ISO/IEC JTC 1/SC29/WG1 **N78016**
78th JPEG Meeting, Rio de Janeiro, Brazil, 29 January – 2 February, 2018

ISO/IEC JTC 1/SC 29/WG 1
(ITU-T SG16)

Coding of Still Pictures

JBIG	**JPEG**
Joint Bi-level Image Experts Group	Joint Photographic Experts Group

TITLE:	Call for Proposals on JPEG 360 Metadata
SOURCE:	JPEG Requirements Subgroup
PROJECT:	ISO/IEC 19566 (JPEG Systems)
STATUS:	Final
REQUESTED ACTION:	Approve for publication
DISTRIBUTION:	Public

Contact:
ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch

Call for Proposals on JPEG 360 Metadata

1 Motivation

The JPEG Committee notes the increasing use of multi-sensor images from multiple image sensor devices, such as 360 degree capturing cameras or dual-camera smartphones available to consumers. Images from these cameras are shown on computers, smartphones, and Head Mounted Displays (HMD). JPEG standards are commonly used for image compression and file format. However, because existing JPEG standards do not fully cover these new uses, incompatibilities have reduced the interoperability of these images, and thus reducing the widespread ubiquity which consumers have come to expect when using JPEG-based images. Additionally, new modalities for interacting with images, such as computer-based augmentation, face-tagging, and object classification, require support for metadata that was not part of the original JPEG scope.

To avoid fragmentation in the market and to ensure wide interoperability, a standard way of interacting with multi-sensor images with richer metadata is desired for JPEG standards. This CfP invites all interested parties, including manufacturers, vendors and users of such devices to submit technology proposals for enabling interactions with multi-sensor images and metadata that fulfil the scope, objectives and requirements outlined in this Call for Proposal.

2 JPEG Systems Approach

Within several standards of the JPEG Ecosystem there is a need for including additional metadata in different formats. Especially in the older JPEG standards there was no clear defined way for supporting third party metadata. For these reasons, various industry consortia defined other ways of metadata support like EXIF or support for storing multiple images like MPO.

We are approaching this broad topic in a measured manner. For the future, metadata and multi-image support shall be offered within the standard itself. The approach needs to be scalable and sustainable, and it also needs to be timely. Ideally, the solution used for JPEG 360 metadata is generalizable to be applied to other sorts of metadata needed to supply emerging and future needs. For example, additional metadata support is relevant for the new upcoming standard JPEG Systems 19566 Part 4: Privacy and Security.

It is the goal within a broader effort of JPEG committee to harmonize system level functionalities across all JPEG standards. For this reason a universal metadata box format shall be defined in a new part of JPEG systems, which can be used according to the JPEG system approach in legacy JPEG 1/JPEG XT standards within the App11 marker as well as in box-based standards. The new part shall cover all aspects of metadata handling in JPEG standards.

Box based formats like JPEG 2000 or the ISOBMFF (*.MP4) allow a high flexibility with low restrictions in their usage. For this reason, future standards using a box-based approach shall be kept if possible. On the other side also legacy formats like JPEG 1 (ISO 10918-1) shall benefit from future extensions. For this reason a new way of integration of boxes in APP markers of JPEG 1 was defined already in ISO/IEC 18477-3 (JPEG XT Part 3); it uses the reserved APP11 marker for future extensions.

3 Metadata Requirements

3.1 Metadata Functional Requirements

Metadata Requirements and Description	
Store multiple images within a single file	The standard shall support the ability to declare and discover the association to multiple images within the same file. The current JPEG file definition allows for only a single image to be contained within it, whereas for JPEG 360, there is need to keep multiple images within the same file. An example are the images from individual cameras which are stitched together to form a single complete spherical (or semi-spherical) image. Another example is a conventional (flat) higher-resolution image (or a pair of such images for the stereoscopic case) which is linked to a region of pixels on a 360 image.
Projection type	The standard shall support the ability to declare and discover the projection type for images contained in the file. It must also support declaration and discovery of the spans of the image within the projection. Several types of representations (projections) are used to map an omnidirectional image to a 2D planar surface (e.g, equirectangular, cubic, pyramidal, dodecahedron, etc). The projection mapping needs to be known with sufficient detail for it to be possible to render an omnidirectional image. This can include geometry description (mesh) and texture coordinates. Images may be only a portion of a projection. E.g., a 180 image is a hemisphere. Further, during editing, an image can become an even smaller portion of the projection, as might be the case in a photo collage onto a spherical projection.
File format of the additional images	The standard shall support the ability to declare and discover the file format used by the additional images contained in the file. Systems-level information about how and where an additional image is stored in the metadata of a JPEG360 file is required to properly extract and parse the additional image. For example, an additional image may be stored using JPEG-1 or JPEG XT file formatting, or in another form.

