JPEG backward compatible coding of omnidirectional images

Martin Řeřábek, Evgeniy Upenik, and Touradj Ebrahimi

Multimedia Signal Processing Group (MMSPG), EPFL, Lausanne, Switzerland

ABSTRACT

Omnidirectional image and video, also known as 360 image and 360 video, are gaining in popularity with the recent growth in availability of cameras and displays that can cope with such type of content. As omnidirectional visual content represents a larger set of information about the scene, it typically requires a much larger volume of information. Efficient compression of such content is therefore important. In this paper, we review the state of the art in compression of omnidirectional visual content, and propose a novel approach to encode omnidirectional images in such a way that they are still viewable on legacy JPEG decoders.

Keywords: omnidirectional image compression, region of interest estimation, virtual reality

1. INTRODUCTION

Omnidirectional image is a specific type of visual content that has become popular thanks to its rich rendering possibilities and increased sense of immersion. Omnidirectional visual information usually represents a field of view covering an entire sphere, a hemi-sphere, or a spherical layer. Alternative approaches exist such as representing omnidirectional visual content on a full or partial cylinder. For the remainder of this paper, we will limit our discussions on spherical capture, although the proposed solution can cope with other alternative omnidirectional representations too. For visualisation of omnidirectional visual content, interactive rendering is popular and offers users the possibility to explore and navigate through the scene and freely change the viewing direction.

In recent years, emerging interest in omnidirectional imaging has shifted focus from computer and robot vision to multimedia and entertainment applications. Such a shift challenges traditional image processing techniques, and brings new requirements on omnidirectional visual content processing such as on compression, on storage, on delivery and on quality assessment. One of the most important of such requirements is on efficient omnidirectional image storage and delivery preserving a high perceived quality and a rich user experience.

Representation of omnidirectional visual content is nowadays performed by converting it into a conventional 2D image with an appropriate projection, e.g. equirectangular, Lambert cylindrical equal-area, dyadic, cubic, etc. The first step, when representing the omnidirectional content, is to define an image formation model, such as a pinhole camera model. A formation model determines the mutual sphere-to-plane and plane-to-sphere projections. Equirectangular is one of the most widely used projections. Projection used to represent an omnidirectional image influences efficiency of the compression algorithm that follows and affects the visual quality of the rendering.

As mentioned above, omnidirectional visual content is often used in an interactive way to fully experience it through advanced rendering capabilities. In general, at one time instance, a user is allowed to visualize only a limited field of view, which can be changed based on interaction. Nevertheless, not all of the available devices, services, and applications allow such an interaction. When advanced rendering capabilities are not supported by the end-user application, a viewport representing properly but also sufficiently the omnidirectional content, should be delivered. In the rest of the paper we refer to such a desired viewport as a Region Of Interest (RoI) on an omnidirectional content. This assumption challenges the current image formats to cope simultaneously with both approaches and to be able to render an omnidirectional image in an interactive way to take full advantage of its rich content, and, at the same time, facilitate rendering of a suitable viewport image when such interactivity is not possible.

Further author information: (Send correspondence to authors) E-mail: {firstname.lastname}@epfl.ch

In this paper, we propose such a format called OmniJPEG, as an extension of the most popular image format - JPEG. The architecture of the corresponding proposed codec is designed with an emphasis on the backward compatibility with the legacy JPEG decoders. Additionally, we assess the performance of the proposed approach. Since effective use of OmniJPEG relies on a precise RoI estimation, the paper also introduces an automatic way to estimate an RoI in omnidirectional images. The proposed RoI estimation of omnidirectional images is based on the well know Graph-Based Visual Saliency (GBVS) model[1] which has proven to be quite efficient for conventional 2D images. Furthermore, an experiment to collect the ground truth RoI of a collection of typical omnidirectional images is conducted. RoI estimation results show that the proposed estimation model can efficiently predict RoI of omnidirectional images, which can be further used for the proposed OmniJPEG coding algorithm.

The reminder of the paper is organized as follows. Section 2 presents related work in omnidirectional compression and omnidirectional RoI prediction. In Section 3, the proposed JPEG backward compatible coding algorithm is described in details. Section 4 introduces an approach to automatically extract an RoI from an omnidirectional image. Conducted experiments to validate the proposed solution are also discussed here. Finally, Section 5 concludes the paper.

2. RELATED WORK

Several approaches to code omnidirectional visual content have been proposed in the past few years. Main approaches to compress omnidirectional images and video presented in literature can be listed as follows: a) adaptive and partial content delivery methods, b) algorithms exploiting or adapting to 2D spherical surface geometry of the content to be compressed, and c) geometric representation or projection based methods as pre- and postprocessing prior to compression.

