Photoacoustic Data and Device Parameters

IPASC-DAM Consensus Document

Last changed: 22.12.2019
Contact persons: Janek Gröhl, Lina Hacker, Ben Cox

The following document aims to define metadata for photoacoustic (PA) devices to enable description, interoperability and data exchange of PA imaging data.

Table of Contents:

Definitions

- **Image:** An array of values varying in two or more spatial dimensions derived from analysis of an imaging signal and corresponding to an array of spatial locations in the imaged object (based on O'Connor 2017). In the context of this document, a PA image refers to the result of mapping raw time series data into the spatial domain.
- **Raw time series data:** A '*raw time series*' refers to the sampled signal from a detection element. '*Raw time series data*' refers to a set of such time series, one for each detection element.
- **Device:** A specific make and model of medical device (hardware and/or software) which is regulated in the USA by FDA Center for Devices and Radiological Health by 510(k) clearance24 and in the EU by CE marking (based on O'Connor 2017).
- **Modality:** A category of imaging device, characterised by a distinct physical principle (based on O'Connor 2017) (e.g. PA imaging).
- **Detection element:** A specific material element capable of converting mechanical energy to electrical energy and in some cases reciprocally converting electrical energy to mechanical energy (IEC TR 60854 1986). Here, it refers to e.g. a piezoelectric crystal or a laser-light interferometer.
- **Illumination element:** A specific material element capable of emitting light to illuminate the target. In this case, it usually refers to an optical fibre that conveys the light generated by a laser source.
- **Frame:** A set of raw time series data corresponding to a specific time point and wavelength.

Further definitions can be found in the IPASC 'Terms and Definitions' consensus document.

Attributes

Each metadata item in this document is characterized by a series of attributes with the purpose of describing and defining its use and boundary conditions. If applicable, further specifications by nested attributes are given.

- **Necessity**: 'Minimal' or 'Report if present' condition for the attribute. Minimal are all parameters that are necessary to reconstruct an image from the raw time series data. Any additional information should be reported in the metadata if by any means feasible.
- **dtype**: data type
- **Units**: SI units of the attribute if applicable
- **Description**: A short description of the attribute
- **Condition**: Constraints of the attribute, if applicable

- Nested attribute: A sub-attribute that further describes an attribute.
- Measurement device attribute: A specific type of a nested attribute that further describes measurement device details if required. Measurement device attributes are always optional. They include:
    - Measurement Device Type: A string literal describing the measurement device for this attribute, e.g. ‘pyroelectric sensor’ or ‘wavemeter’.
    - Measurement Device Manufacturer: A string literal describing the manufacturer of the measurement device, e.g. ‘Thorlabs’.
    - Measurement Device Serial Number: A string literal comprising the serial number of the measurement device.
    - Calibration Date: A timestamp referring to the date when the measurement device was last calibrated.

Minimal parameters

The minimal parameters refer to a set of parameters which is minimally needed to reconstruct an image from PA raw time series data. In specific, the minimal parameter set contains:

- Container Format Metadata: The *container format metadata* refer to the inherent features of the file format which specify mandatory parameters of the file format. They include the UUID (1), the type of compression (2) and the type of encoding (3).

- Binary Data Metadata: The *binary data metadata* refer to the metadata that makes the binary data machine-readable. They include specifications on data type (1), dimensionality (2) and the sizes (3) within each dimension.

- Sampling Rate: The *sampling rate* refers to the rate at which samples of the analog signal are taken to be converted into digital form.

- Acquisition Wavelengths: The *acquisition wavelengths* field is a 1D array that contains a list of all wavelengths used for the image acquisition.

- Detector Positions: The *detector position* defines the position of the detection element in 3D cartesian coordinates $[x_1, x_2, x_3]$.

Part 1 - Photoacoustic Imaging Device

The first part of this document defines all information necessary to describe a PA imaging device. It aims to build the foundation of a database storing information of commercial and home-built PA devices and enabling data acquisition modelling of respective systems. Each system will be assigned with a unique identifier which is referenced in the recorded PA data.
A PA imaging device consists of one or more illuminators which deliver light into the tissue as well as one or more detectors that are responsible for measuring incoming sound waves generated by the PA effect. All coordinate positions of the individual elements will be given in cartesian coordinates relative to the origin of the transducer. The field of view of the imaging device will be given in the same coordinate system.