Metadata Requirements and Description	
Coding format	The standard shall support the ability to declare and discover the image coding format used to compress the images contained in the file. The additional images being stored in the JPEG 360 file are likely to have been compressed in order to reduce the overall file size. In order to use the compressed images, we need to know which compression has been applied. These may already be included in the definitions of storage format.
Pixel format	The standard shall support the ability to declare and discover the pixel format of the images contained in the file. Further details of the compression, such as pixel component interleave ordering, pixel color space, and color gamut, are needed to reproduce the uncompressed images. These may already be included in the definitions of the compression codec and storage format.
Viewport position upon opening image	The standard should support the ability to declare and discover the initial position in the projection rendering which is viewed when the image is opened. The ability to set the viewport position allows for the primary subject of the image to be seen first when the image is rendered.
Associate a region of pixels to an interactive action	The standard should support the ability to declare and discover the association of a region of pixels with an action, such as a mouse or controller, to create interactivity with image elements. This feature allows for interactions by the user with elements seen in the image. For example, putting cursor over a person’s face could bring up an overlay of that person’s name, or clicking on a door seen in the image could cause the image to be changed to an image taken in the adjacent room of a house.
Camera pose relative to Earth reference	The standard should support the ability to declare and discover the pose of the camera relative to a local reference frame. Cameras are increasingly provided with a variety of sensors, such as a gravity sensor and Eath’s magnetic field, from which the camera’s overall orientation can be determined. This information can be helpful with image stabilization, or setting the nominal image horizon.

Metadata Requirements and Description	
Location data	The standard should support the ability to declare and discover a location that is associated to an image, or a set of images, or a set of locations to a set of images. The notion of location occurs at many abstraction levels, ranging from a set of longitude, latitude, altitude coordinates to building address, neighborhood, city, region, country, continent, planet, solar system, galaxy, etc. Location is often combined with other sensor data, such as pose, to further refine the oberved scene.
Date & Time	The standard should support the ability to declare and discover a timestamp that is associated to an image, or a set of images, or a set of timestamps associate to a set of images.
Camera sensor data	The standard should support the the ability to declare and discover associated sensor data. Many cameras have built-in or access to nearby sensors to monitor acceleration, velocity, temperature, underwater depth, humidity, heart rate, respiration rate, blood pressure, GPS, etc.
Relative sensor arrangement for multi-sensor capture	The standard should support the ability to declare and discover relative sensors positions a suitable coordinate system, including a 2D grid array. Assuming the availability of depth information, this information should allow to transform the image of one camera/sensor to the viewing position of another camera/sensor. This allows for stitching multiple images into a larger image, such as a spherical image, or partial spherical image, cube-map image, etc. under possible consideration of occuring parallax. Examples of multisensor cameras include Google VR180), and spherical (Ricoh Theta series). It is also possible to use a time-series of captures using a typical 2D camera which can be stitched together to create panoramas to various projections.
Exposure time	The standard should support the ability to declare and discover the exposure time that is associated to each image captured within the same file.

Metadata Requirements and Description	
Embed video clips	The standard shall support the ability the ability to declare and discover multiple videos within the same file. 360 cameras typically can capture high resolution still and lower resolution video images. The viewer's experience and understanding of the image can be enhanced if it's possible to see a relevant video clip that provides more context for the image.
Embed audio-only clips	The standard should support the ability to declare and discover an audio clip associated to an image or a set of images within the same file. A brief clip of sounds captured concurrent with the image capture can be used to enhance the experience of the image; e.g., the trickling of water, birdsong, and spoken narrative to explain the image content.
Director's view for scripted experience	The standard should support the ability to declare and discover scripted experience of the 360 image within the same file. The viewport of a spherical image is usually smaller than the total viewable area, and several items of interest may fall outside the initial viewport setting. A script that guides the viewport through the interesting portions of the image (a director's cut) can increase the viewer's enjoyment and understanding of the image.
Lens parameters	The standard should support the ability to declare and discover the optical system parameters that describe the image formation onto the sensor within the same file. In order that the images can be stitched and properly mapped to the desired projection requires knowledge of the distortions, focal length, modulation transfer function, and other optics information.
Embed stereo 360 images	The standard should support the ability to declare and discover the association to stereo images within the same file.
Embed stereo 360 video clips	The standard should support the ability to declare and discover the association to stereo video clips within the same file.
Support JPEG Systems Part 4 for Privacy and Security	The standard should support the ability to declare and discover the metadata needed to support JPEG Systems Part 4 – Privacy & Security.
Default render mode wrt available images and display devices.	The standard should support the ability to declare and discover the preferred characteristics of the display device. If the image was created to be experienced on a head mounted device (HMD), and both a 2D monitor and HMD are connected as display devices, the image should be preferentially shown on the HMD.

Metadata Requirements and Description	
Stitching software & version	The standard may support the ability to declare and discover the relevant information about the origin and version of software which was used to generate the image projections, people tags, object tags, etc. Stitching algorithms are in rapid transition. If the component images are available, and an improved stitching algorithm is available, the stitched image can be updated with a better quality version.
Link to external data sources for the additional images	The standard may support the ability to declare and discover links to images which are stored in a separate file (local or remote). For sharing images through email, or for uploading to photosharing sites, it may be desirable to reduce the overall size of the file by removing some of the embedded files, and replacing them with links (e.g., URIs) so to represent the deleted images. Likewise, to protect anonymity, it might be helpful to remove face taggings, but to keep a reference as to where those tags are available.
Embed HTML5 scripts	The standard may support the ability to declare and discover the HTML5 scripts. When using a browser to view images, customing the experience through HTML5 code improve the experience, for example, to create a scrapbook feel.
Support encrypted data	The standard may support the ability to declare and discover the presence of encrypted data. Hardware and software vendors may want to protect metadata they have added to the images, and to make it accessible under certain restrictions. End users may want to protect sensitive metadata from unintended uses.
Embed spatial audio clips	The standard may support the ability to declare and discover the presence of spatial audio. Spatial audio, aka 3d audio, clips give a more immersive experience and can be used as another way to guide users through a 360 image.
Support for medical images	The standard may support the ability to declare and discover the presence of an atypical image in the same file.
Support authenticated data	The standard may support the ability to declare and discover the authenticity of the images and/or metadata in the file.
Dynamic Viewport -- programmatic	The standard may support the ability to declare and discover the presence of a dynamically generated script for the viewport behavior. Similar to the director's cut, except this script is external to the image file. At the time the image is opened, the metadata provides an API through which the viewport behavior is scripted.