One of the very first attempts to introduce a solution to code an omnidirectional or a panoramic image data was developed by Apple Inc. The so-called Quicktime VR,[2] which refers to both a file format and visualization software, allows for creation and display of panoramic images. More specifically, it proposes to store 360 degree cylindrical panoramic images divided in tiles. While displaying panoramic images with a specially designed viewer, only the tiles visible in the current viewport are decoded. However, the proposed panoramic image format did not have backward compatibly with conventional image viewers, such as legacy JPEG decoders, and it did not consider any RoI estimation to provide users with a default viewport.

The idea of using RoI to code omnidirectional visual content recently appeared in work considering adaptive coding and partial delivery methods.[3] This approach proposes to deliver only a part of the omnidirectional content which is being viewed. Each frame, after an equirectangular projection, is divided into tiles or regions which are then coded separately with different quality according to an adaptive model. The tiles covered by the portion of the frame which is being viewed are encoded with the highest possible quality. The quality of other regions is determined considering their probability of being viewed next. However, authors did not investigate any statistical model to predict next candidate tiles which will be viewed. Additionally, it did not consider that an omnidirectional image can be divided into tiles corresponding to a predicted RoI.

Coding by taking into account a specific 2D spherical surface has also been investigated in the past.[4] Assuming that a raw image from all different types of cameras after stitching can be mapped into a sphere, authors proposed a generic compression method based on decomposition over a dictionary of geometric atoms. A redundant dictionary is built over two generating functions (low frequency and high frequency) extended with scaling and affine transformations on a 2D sphere. The coder performs matching pursuit algorithm to select atoms from the dictionary, sorts the atoms along the decreasing magnitude of their coefficients, and then applies adaptive quantization. The proposed codec outperforms JPEG 2000 at low bitrates, however, yields to it at high bitrates.

Representing omnidirectional visual content using geometric projections, which produce less amount of data, is another approach. Example of such an approach can be found in,[5] where authors propose a rhombic dodecahedron (RD) mapping model. This convex polyhedron was chosen considering the restriction that the faces should be of quad-based nature, which allows constructing unfolded rectangular images for encoding. The model provides almost uniform pixel distribution without significant oversampling or undersampling, which allows applying DCT

and wavelet based coding more efficiently when compared to alternatives, such as cubic mapping. However, this method has not been widely adopted, because of its complexity.

Regarding solutions with JPEG backward compatibility, several such approaches[6,7] have been proposed in the past. However, none of them targets omnidirectional image compression.

As for estimation of an RoI in an omnidirectional content, there have been little to nothing in literature in the past. The few efforts in such directions have mainly focused on the definition of visual attention models. As an example, we refer to a computational model of dynamic visual attention on a sphere combining several features to detect salient locations in omnidirectional image sequences.[8,9] This solution appears to exhibit a much higher complexity when compared to visual attention models in conventional content without a rigorous analysis of its performance.

3. OMNIJPEG CODING SCHEME

In this section, the proposed OmniJPEG coding scheme is described in details, including its coding syntax and architecture. Furthermore, an implementation of a prototype application allowing OmniJPEG coding is described and its performance is analyzed.

3.1 Overview

As mentioned above, OmniJPEG exploits the RoI within an omnidirectional image, and relies on properties of the projection into a plane, to visualize an omnidirectional image content from a representative viewport by legacy JPEG decoder, while taking at the same time full advantage of the omnidirectional content by an extended decoder. Although, the proposed architecture for OmniJPEG can cope with various capture, representation, and projection alternatives, the prototype software implemented in the framework of this paper only supports 360 omnidirectional images of a spherical model projected to an equirectangular representation.

3.2 Proposed coding architecture

The proposed scheme for JPEG backward compatible coding of omnidirectional images aims at providing the ability to visualize omnidirectional content in cases when advanced omnidirectional viewing is not available. Rendering of omnidirectional image content requires efficient implementation of sphere-to-plane projections, such as the widely used equirectangular or cubic projections. Therefore, to successfully and interactively visualize a desired viewport, a 3D graphics powered viewer is necessary. To display omnidirectional image information with a conventional 2D image viewer without a need for an advanced rendering algorithm, a predefined viewport representing limited field of view of the original content has to be identified a priori. This viewport can then be encoded with a conventional 2D image compression algorithm and consequently displayed after decoding.

The block diagram of the proposed JPEG backward compatible coding scheme for omnidirectional image content, OmniJPEG, is presented in Figure 1. Input is an either compressed or uncompressed omnidirectional image in one of the sphere-to-plane geometric projections (such as equirectangular or cubic). Input data is first read into a memory block and, if needed, decompressed. Afterwards, RoI of an input image is extracted and its corresponding viewport is generated and geometrically corrected. Then, a viewport image is compressed with a conventional JPEG encoder and stored in the JPEG File Interchange Format (JFIF),[10] which can then be displayed by any legacy viewer. Subsequently, the entire omnidirectional image is stored in the same JFIF file as a metadata.

Block diagram of OmniJPEG decoder is presented in Figure 2. The process of decoding consists in extracting the original omnidirectional image from metadata of JFIF file. Obtained image can be displayed by an omnidirectional image viewer supporting OmniJPEG coding, or stored in a file in order to be displayed later by an omnidirectional visualization software.

