual content," in *9th International Conference on Quality of Multimedia Experience (QoMEX 2017)*, 2017.

[21] Evgeniy Upenik, Martin Rerabek, and Touradj Ebrahimi, "Testbed for subjective evaluation of omnidirectional visual content," in *2016 Picture Coding Symposium (PCS)*, 2016.

[22] "ITU-R BT.500-13: Methodology for the subjective assessment of the quality of television pictures," 2012.

[23] Francesca De Simone, Lutz Goldmann, Jong-Seok Lee, and Touradj Ebrahimi, "Towards high efficiency video coding: Subjective evaluation of potential coding technologies," *Journal of Visual Communication and Image Representation*, vol. 22, no. 8, pp. 734–748, 2011.