White Paper

of

CIPA DC-x006- 2008

Stereo Still Image Format for Digital Cameras

This translation has been made based on the original white paper. In the event of any doubts arising as the contents, the original document is to be the final authority.

Published by:
Camera & Imaging Products Association

1. Introduction

⋆ Background of the standardization

Stereo photography, which records stereo-viewable subject image by taking two complementary photographs together, leveraging the fact that human vision is getting a stereoscopic perception by the parallax between the left and right eyes, was born at the same time of the birth of photographic technology. The enduring popularity of stereo photography entertainment events indicates a very high level of untapped demand for systems that enable stereoscopic viewing of recorded images. However, apart from several boom cycles, stereo-photography has never really taken hold in the general consumer photography market. This has been attributed to a range of factors, but probably the most important factor is technical difficulty and equipment complexity involved in camera taking, printing and appreciation on the display, compared to general photographic technology.

Meanwhile the recent advent of certain models of mobile phones and computers equipped with 3D (i.e. stereoscopic viewable) displays suggests that 3D representation applicable equipments may well be embraced by consumer markets.

While binocular type 3D photographic images (called stereo images in this standard) have been used by certain users, the applications of stereo images seems to be spreading into diverse range now, because of the combination of recent rapidly growing popularization of digital still cameras and the advent of various displays mentioned above.

In order to make stereo images more appealing and accessible, it will be necessary to have the environment for easy and effective creating and viewing stereo images while also enhancing the viewing experience. In particular, it is important to construct the infrastructure for bridging the gap between the creation of stereo images on digital still cameras (which are enjoying spectacular growth at present) and the tools used to view them (such as display devices and printing systems).

However standard stereo image format has not been established. If this situation is allowed, manufacturers would simply develop a number of different own formats. This would not only hinder the real popularization but also create unnecessary confusion with existing services such as photography (printing service) and media service possibly.

Considering the above-mentioned situation, this document that is the stereo still image format for digital still cameras designed to provide a standardized format for stereo images and also to promote effective utilization of stereo image data, has been established. Specifically, it sets out regulations for the recording of stereo images into image files together with associated data (Stim tags) used to coordinate the camera with the viewing device.

⋆ Objectives of this standard

- The target is a format compatible with the existing popular binocular stereo image typically captured by a digital still camera fitted with a stereo imaging optical adaptor, as shown in the Figure on the next page. It could also potentially apply to data created in the same format with authoring tools or equivalent.
- The aim is to create a stereo image format that is independent of the camera type or model used to record the image.

- A framework for ensuring consistent display quality on a variety of different 3-D display devices will be provided.
- Attention should be paid to have no impact on the usage of incompatible devices.