On the other hand, Figure 3 depicts how an omnidirectional image encoded by the proposed OmniJPEG format can be decoded and displayed by both, legacy JPEG viewers or decoders supporting OmniJPEG decoding, respectively. Generally, an omnidirectional image can only be displayed by dynamically rendering the particular
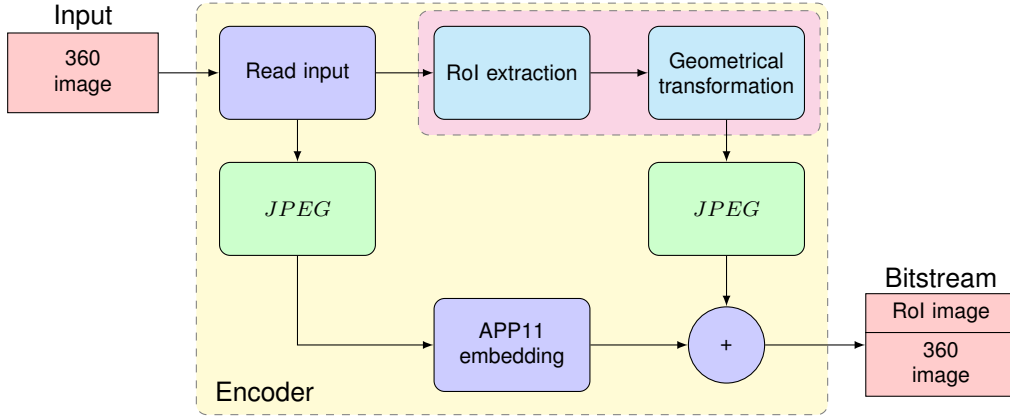
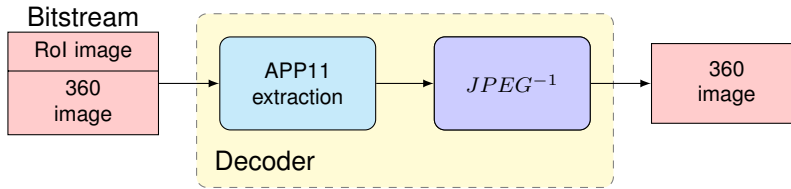Figure 1: Block diagram of the proposed OmniJPEG encoder.

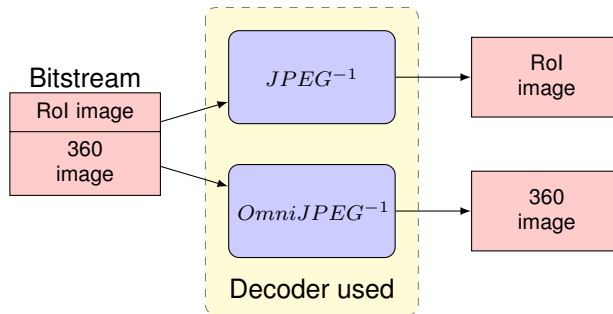Figure 2: Block diagram of the proposed OmniJPEG decoder.

Figure 3: Flexible visualization of an omnidirectional content encoded by the proposed OmniJPEG format.

viewport with a special software allowing users to freely change the view direction. However, when OmniJPEG encoded image file is opened with a conventional legacy viewer, it can only access the predefined viewport corresponding to a particular RoI specified during encoding. On the contrary, when an encoded image is opened by omnidirectional viewer, which supports OnmiJPEG, the application has access to an entire omnidirectional image stored in OmniJPEG metadata.