Metadata Requirements and Description	
In-image data tags	The standard may support the ability to declare and discover the metadata which is derived from the images within the same file. The meeting attendance list can be constructed by finding and identifying the faces seen in a 360 image taken from the middle of a conference room table; this list can be included within the file for future reference.

3.2 Metadata Constructional Requirements

Metadata Requirements and Description	
Platform independence	The standard shall be usable on a wide variety of hardware and operating systems.
Compatible to JPEG-1 compliant decoder	The standard shall be compatible with JPEG-1 compliant decoder in that method used to add the metadata does not degrade the expected operation of a JPEG-1 compliant decoder. When using a JPEG-1 compliant decoder to read a file with the new metadata, the JPEG-1 compliant decoder is able to render the expected image, and it ignores the new metadata.
Compatible to JPEG XT compliant decoder	The standard shall be compatible with JPEG XT compliant decoder in that method used to add the metadata does not degrade the expected operation of a JPEG XT compliant decoder. When using a JPEG XT compliant decoder to read a file with the new metadata, the JPEG XT compliant decoder is able to render the expected image, and it ignores the new metadata.
Compatible to be used within JPEG Box File Formats	The standard shall be compatible and fully functional with the JPEG Box File structure. The proposal may also work with files structured differently from the JPEG Box File Format, but this is not mandatory.
Compliant to JPEG Systems guidelines ISO/IEC 19566-1	The standard should be compliant to the guidelines in ISO/IEC 19566-1.
Compatbility with existing JPEG-1 metadata functionality	The standard should provide for compatibility with the functionality provided by existing metadata.
Use existing ISO metadata formats.	The standard should use existing standards with respect to the presentation of metadata information such as XMP.

Metadata Requirements and Description	
Compatibility with long term metadata standard roadmap	The standard should support the addition of metadata required to support new usages.
Metadata contained within a BOX, and be capable to reference (meta)data in different BOXes.	The standard should support metadata being contained within a JPEG Systems compliant BOX format, and should have the ability to refer to (meta)data contained in other BOXes. Metadata will be contained within a well-known BOX. This metadata can refer to data and metadata contained in other BOXes within the same file, or BOXes in other files. It is also possible that when referring to an image, that referenced image may also contain metadata; i.e., recursive metadata.
Robust to a variety representations of the metadata.	The standard should provide robust with respect to formats and data type differences to represent the same information.
Usable with JPEG Box File Formats AND non-JPEG file formats	The standard may work also be compatible to non-BOX format files.
Non-Box Image File Directory (IFD) container	The standard may work also be compatible to these particular non-BOX format files without regard to other non-BOX format files.

3.3 Royalty-Free Goal

The royalty-free patent licensing commitments made by contributors to JPEG 1 has arguably been instrumental to its success. JPEG expects that similar commitments would be helpful for the adoption of JPEG 360.

4 Proposal Submissions

4.1 Submission Requirements

Proposals must be submitted by March 30, 2018 according the following guidelines.

The submission shall include:

- A high-level technical description of the submission,
- For each of the requirements listed above, a response of YES, NO, or PARTIAL
    - YES – fully meets the requirements
    - NO – no support for this requirement in submission,
    - PARTIAL – some aspects are supported,
- Elaboration of YES/NO/PARTIAL by the proposed solution with respect to the relevant requirement,
- Software architecture overview, including block diagrams of metadata and file structure,

- An analysis demonstrating that the submission meets all requirements or at least can be transformed to the desired structure,
- Source code and documentation for the proposal need to be submitted via ISO LiveLink as an input document to the 79th JPEG meeting. If a submitter does not have access to ISO LiveLink, please contact one of the people listed below.
- If possible, sample code demonstrating support for one or more use cases.

Materials submitted to, or made available as part of, the JPEG 360 Metadata activity:

- Shall not be used for any purpose other than strictly necessary to develop the JPEG 360 metadata standard. Unsanctioned uses include competitive comparison, marketing, external publication, etc.

4.2 Evaluation of proposals

Assessment will be based on the per-requirement response (YES, NO, or PARTIAL), along with the additional elaborations provided by the submitter as described in the Submission Requirements.

Considerations will be given to the perspectives of the JPEG Systems and JPEG Requirements subgroups. Physical attendance of at least one representative of the proponents at the JPEG meeting where evaluations are carried out is required.

4.3 Contribution to Standardization

Proponents are informed that based on the submitted proposals, a standard specification will be created. If they submit a proposal and (part of) the proposed technology is accepted for inclusion in the standard, they will be expected to attend subsequent WG1 meetings and contribute to the creation of the relevant documents. Within this process, evolution and changes are possible as several technologies may be combined to obtain a better performing solution.