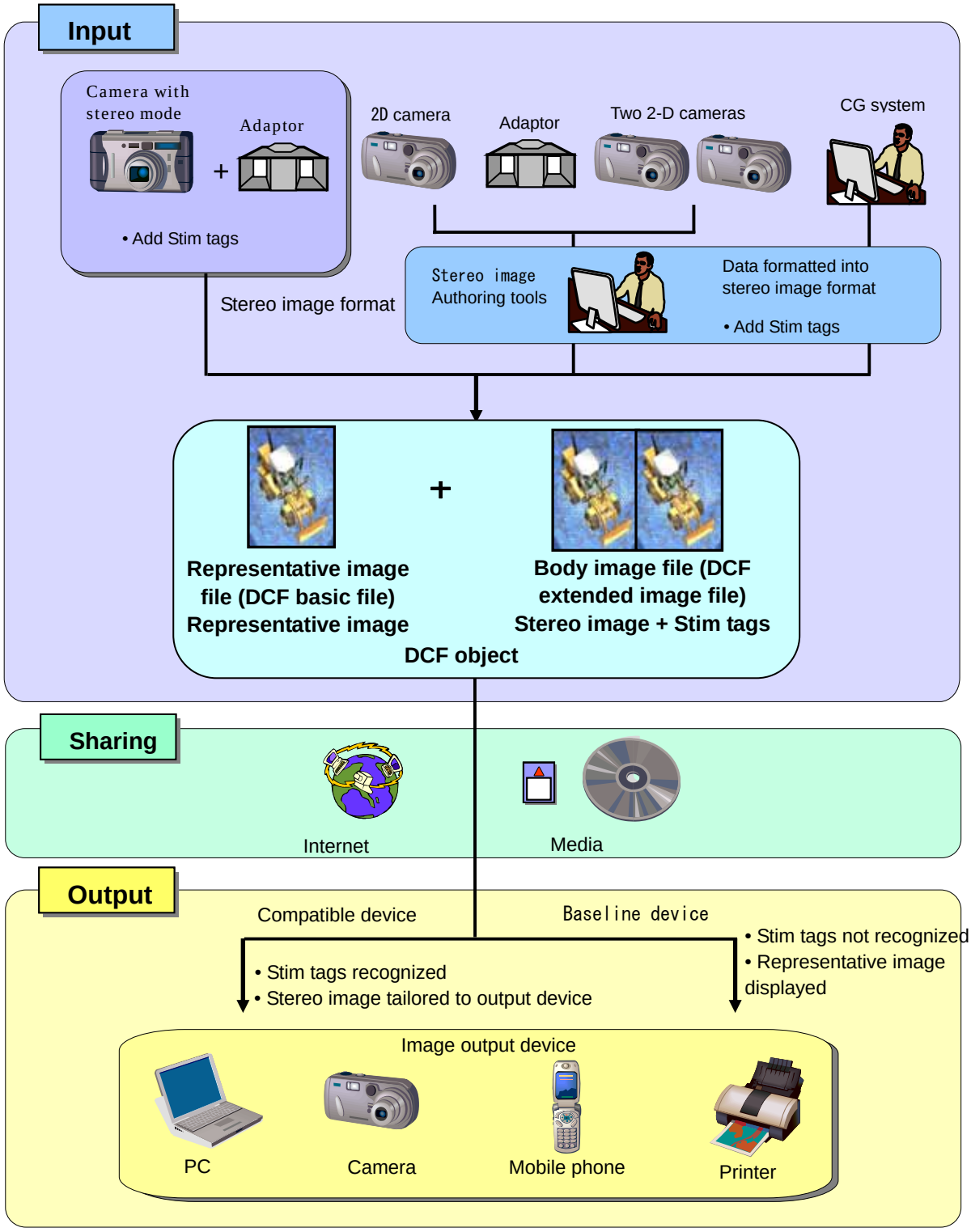

Figure: Typical stereo image applications,

2. Outline

Basic concept of this standard is that an image data themselves are recorded as existing 2D (plane) image, and at the same time the attached information, which is necessarily (otherwise useful) to view the image stereoscopically (i.e. to appreciate as a stereoscopic vision), is combined with the image into one file as some tag data. This enables the viewer to get proper and effective representation quality at the playback (or appreciation).

As above mentioned existing 2D image, Jpeg compression is adopted, and the most common location in binocular type stereoscopic image known as the side-by-side format (which locates each L or R viewpoint image on the left and right) is used.

As tag information, various items are defined for example;

- Which the alignment of the left and right images is, parallel or cross
- Image cropping area that would be adequate for stereoscopic viewing.
- Information about camera arrangement at taking picture.

This white paper only lists the names of the tags as the simplest introduction.

3. Scope

This standard specifies the formats to be used for images and metadata related to stereo image (Stim tags), in the case of recording stereo images as image files in digital still cameras and similar devices and systems.

* Note that stereo images in this standard are limited to the binocular (i.e. two viewpoints) type aligned viewpoint images.

4. File structure

4.1. Purpose

A body image file for stereo images in this standard employs a unique file extension , which may affect compatibility with image viewers that are not compliant with this standard. For this reason, the DCF camera file system rule is leveraged in order to prevent confusion and ensure efficient operation with ordinary image capture and viewing devices (such as digital cameras) by users.