3.3 Implementation

To test OmniJPEG encoding process, a prototype software was implemented. The implemented software allows to encode and decode image files with the proposed OmniJPEG algorithm. Assumed input is an omnidirectional image represented in equirectangular projection. The software reads an input image to a memory block and decompress it, if already compressed.Then, it extracts the predefined RoI and renders its corresponding viewport,
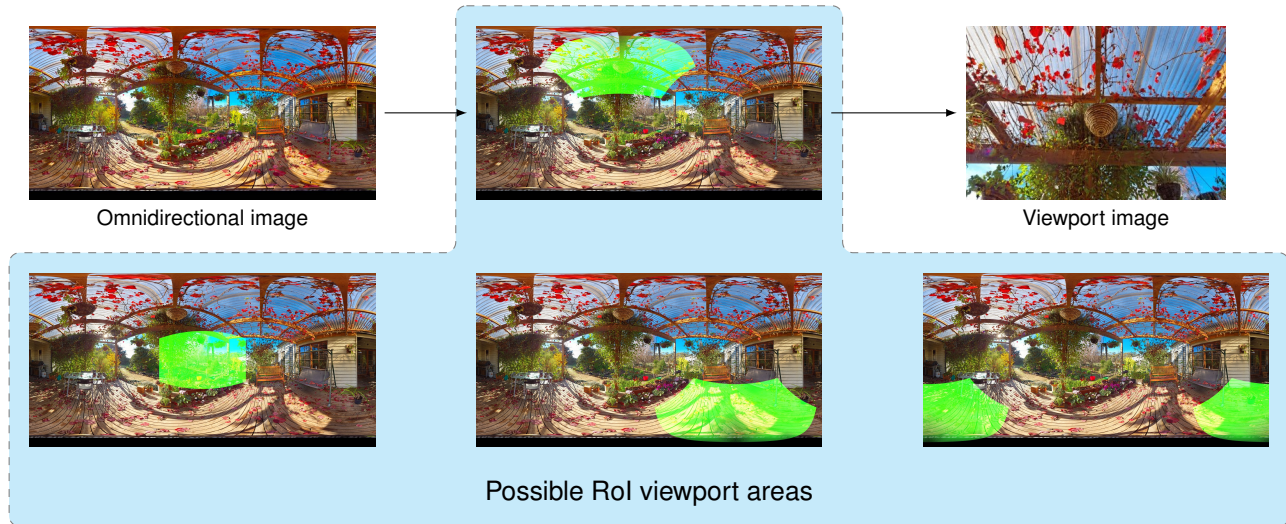
Figure 4: Examples of visualization of RoI in equirectangular projection and its counterpart viewport.

performing geometric transformations necessary for the given position within the omnidirectional image. Produced viewport is subsequently compressed with JPEG and stored in a JFIF file. The original omnidirectional image is stored in the APP11 application marker in JPEG format.[11] The developed application requires the following parameters:

- pan, as a horizontal spherical coordinate in degrees,
- tilt, as a vertical spherical coordinate in degrees,
- vertical field of view in degrees, and
- aspect ratio of the desired viewport.

The implemented prototype uses the open-source library LibJPEG to cope with JFIF file format and JPEG compression. Additionally, a graphical user interface has been developed to help in the process of OmniJPEG encoding.

To apply the geometrical transformation and to produce undistorted viewport for specified RoI of omnidirectional image, we use the pinhole camera model. Omnidirectional image represented in equirectangular format is mapped to a sphere first. Then, perspective projection transformation, which removes the geometric distortions, is applied either to a manually defined or to an automatically extracted RoI. Examples of RoI areas in equirectangular projection and their viewport counterparts extracted with the developed prototype are shown in Figure 4. Desired RoI, highlighted in green color (see Figure 4), is processed, all geometric distortion related to sphere-to-plane projection are removed, and the corresponding viewport representing omnidirectional image is produced.

3.4 Performance Evaluation

Table 1 contains results of proposed backward compatible coding scheme for omnidirectional visual content. For five different contents in their original resolution (see Section 4.2.1 for details), we calculate the file size overhead after encoding with OmniJPEG. Three different relative viewport sizes with vertical and horizontal field of view, corresponding to 30×60, 60×90, and 90×120, are used. As expected, file size is increasing with increasing size of the desired viewport. Another factor affecting the file size overhead is the actual content in the viewport, because, for example, the amount of details in the picture can decrease JPEG compression efficiency. Results show an average overhead of approximately 8% for desired viewport size of 60×90.

Table 1: OmniJPEG performance for tested dataset (see Section 4.2.1). Overview of file size overhead for different sizes of extracted RoI.

Content		C1	C2	C3	C4	C5
file size [MB]	original file	6.9	5.6	4.4	4.5	0.3
	30x60	7.1	6.0	4.5	4.6	0.3
	60x90	7.4	6.6	4.6	4.7	0.3
	90x120	7.9	7.5	4.8	4.8	0.3
file size overhead [%]	30x60	3.2	5.7	1.6	1.5	3.4
	60x90	7.2	17.9	4.8	3.7	6.9
	90x120	14.2	33.6	9.8	6.6	10.3
resolution WxH [px]	original file	6000×3000	6512×3256	4000×2000	4096×2048	2048×1024
	30x60	1000×500	1085×543	667×333	683×341	341×171
	60x90	1500×1000	1628×1085	1000×667	1024×683	512×341
	90x120	2250×1500	2442×1628	1500×1000	1536×1024	768×512

4. REGION OF INTEREST OF OMNIDIRECTIONAL IMAGES

In this section, proposed method to estimate and extract RoI from an omnidirectional image content is presented. Then, the experiment to collect a ground truth for RoI obtained manually is described. The RoI prediction performance is then analysed and compared to experimental results.

4.1 Region of Interest Estimation

Estimation of RoI of an omnidirectional image data is linked to a prediction of desired viewports which are more likely to be visualized by users. Interactive consumption of an omnidirectional visual content allows users to actively scan the image data and visualize individual viewports. At each time instance, users can see only one viewport maintaining visual fixation or visual gaze on various locations. Visual attention within each viewport can be modeled by available saliency models for regular 2D images predicting scene locations where users may fixate. However, to take into account an interactive consumption and active image data scanning the size and visualization frequency of the displayed viewport has to be taken into account.