5 IPR conditions (ISO/IEC Directives)

WG1 requests submissions of IPR statements relevant to all JPEG standards under development in accordance with the Common Patent Policy for ITU-T/ITU-R/ISO/IEC. Further details can be found at www.itu.int/ITUT/dbase/patent/patent-policy.html or ISO/IEC Directives Part 1, Appendix I.

6 Workplan for JPEG 360 Metadata Recommendation

The table below specifies the JPEG 360 workplan, including submission deadlines. The technical committee intends to prioritize consideration of those submissions that meet these deadlines.

October 2017	Draft Call for Proposals.
January 2018	Final Call for Proposals.

February – March 2018	AHG meetings scheduled per ISO requirements as needed to prepare the assessnment criteria for evaluating submissions.
March 02, 2018	Pre-registration for proponents to indicate intention to submit a proposal. (Establish account for uploading, clarify items in the CfP, etc.)
March 26, 2018	Submission period for proposals closes on March 26, 2018, 11:59PM Pacific Time.
April 9-10, 2018	AhG meetings with in-person presentations by proposers.
April 11 - 15, 2018	Complete evaluation of submissions, and issue a WD for ISO/IEC 19566-6: Information technologies – JPEG Systems -- Part 6: JPEG 360.
July 2018	CD ISO/IEC 19566-6: Information technologies – JPEG Systems -- Part 6: JPEG 360 (Request 2-month balloting at 80th JPEG Meeting)
October 2018	DIS ISO/IEC 19566-6: Information technologies – JPEG Systems -- Part 6: JPEG 360 (Request 5-month balloting) at 81st JPEG Meeting.
April 2019	IS ISO/IEC 19566-6: Information technologies – JPEG Systems -- Part 6: JPEG 360 at 83rd JPEG Meeting. (Assumes zero "no" votes, otherwise FDIS.)

The above schedule is subject to change, depending on the nature of proposals that are received and the possible need to integrate or merge elements from separate proposals.

7 Further information

JPEG 360 Ad hoc Group

The JPEG 360 Ad Hoc Group was established at the 77th JPEG meeting in October 2017, following and will be used as a forum to coordinate and exchange information until 78th JPEG meeting in January 2018. In particular the JPEG 360 Ad Hoc Group is charged with updating this Draft CfP, collecting use cases and evidence for both requirements and technologies, and facilitating the process leading to the evaluation of proposals, as identified in the above timeline. The AhG then submits recommendations to be approved at the 78th JPEG meeting. All interested parties are requested to register to the email reflector of the JPEG 360 Ad Hoc Group at: https://listserv.uni-stuttgart.de/mailman/listinfo/JPEG-360

Contacts

Touradj Ebrahimi (JPEG Convenor)
Email: Touradj.Ebrahimi@epfl.ch

Fernando Pereira (JPEG Requirements Subgroup Chair)
Email: fp@lx.it.pt

Siegfried Fößel (JPEG Systems Subgroup Chair)
Email: Siegfried.foessel@iis.fraunhofer

Andy Kuzma (JPEG 360 Ad hoc Group Chair)
Email: andy.kuzma@intel.com

Annex A – Use Cases

JPEG 360 should enable use cases based on multi-sensor images captured, e.g. omnidirectional cameras, dual-camera smartphones, or stereoscopic cameras. The following summarizes some, but not necessarily all, use cases specifically enabled by JPEG 360.

1 Action Sports Enthusiast

François uses 360 degrees panorama camera to record his paragliding adventures. His flight equipment (altimeter, barometer, compass, GPS, etc.) are streamed to the camera for capture. The 6DOF sensor for the camera rig is also streamed for capture. And the biometric sensors for the pilot are streamed to the camera for capture. Let's assume the camera is set to capture a high resolution 360 degrees panorama image every 1 second, along with the aggregated sensor data.

When François returns to ground, the camera data is uploaded to a personal computer and/or cloud service. The basic sensor data can be used to link an image to a position on a map. François, however, subscribed to an enhanced service, running on his personal computer which able to determine which landmarks are visible in an image and augment with image with additional labels; e.g., Mont Blanc. The derived metadata components (Mont Blanc's position in the image, and the position of the label in the image) are new metadata that are to be added to the François' images to create enhanced interactions. François' children enjoy learning geography when they see the images.

2 Engineering Smart Assistant

Tom is an engineer in charge of structural engineering for reinforced concrete. As there are many aged reinforced concrete constructions to be regularly inspected, he really feels that it is necessary to raise the efficiency of inspection. Tom has been storing a huge number of sample pictures regarding deterioration of concrete constructions. As a new technology trend, Tom has heard that advances in image recognition technology will become practical on mobile devices. He decided to take in the new technology to improve its effectiveness. Tom took pictures with his smart phone, being careful to capture them with high fidelity to improve the analyses. Below is a sample picture applied for a dam, but the images and analyses are similar for tunnels and bridges.

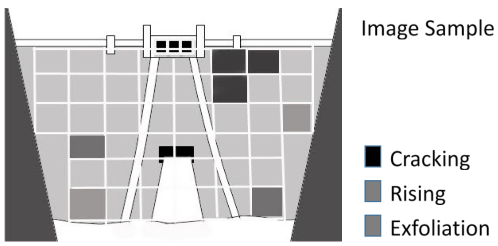

Figure 2.1 Engineering Smart Assistant

3 Memento Photo Collage

Alice has returned to Sydney after being a bridesmaid at her best friend's destination-wedding in Cusco, Peru. Alice had a great time at the wedding, and really enjoyed a self-guided tour of the city using a smart guide developed by a Cusco native. Alice added a few days to visit Machu Pichu, and added a few days in Lima to explore the culture, food, nightlife, and beaches. Alice is very taken with the traditional Peruvian music, and also really liked a couple popular songs on the radio, so she downloaded those for her music collection.