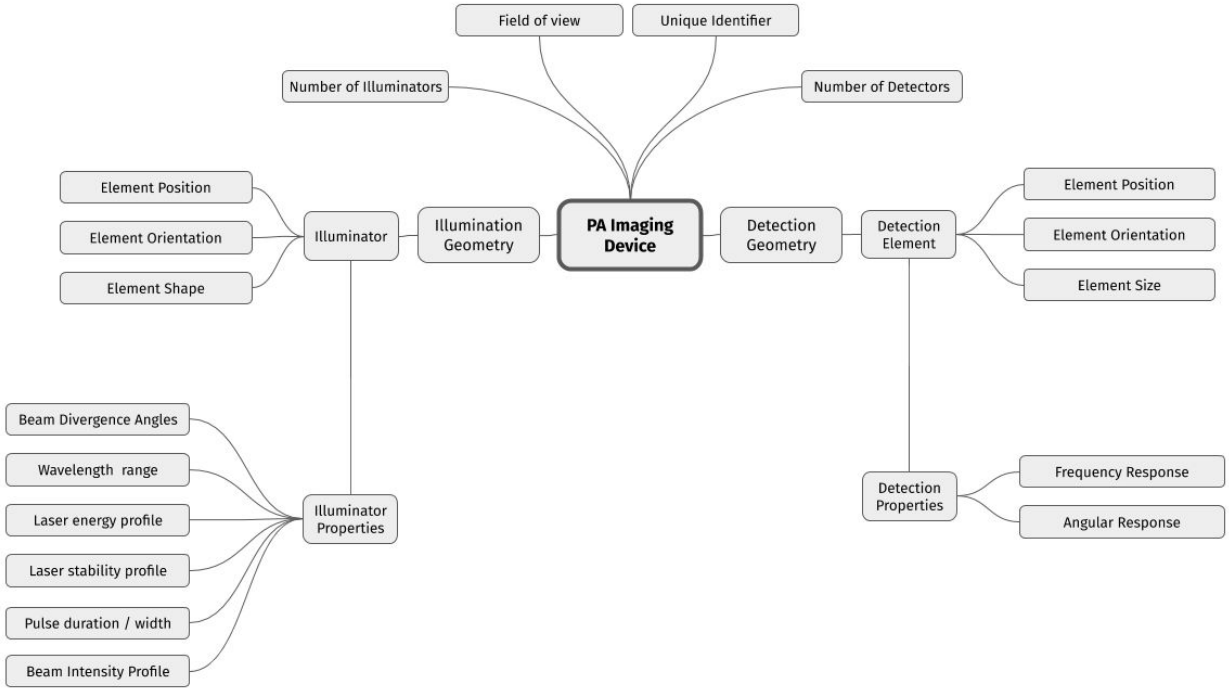

Figure 1: Graphical overview of the proposed main elements of PA Imaging devices and the specific information that these elements hold. The main elements are (1) the illumination geometry and (2) the detection geometry.

UUID (Universally Unique Identifier)

Necessity: Minimal
dtype: String

Description: The *UUID* is a universally unique identifier to the device description that can be referenced.
Condition: 128-bit Integer displayed as a hexadecimal string in 5 groups separated by hyphens, in the form 8-4-4-4-12 for a total of 36 characters. The UUID is randomly generated using the UUID Version 4 standard.

Field of View

Necessity: Report if present
dtype: Double Array
Units: Meter [m]
Description: The *field of view* defines an approximate cube of the area detectable by the PA imaging device in 3D cartesian coordinates [x_1, x_2, x_3]. The field of view always starts in the origin of the coordinate system and expands in the positive x_1, x_2, x_3 directions.

Number of Illuminators

Necessity: Report if present
dtype: Integer
Units: Dimensionless Unit
Description: The *number of illuminators* quantifies the number of illuminators that are used in the respective PA imaging device. Each of these illuminators is described by a set of illumination geometry parameters.

Number of Detectors

Necessity: Minimal
dtype: Integer
Units: Dimensionless Unit
Description: The *number of detectors* quantifies the number of transducer elements that are used in the respective PA imaging device. Each of these transducer elements is described by a set of detection geometry parameters.

Illumination element

Illuminator Position

Necessity: Report if present
dtype: Double Array
Units: Meter [m]
Description: The *illuminator position* defines the position of the illuminator in 3D cartesian coordinates [x_1, x_2, x_3] .