Since generation of saliency map of every possible viewport is a time consuming process, we apply the model addressing the fixation prediction on the entire equirectangular representation of omnidirectional content and create only one saliency map per content. To generate the saliency map, a Matlab implementation of GBVS [*] was used. This model uses graph algorithms and a measure of dissimilarity to outperform its competitors when achieving visual fixation prediction and efficient saliency computation. To address the users interaction with omnidirectional content in terms of RoI extraction a sliding window method was applied. More specifically, a moving average filter with parameters corresponding to viewport size was applied on the acquired saliency map. Coordinates with maximum filtered saliency value locate the middle of the desired viewport or RoI.

The size of a sliding window is determined by the portion of the image displayed in a viewport. We selected $90° \times 60°$ as a size of the sliding window for our computations since most of the available 360 video and image viewers support such size of the rendered viewport as a default option. The above selected size in visual degrees corresponds to one third and one fourth of original equirectangular image height and width, respectively. It should be noted that the implementation of sliding window needs to take into account the 360 nature of omnidirectional image data. More particularly, the sliding windows must be applied on the wrapped saliency map, where its both left and right edges are connected unlike its top and bottom edges. The wrapped saliency ma can be imagined as a cylinder surface without its bases.
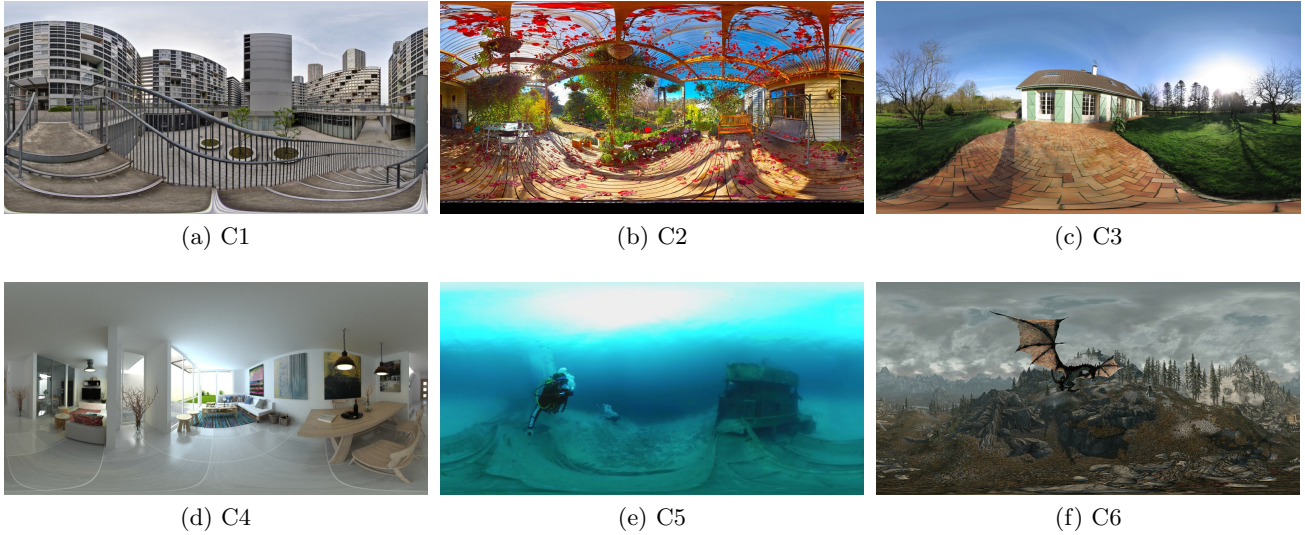
---

[*]`http://www.vision.caltech.edu/~harel/share/gbvs.php`

(a) C1   (b) C2   (c) C3

(d) C4   (e) C5   (f) C6

Figure 5: 360 equirectangular test images used in experiments.

4.2 Experiment

To validate the results of RoI estimation, we conducted a subjective experiment, where ground RoI was collected. Ground truth consists of the tracking data of each user, which allows extraction of viewports to be visualized during the media consumption. Collection of the subjective data was performed using proprietary tool developed for this purpose.

4.2.1 Dataset

The dataset described in this section was used not only for subjective experiments but also for automatic RoI estimation and RoI extraction using the method described in Section 4.1.

Six different full panoramic and publicly available images in equirectangular representation (see Figure 5) were used in subjective experiments[†]. Contents C1-C5 were used as test images, whereas content C6 was used for training purposes. Test images were selected based on a careful analysis of their content. More specifically, the amount of spatial information (SI)[12] and colorfulness (CF) index[13] were computed for each image content. The distribution of test images along their corresponding SI and CF values are presented in Figure 6. Images were originally available in JPEG format with transparent visual quality in their native resolution presented in Table 1. Prior to the experiments, all images were resampled to the resolution of 4096×2048 pixels.