On the long flight back to Sydney, Alice went through her photos and created a collage of photos and used the newly acquired music as the collage soundtrack that would play when Alice played the collage. The application she used to create the collage also allowed Alice to link to the full resolution image version of each of the photo tiles in her collage. The entire memento photo collage, including high resolution photos and soundtrack, was bundled into a single JPEG file so that it could be easily shared to her electronic photo frame and share the memento with friends; everyone could at least see the photo collage, and those with newer photo frames got to hear the soundtrack and enjoy the full-screen versions of the photos.


Figure 3.1 Momento Photo Collage

4 Navigating in Omnidirectional Images

Heidi captures a 360 picture of her grandfather's house and garden. Later, when she reviews and edits the image, Heidi selects a region of interest (ROI) within the omnidirectional photo which can be viewed using a legacy JPEG viewer. Heidi also adds the default position the viewer should see when opening the 360 image. She sends this composite image to her friend Klara as an email attachmnent.

Klara receives Heidi's email. Klara is able to view Heidi's the ROI portion of photograph with a conventional JPEG viewer on her PC. However, her PC's photo viewer application supports the JPEG 360 photos, so when she uses that application, Klara can choose to view either the ROI image or the 360 image. Klara's phone also has a photo viewer application that supports JPEG 360 photos. Klara looks at Heidi's photo with using her smartphone., and then decides to switch to her PC to use her head mounted display (HMD) to have a more immersive and enjoyable experience which brings back memories of her life in the Alps. Klara decided that in the future, she wants to open the 360 image to a different default view, and she edits this property of the image. If she wants, Klara could change the ROI image by having the 360 photo editor replace it with her preference.

Heidi emails the URL of the omnidirectional photo to her childhood friend Peter, so Peter can access it with a 360 web-viewer using a browser on his computer. He can interact with the picture naturally without any noticeable delays because the compression technology optimizes the transmission efficiency.

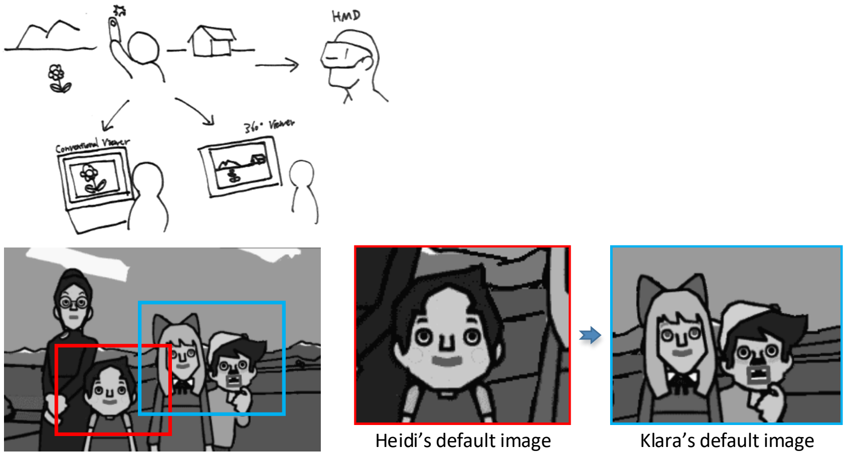

Figure 4.1 Navigating in Omnidirectional Images

5 Adding additional media data to omnidirectional images

Mr. Urashima Tarō captures several pictures in a domed garden in the Ryugu castle with his 360 camera and also with a high resolution DSLR camera. He captures images of the same scene in the garden with different resolutions, and he also captures videos of a waterfall and a river which include audio. After collecting several images and videos, he embeds this image and video data into a 360 JPEG file, and he makes an omnidirectional image with multiple media data.

Tarō send his omnidirectional photograph data to Ms. Oto-hime. Oto-hime checks it with her 360 viewer (e.g. HMD, handheld device, etc.). She looks in all the directions of the omnidirectional image. She enlarges a particular area of his omnidirectional photo image, and she can see some zoom-upped area in a higher resolution where high resolution photo data exist. When she picks several image areas such as waterfall area of his photo, she watches a moving picture of that part and listen to the sound of it where the video and audio data exist.

Tarō stores several enhanced omnidirectional images in the cloud. It goes fast, because the files are optimized to reduce the amount of data transfer. Momotarō, Tarō's old friend, can review the photos without waiting for downloading of large amounts of data.

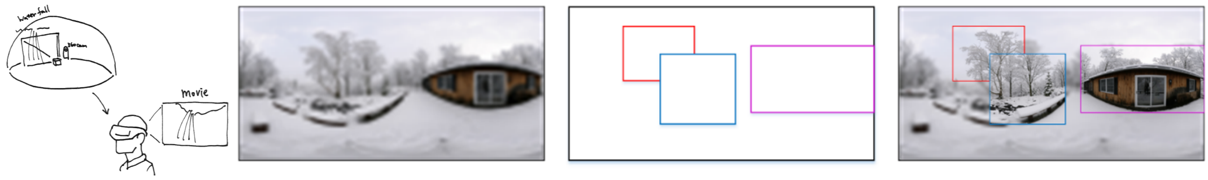

Figure 5.1 Adding additional media data to omnidirectional images

6 Time-Lapse Omnidirectional Animation

Ms. Rottenmeier captures a 360 picture of an aurora scene every five minutes. Then she makes an animated omnidirectional image of aurora. She also captures a 360 picture of a Japanese garden every day during the year, which allows her to compose an animated 365 days time-lapse image of the Japanese garden. After completing her collection of several omnidirectional scenes, she puts this data on her website.