Thus, the DCF object consists of the body image file and a representative image file containing the corresponding representative image (see below). This structure supports unified treatment at standard file operations such as copy, move and delete for the files. The representative image is substituted for viewing on DCF readers, while Exif tag information in the representative image file can also be accessed if required.

4.2. Body image file

The body image file is an image file having the unique file extension, defined as a DCF extended image file. It consists of the DCF object, with the DCF basic file (i.e., the representative image file) .

See the next section about the format of the body image file.

4.3. Representative image file

Recording a body image file shall involves recording the representative image (any one of the L and R viewpoint images, which is used for 2D playback/display) specified by the representative image tag as a DCF basic file, which consists of the DCF object, with the corresponding body image file.

The image size of the representative image may be any (i.e., resizable).

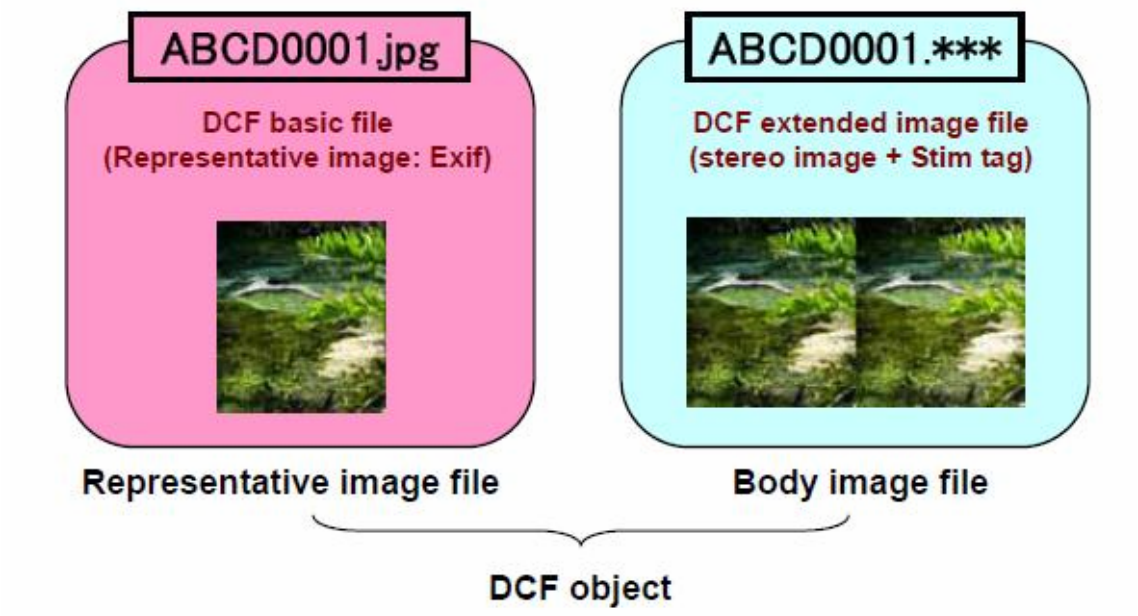

Figure: File structure

5. Body image file format

The body image file format for stereo images (Stim) specified in this standard conforms to the JPEG Baseline DCT format stipulated in ISO/IEC 10918-1. An application marker segment for the Stim tag (APP3) is also inserted.

The APP3 marker segment consists of the APP3 marker, the Stim identifier code, the Stim header and the Stim tag (APP3 body), as shown in Figure below.