4.2.2 Methodology

The experiments consisted in two parts interleaved in one experimental test session. In both parts of the experiments, subjects viewed each content once, however, different consumption time was set for different parts of experiments. More specifically, two parts of experiments to collect tracking information and viewing direction of each subject were designed in the following way.

[†]C1 - `https://www.flickr.com/photos/heiwa4126/3551485972/sizes/l`

C2 - `http://www.lemis.com/grog/photos/Onephoto.php?image=Photos/20120601/verandah-centre-equirectangular.jpeg&size=4`

C3 - `https://www.flickr.com/photos/alexispz/376913871/sizes/l`

C4 - `https://www.flickr.com/photos/141405686@N05/25862086610/sizes/l`

C5 - `https://www.youtube.com/watch?v=aQd41nbQM-U`

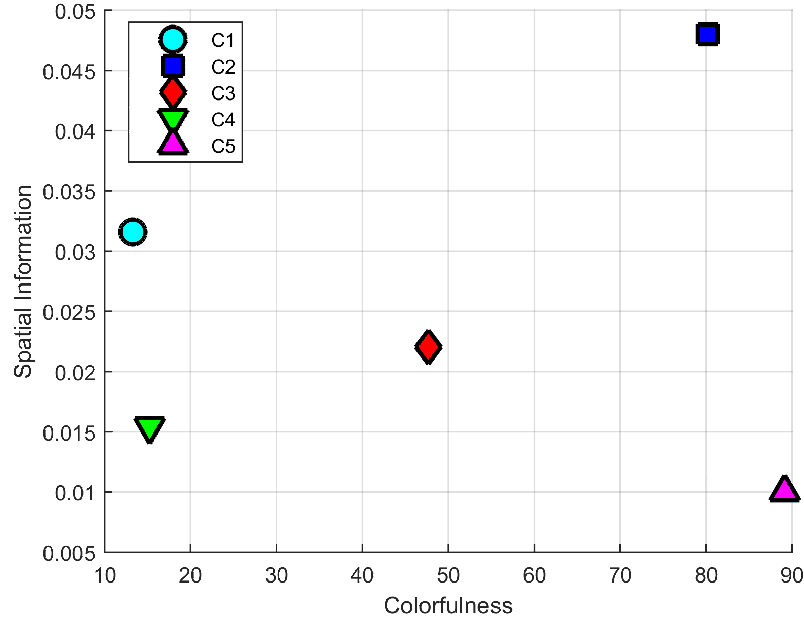C6 - `https://www.flickr.com/photos/pedroscreamerovsky/6806942120/`

Figure 6: Spatial information index and colorfulness index for each test content.

- **Free viewing experiment**: First, subjects were asked to navigate through each full panoramic image freely with emphasis on screening of an entire picture. This helps subjects to get familiar with the content as well as to be aware of regions of the image which they like most. Each content was visualized for 30 seconds in the first experiment. This amount of time was shown to be enough to comfortably scan over the entire panoramic image.

- **Targeted viewing experiment**: After performing the first viewing, subjects were immediately asked to watch the content again. This time, subjects were asked to navigate, in a limited amount of time, directly to the area with the most appealing visual information. In this part of the experiments, each content was presented for 20 seconds only.

Thus, subjects watched each content continuously for approximately 50 seconds with a short break after first 30 seconds. Before second part of the experiments for each content, subjects were reminded about the goal of targeted viewing part and its limited duration. Additionally, a training session was organized before the experiments to allow participants to get familiar with the visualization interface and especially with the way how to navigate and scan over the image content. The training sample was not used in the experiments in order not to bias subjects.

4.2.3 Test environment and equipment

The subjective experiments were performed in home-like scenario conditions with semi-controlled lighting system assuring no screen reflections. To display the full panoramic images, an iPad4 with native resolution of 2048×1536 pixels was used. Viewing distance was set to approximately 40 cm, which corresponds to a foveated vision area ($2°$ of visual angle) with resolution of 140 pixels in relation to native resolution. Full panoramic content was rendered using a testbed, specifically implemented for these experiments. Testing interface allows user to navigate through the image content by swiping their finger over the screen of the iPad4, to track the viewing direction, to set the size of viewport and its projection, and to limit the visualization time. The viewing direction of each subject and content was recorded via two coordinates: pan and tilt, representing horizontal and vertical coordinates of the middle of the viewport, respectively. In our experiments, omnidirectional visual contents were displayed with a vertical field of view set to 60 degrees.

Figure 7: Resulting saliency maps for each content. Individual content, estimated saliency map, and saliency maps from first and second experiment are shown in first, second, third and fourth columns, respectively.

4.3 Results

Below, the results of RoI prediction as well as results of conducted subjective experiments are presented and discussed.

In the second column of Figure 7, the saliency maps generated with GBVS for each test image are shown. The intensity distribution of these saliency maps corresponds to probability with which the content in equirectangular representation would be consumed if rendered and displayed as a regular 2D image.