Peter browses Ms. Rottenmeier's website by looking at the still photo images, and he decides to see her aurora's animated image data. After accessing the complete image data, he can play the animated aurora omnidirectional image with his HMD and 360 viewers on his PC and his smart phone. He enjoys Rottenmeier's omnidirectional animation of aurora.

Figure 6.1 Time-Lapse Omnidirectional Animation

7 Computer-Savvy Tour Guide

Francesca earns a living by guiding walking tours through Padova; she's a native to Padova, and has been fascinated by its sights and history since she was a little girl. She offers an application that can be used by tourists for self-directed tours, and another application that can add labels and annotations to images captured by tourists after-the-fact. To create her catalogs, Francesca walked around with a 360 camera to the highlights she wants to capture, and then captures an image that includes the location data from her trekking GPS which is more reliable and accurate than the GPS in her phone.

When Francesca gets back home, she uploads her images onto her personal computer, and organizes them into catalogs for walks of different length of time, topics of interest (history, architecture, food & wine, restaurants, music, etc.); each image is tagged to at least one catalog, and often to several catalogs. Then Francesca opens each image and adds explanatory comments and image labels to call out the interesting facts about each place. The application then derives a signature for each of the tagged visual elements in the images, and using the GPS data, calculates positional data from which the tagged visual element was observed. Then Francesca adds links to her images, and links to additional online information and images for further details. Finally, the application extracts the positional data, image labels, image element signatures, and the Francesca's comments, and the additional links into a metadata set for each catalog that she offers, and copyright and authentication metadata is added. When a catalog is purchased, a customer-specific encryption is applied to the catalog; the key is provided separately.

8 Erudite Tourist

Gunther feels that wandering the city, while enjoyable, would be more enjoyable with the advice of guide, but Gunther dislikes large guided tours, and would rather spend his budget on a nice meal than getting personal tour guide. Lucky for Gunther, he noticed Francesca's self-guided tour, and he bought two of her catalogs: i) the two hour historical tour of central Padova, and ii) her latest guide to street food in Padova. Both catalogs contain very similar information, including visual signatures for the food served by the street vendors. Francesca's metadata catalogs are uploaded to his phone and his personal computer.

Gunther borrows a 360 degrees panorama camera from his friend François to use on his vacation in Italy. He straps it to his back and starts wandering Padova. The camera captures his walk by taking an image every 30 seconds, or when Gunther presses a button on a remote control in his hand in case something catches his interest. The camera also collects the time and position from the mobile phone Gunther is carrying, as well as the 6DoF and compass data from the camera sensor.

As Gunther walks, a mapping application shows him Francesca's tour path based on the GPS data from Francesca's file. He can stop anytime, take a side street, skip a sight, or stop for coffee, and then continue at his own pace.
When Gunter finishes the walking tour, he uploads the images to his personal computer. Using the sensor metadata captured as we walked around, and combining with Francesca's metadata in her catalog, the appropriate metadata is carried over from Francesca's catalog into Gunther's images, though some of it must be recomputed due to positional differences of the observations visual high lights. This merge/update of metadata become of Gunther's images. It might also include Francesca's copyright for her commentary. Gunther can decide whether or not to keep Francesca's link to additional on-line information about the sights – probably will for now since he may want to learn more after he returns home from his vacation.

9 Multiple Resolutions Omnidirectional Images

Mr. Schmitz captures several 360 pictures of the scene in different resolution with his omnidirectional camera. The camera then fuses these images into an omnidirectional image file supporting multiple resolution omnidirectional data. Mr. Schmitz zooms in and out the omnidirectional image he has just taken, and he can see that his omnidirectional picture has functionality of zooming in any direction. He stores it on his web site.

Mr. Momotarō is having a look at Mr. Schmitz's photos available online with an HMD viewer. Adolf's web site offers low resolution photo image quality when browsing through his collection of omnidirectional photographs. Momotarō picks several photos and decides to access the full data of these photos. He can also display high resolution photos with his 360 degree viewing devices.

10 Layered 360 Images

Miss Orihime tooks many 360° pictures at almost the same time in many locations, and then she puts these all 360° pictures to as single format. Mr. Kintarō take a look to Orihime's photographic image with his smartphone, and he find one house in her photo. He wanted to know what is behind the house, he use the transparent mode and select the house, and then the display application erased the selected house from the scene. After the house disappears, he can see a classic car behind it.