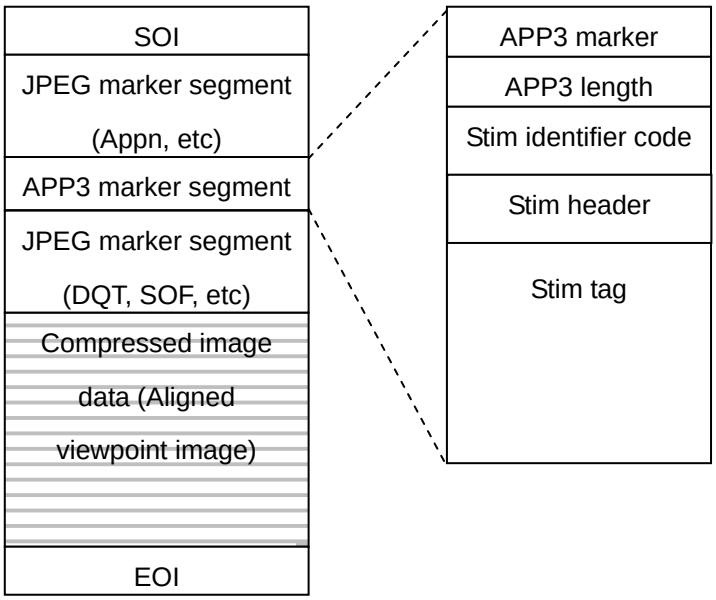

Figure: Structure of body image file incorporating Stim tag

6. Stereo image

In this standard, a stereo image refers to an aligned viewpoint image representing a side-by-side combination of two viewpoint images (monocular images corresponding the left-eye and right-eye viewpoints) in the same plane. The viewpoint images are rectangular in shape.

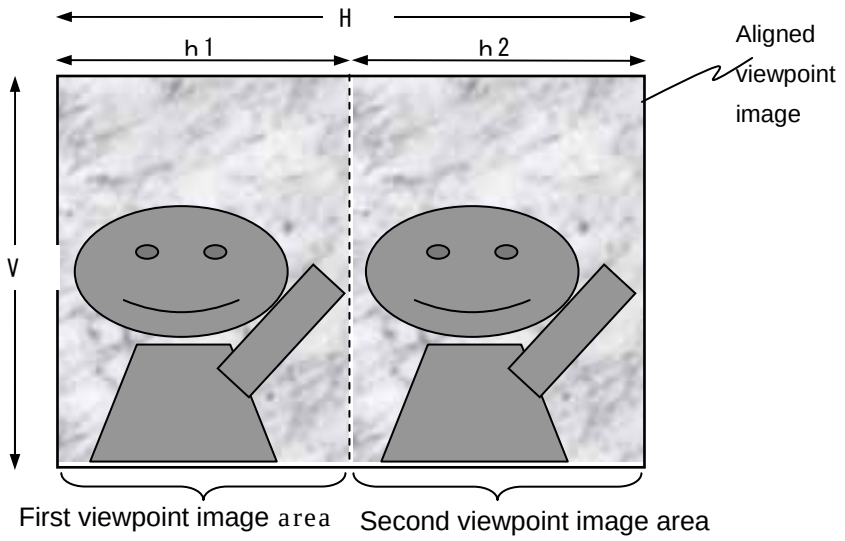

Figure: Structure of stereo image

In this standard, the term “aligned viewpoint image” is used to denote the overall image, while “first viewpoint image area” and “second viewpoint image area” denote the separate image areas. This convention, as illustrated in the diagram above, helps to delineate the different image areas in the recorded image. The aligned viewpoint image consists of the first viewpoint image area on the left-hand side together with the second viewpoint image area on the right-hand side.

7. Stim tag elements

When stereo image data is recorded, the Stim tags are stored in the prescribed locations within the recorded data. The following table lists the names of the elements defined as Stim tags. The Support level of each tag such as 'Mandatory' or 'Optional' is specified.

Table: Name of Stim tag elements

Tag name
StimVersion
ApplicationData
ImageArrangement
ImageRotation
ScalingFactor
CropSizeX
CropSizeY
CropOffsetX
CropOffsetY
ViewType
RepresentativeImage
ConvergenceBaseImage
AssumedDisplaySize
AssumedViewDistance
RepresentativeDisparityNear
RepresentativeDisparityFar
InitialDisplayEffect
ConvergenceDistance
CameraArrangementInterval
ShootingCount