To be able to compare results of RoI estimation with results of our experiments, collected tracking information are converted into saliency maps. Saliency maps are created by filtering the tracking information corresponding to visual fixation with a Gaussian kernel corresponding to foveal vision. All recorded coordinates of viewing directions were considered as gaze points and classified as visual fixation in our analysis. Moreover, we assumed the foveal vision with respect to native resolution of rendering device and full panoramic content resolution. Therefore, σ parameter of the Gaussian kernel was set to 70 pixels.

Third and fourth columns of Figure 7 present average saliency maps resulting from free and targeted viewing experiments, respectively. Whereas the first free viewing experiment reveals general consumption behavior of each subject, the second experiment, focused on targeted viewing, shows the actual highlighted segments of each content in terms of visual attraction. The experimentally detected coordinates of maximum saliency do not always correspond to estimation results. However, the maximum of a saliency map does not, in this case, reflect the actual position of RoI. As explained above, it is due to the fact that it does not take into account the nature of omnidirectional content rendering, more specifically, the size of the rendered viewport. The RoI can

Table 2: Comparison of tilt and pan coordinates of the middle of estimated and experimental RoI. Ex1 and Ex2 row shows results for free and targeted viewing experiment, respectively.

Content	C1		C2		C3		C4		C5	
	tilt [°]	pan [°]	tilt [°]	pan [°]	tilt [°]	pan [°]	tilt [°]	pan [°]	tilt [°]	pan [°]
model	**22.14**	**93.53**	**10.89**	**33.41**	**10.69**	**-26.42**	**-11.61**	**9.51**	**-0.18**	**-62.75**
Ex 1	22.14	129.38	-11.61	21.46	-0.56	-2.23	-11.61	21.45	-22.68	-50.79
Ex 2	**33.39**	**105.48**	**-11.61**	**21.46**	**-0.56**	**-38.27**	**-11.61**	**9.51**	**-0.18**	**-62.75**

be extracted by applying the sliding window approach, described in Section 4, and by subsequent detection of the maximum saliency coordinates.

Table 2 compares the pan and tilt coordinates of RoI resulted from automatic prediction and from both parts of conducted experiments, assuming that the middle of the equirectangular full panoramic image has coordinates of $[0, 0]$. Additionally, Figure 8 visualizes automatically estimated and actual (results of targeted viewing experiment) RoI and corresponding viewport for each content. When looking at the values in Table 2 and when visually comparing RoI areas and corresponding viewports for each test image, we see that the representative part of full panoramic image content can be successfully predicted in an automatic way.

5. CONCLUSION

In this paper, we propose OmniJPEG as a JPEG backward compatible solution to encode the omnidirectional images. To ensure the JPEG backward compatibility, OmniJPEG exploits the RoI from omnidirectional images as well as properties of equirectangular projection, while at the same time also providing the complete equirectangular information to preserve the capability of correctly rendering the 360 image with appropriate devices. Additionally, the methodology for an automatic estimation and extraction of the RoI is presented and evaluated. Results with the tested image dataset show that the RoI on an omnidirectional image can be successfully estimated with a relatively high accuracy. Nevertheless, the conducted experiments as well as developed prototype implementation takes into account only equirectangular representation of spherical omnidirectional images. Other equirectangular formats, as well as an exploration of the possibility of various sphere-to-plane mapping, which removes the redundant data from equirectangular representation, will be addressed in future.

ACKNOWLEDGMENTS

REFERENCES

[1] Harel, J., Koch, C., and Perona, P., "Graph-based visual saliency," in [*Advances in neural information processing systems*], 545–552 (2006).

[2] Chen, S. E., "QuickTime VR: An image-based approach to virtual environment navigation," in [*Proceedings of the 22Nd Annual Conference on Computer Graphics and Interactive Techniques*], *SIGGRAPH '95*, 29–38, ACM (1995).

[3] Rondao Alface, P., Macq, J.-F., and Verzijp, N., "Interactive omnidirectional video delivery: A bandwidth-effective approach," **16**(4), 135–147 (2012).

[4] Tosic, I. and Frossard, P., "Low bit-rate compression of omnidirectional images," in [*Picture Coding Symposium, 2009. PCS 2009*], 1–4 (2009).