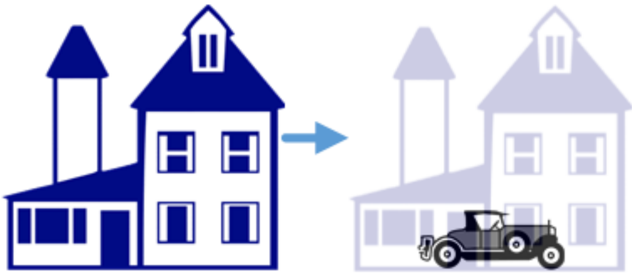

Figure 10.1 Layered 360 Images

11 Linked 360 Images

Mr. Kintarō checks the 360° image rooms on some real estate's website for renting a house in the Shibuya area. He selects one picture and captures a look at the room with his HMD. He finds a door in the room. He clicks on the door, and his point of view moves to the next room which is a kitchen. He looks in several directions in the kitchen and he finds that there is a cabinet. After he clicks on the cabinet door, the cabinet door opens and he can check the cabinet inside.

Figure 11.1 Linked 360 Images

12 360 Image Stabilization

Mr. Urashima Tarō creates a series of continuous scenes by continuously shooting 360° pictures of the Ryugu castle for several minutes in order to keep some memories. When Mr. Momotarō watches the pictures of the Ryugu castle taken by Tarō, he feels uncomfortable because the castle is shaking up and down. Mr. Tarō shot continuously without using a tripod, and so the Ryugu castle shook up and down. Fortunately, Tarō's photographs of the Ryugu castle has data of gravitational directions. Momotarō can correct the shake of the Ryugu castle easily by using that gravity data, so he won't experience any dizziness when watching Urashima Tarō's picture series.

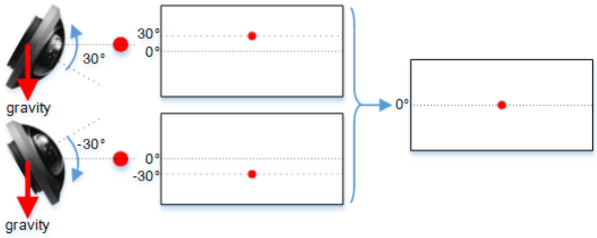

Figure 12.1 360 Image Stabilization

13 Efficient and Low Latency Viewport Transmission

Mr. Lee Min-Hyeong, a 360° image/video photographer, captures panoramic photos at the final race of women's 10 km speed skating relay at the Pyeongchang Winter Olympics with 8K (7680x4320) or higher resolution. After taking several picture sequences at the Winter Olympic Games, he stores this image data on his company's server, which delivers the data to their customers.

Ms. Jeong Yoo-Jin, one of the customers, takes a look at his sequence of speed skating relay pictures with her HMD. The display system has been created assuming a viewing angle of 90 degrees. The viewport has 2K resolution (1920x1080) and it is transmitted in high quality from the server side to HMD, while the data around the viewport is blurred. It becomes possible then to receive clear image at high speed. This technology allows Mr. Jeong Yoo-jin to view comfortly the Olympic Winter Games photos and videos.

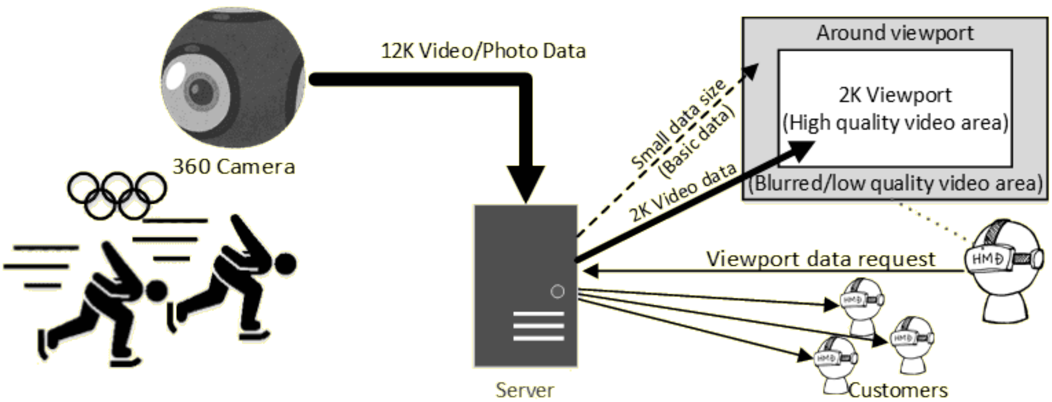

Figure 13.1 Efficient and Low Latency Viewport Transmission

14 Guided Viewport Experience

Mr. Kitarō tooks several spherical photographs, and he decides to add information to the photograph about displaying it in his favorite way. With this information other people can see his photograph in his recommended viewing way. He also puts three thumbnails of his most favorite scenes to the captured spherical image for instant reviewing with web browser, so everyone can immediately access his favorite scenes.

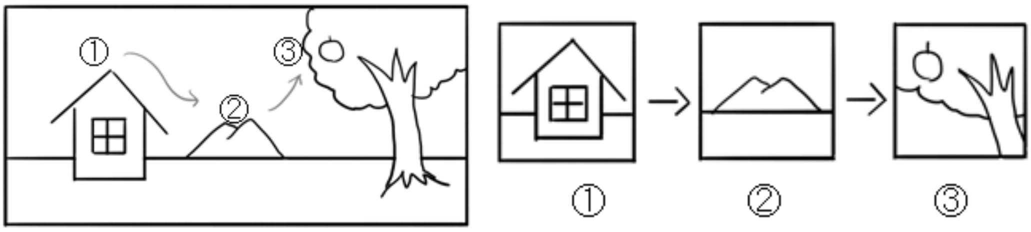

Figure 14.1 Guided Viewport Experience

15 Virtual Museum Tour with Explanatory Materials

Miss Neko-Musume goes to the web site of the National Museum to see modern art from her convenient location. Using her HMD, she goes through and see several pieces of art in the rooms of the museum via real/virtual corridor which connects the rooms. One picture catches her eyes. She wants to know what this painting art is, and she decides to display a detailed commentary of the picture. Then an explanatory note appears on the wall next to the picture, and she is able to know more about the painting art.