[5] Fu, C.-W., Wan, L., Wong, T.-T., and Leung, C.-S., "The rhombic dodecahedron map: An efficient scheme for encoding panoramic video," **11**(4), 634–644 (2009).
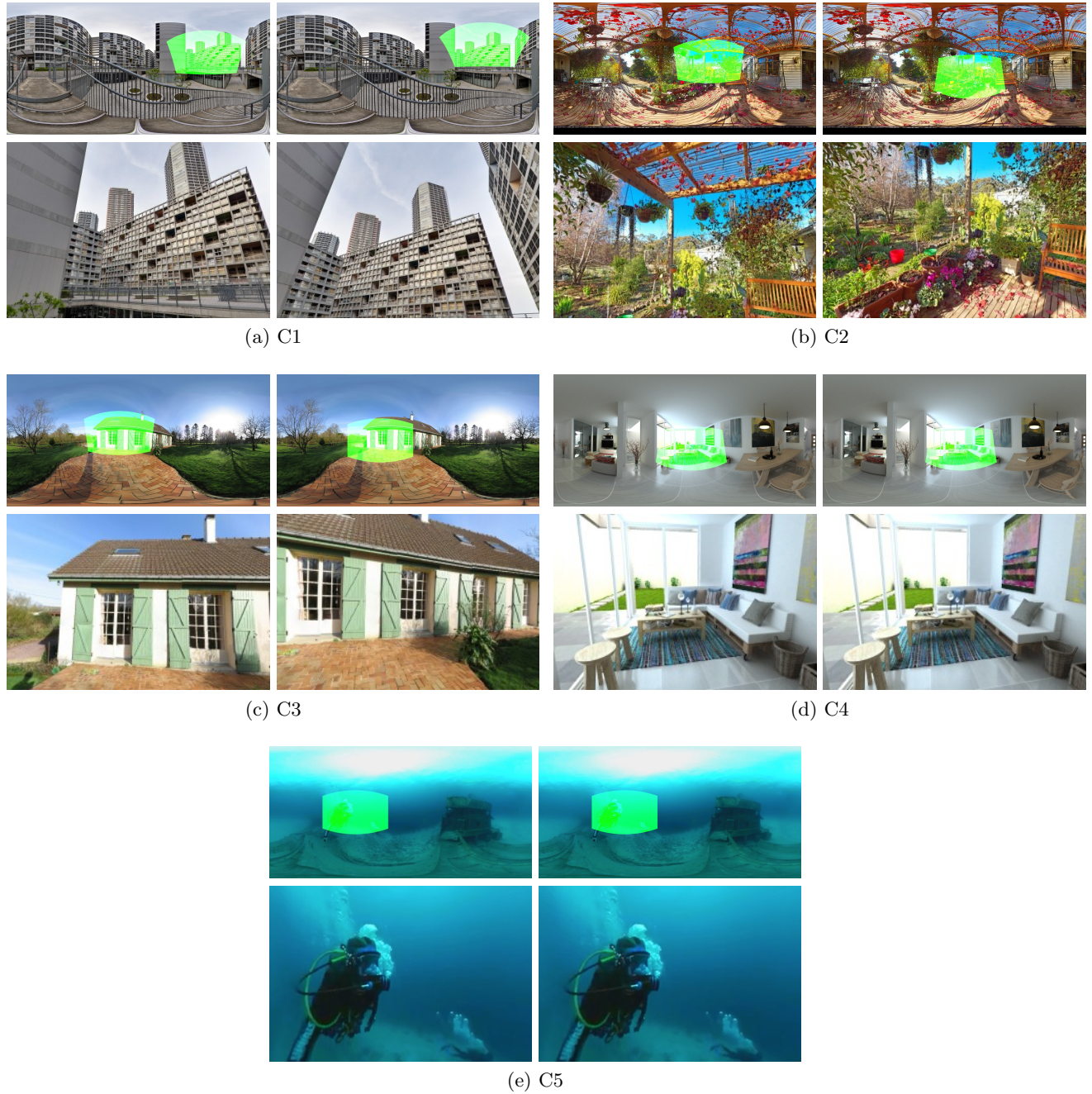
(a) C1

(b) C2

(c) C3

(d) C4

(e) C5

Figure 8: Visualization of estimated and actual RoI in equirectangular and viewport representation for each content. Left columns of each subplot show automatically estimated RoI using the saliency prediction model, whereas right columns present experimentally obtained RoI.

[6] Hanhart, P., Korshunov, P., Rerabek, M., and Ebrahimi, T., "JPEG backward compatible format for 3d content representation," 1–4, IEEE (2013).

[7] Korshunov, P. and Ebrahimi, T., "Context-dependent JPEG backward-compatible high-dynamic range image compression," **52**(10), 102006–102006 (2013).

[8] Bogdanova, I., Bur, A., and Hugli, H., "Visual attention on the sphere," **17**(11), 2000–2014 (2008).

[9] Bogdanova, I., Bur, A., Hugli, H., and Farine, P.-A., "Dynamic visual attention on the sphere," **114**(1), 100–110 (2010).

[10] ITU-R T.871, "Information technology - digital compression and coding of continuous-tone still images: JPEG file interchange format (JFIF)," (2011).

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

[12] ITU-T P.910, "Subjective video quality assessment methods for multimedia applications." International Telecommunication Union (April 2008).

[13] Hasler, D. and Suesstrunk, S. E., "Measuring colorfulness in natural images," in [*Electronic Imaging 2003*], 87–95, International Society for Optics and Photonics (2003).