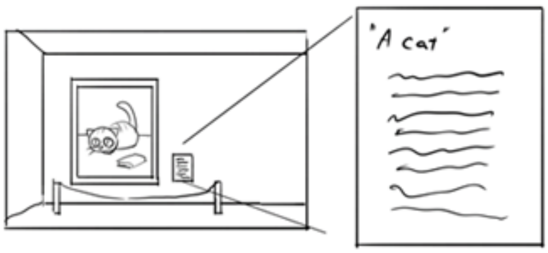

Figure 15.1 Virtual Museum Tour with Explanatory Materials

16 Camera Position and Orientation

Mr. Kitarō captures several photographs using his 360° camera which stores its shooting position, direction and time. With this data he can make a relative map easily. And he is also able to see which photograph is taken at which position and when.

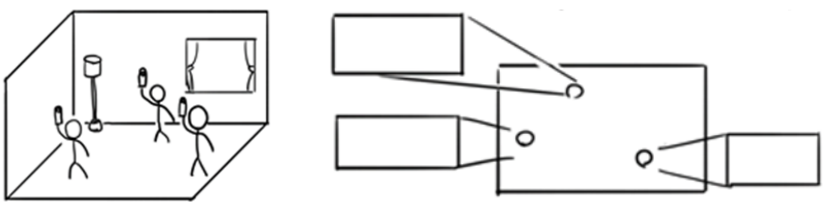

Figure 16.1 Camera Position and Orientation

17 Selective Filtering Indication

Harry buys a smartphone with a 360 camera. In order to draw visual attention to a portion of the 360 image, Harry can define a masking region of interest ROI when editing his 360 photos such that where the mask is defined to overlay the captured image, a visual blurring effect is applied. This allows Harry to send the image + mask to his girlfriend Sarah to put the attention to the interesting part. If, however, Sarah wants to see what was in the background without the blurring effect, she can disable the effect and see the original captured image.

18 Stereoscopic 360 Image

Frank captures an image with a stereo camera system during his vacation. He put the image on facebook, so that his friends can check his visited location. His friend Harry wants now to have a more immersive impression of the location, so he puts on his HMD to watch the image in stereo.

19 Multiple Type Sensor Images

Miss Orihime has a smart phone with dual type image sensors which can capture coloured image and infrared image. These image sensors are arranged in parallel so that the same object can be taken at the same time. Her smartphones can make wonderful photos from these images, however she saved these sensed image as one file format, and she apply other algorithms and sent it to the web service.

20 In-image Face Tagging for Photos

The photographer Jean Stringer takes a 360 degree photo for the photo department of the news agency AllNewsNow at a meeting of the heads of the 28 EU countries in Brussels. The 26 persons present at the time of the shot are seated around a table of elliptical shape and the 360 degree camera is placed in the middle of it. (Note: the illustration image does not meet this position, it is placed at the edge of the table.)

The photo editor of AllNewsNow has to add the names of the heads-of-state as metadata. As the typical viewer software of 360 degree photos opens only a limited viewport to the spherical view the photo editor wants to associate the name of a person with her or his face individually. First he draws a virtual rectangular numbered frame around the face of each head-of-state looking towards the camera in the 360 degree photo.

Figure 20.1 Captured Photo with Face Position Indications

Then he adds the name to the Person Shown metadata field of each numbered frame.

AllNewsNow Photo/Video Metadata Editor's Desk

Frame	Metadata of that frame
#1	Person shown: Ms Aname
#2	Person shown: Mr Cname
#3	Person shown: Mr Hname
#4	Person shown: Mr Kname
#5	Person shown: Mr Lname
#6	Person shown: Mr Mname
#6	Person shown: Mr Tname
#8	Person shown: Mr Uname
#9	Person shown: Mr Zname
#10	

Figure 20.2 Association of Person's Name to Frame Number

The information of each frame is transformed into an individual set of XMP metadata and all sets are embedded into the JPEG photo file.

A client of AllNewsNow launches a 360-degree-photo-viewer software to have a look at this JPEG photo. While moving the viewport across the sphere she can click on the M key or touch the screen of the tablet or smartphone and the viewer software shows the names next to the faces which are currently shown:

View in one direction …

Figure 20.3 FOV in One Direction

… view in another direction:

Figure 20.4 FOV in A Second Direction

… view in another direction:

Figure 20.5 FOV in A Thrid Direction

This definitely helps to distinguish which faces out of the many heads-of-state around the table are currently shown in the viewport and this improves the acceptance by consumers, and this improves the business value of 360-degree photos with metadata for different regions inside it.

AllNewsNow highly appreciates a very similar way of editing metadata for frames inside a 360-degree video.

IPTC considers it as key technical requirement for metadata associated with virtual frames to have a standardized reference point or reference coordinates system for a 360-degree spherical view. This is the origin for the definition of other points in this sphere relative to it. These points can be used to define such virtual rectangles.

Requirements

• To specify the relative location of tagged image elements a standardized origin is needed.