Illuminator Orientation

Necessity: Report if present

dtype: Double Array
Units: Radians [rad]
Description: The *illuminator orientation* defines the rotation of the illuminator in 3D cartesian coordinates [r_1, r_2, r_3] . It is the normal of the planar illuminator surface.

Illuminator Shape

Necessity: Report if present
dtype: Functional
Units: Meter [m]
Description: The *illuminator shape* defines the shape of the optical fibres, so it describes whether the illuminator is a point illuminator, or has a more continuous form. Illuminators can only have planar emitting surfaces.

Illumination Element Properties

Wavelength Range

Necessity: Report if present
dtype: Double Array
Units: Meters [m]
Description: The *wavelength range* quantifies the wavelength λ range that the illuminator is capable of generating by reporting three values: the minimum wavelength λ_{max} , the maximum wavelength λ_{max} and a metric for the accuracy $\lambda_{accuracy}$: $(\lambda_{min}, \lambda_{max}, \lambda_{accuracy})$. This parameter could for instance be (700, 900, 1.2), meaning that this illuminator can be tuned from 700 nm to 900 nm with a precision of 1.2 nm.

Laser Energy Profile

Necessity: Report if present
dtype: Double array
Units: Normalized units (to the maximum intensity)
Description: The *laser energy profile* field is a functional of wavelength (nm) that represents the laser energy of the illuminator with regard to the wavelength. Thereby, systematic differences in multispectral image acquisitions can be accounted for.
Condition: The *laser energy profile* functional is well defined and non negative in the *Wavelength Range*.

Laser Stability Profile

Necessity: Report if present
dtype: Double array
Units: Joule [J]

Description: The *laser noise profile* field is a functional of wavelength (nm) that represents the standard deviation of the pulse-to-pulse laser energy of the illuminator with regard to the wavelength.
Condition: The *laser noise profile* functional is well defined and non negative in the *wavelength range*.

Pulse Duration / Width

Necessity: Report if present
dtype: Double
Units: Seconds [s]
Description: The *pulse duration* or *pulse width* describes the total length of a laser pulse, measured as the time interval between the half-power points on the leading and trailing edges of the pulse.

Beam Intensity Profile

Necessity: Report if present
dtype: Functional
Units: Normalized units (to the maximum intensity)
Description: The *beam intensity profile* is a function of a spatial position that specifies the relative laser beam intensity according to the planar emitting surface of the *illuminator shape*.

Beam Divergence Angles

Necessity: Report if present
dtype: Double
Units: Radians [rad]
Description: The *beam divergence angles* represent the opening angles of the laser beam from the *illuminator shape* with respect to the orientation vector. This angle represented by the standard deviation of the beam divergence.

Detection Element

Element Position

Necessity: Minimal
dtype: Double Array
Units: Meter [m]
Description: The *element position* defines the position of the detection element in 3D cartesian coordinates [x_1, x_2, x_3] .

Element Orientation

Necessity: Report if present

dtype: Double Array
Units: Radians [rad]
Description: The *element orientation* defines the rotation of the detection element in 3D cartesian coordinates [r_1, r_2, r_3] in radians.

Element Size

Necessity: Report if present
dtype: Double Array
Units: Meter [m]
Description: The *element size* defines the size of the detection element in 3D cartesian coordinates [x_1, x_2, x_3] relative to its position and orientation.

Detection Element Properties

Frequency Response

Necessity: Report if present
dtype: Functional (of frequency in Hertz [s^{-1}])
Units: Normalized Units (to the maximum intensity)
Description: The *frequency response* is a functional that characterizes the response of the detection element to the frequency of the incident pressure waves.

Angular Response

Necessity: Report if present
dtype: Functional (of incident angle in radians)
Units: Normalized Units (to the maximum efficiency)
Description: The *angular response* field characterizes the angular sensitivity of the detection element to the incident angle (relative to the elements orientation) of the incoming pressure wave.

Part 2 - Photoacoustic Raw Time Series Data

PA raw time series data refer to the signals recorded by the detection elements of the PA system. On a data level (Figure 2), a PA raw time series data is a block of binary data. This binary data is characterised by three main components: (1) The metadata of the file container, including a *Universally Unique Identifier* (UUID) of the data and a UUID reference to the PA imaging device that was used, (2) the binary data metadata that makes the binary data machine-readable, and (3) the acquisition metadata that enables interpretability of the recorded data and gives context on the imaging settings.

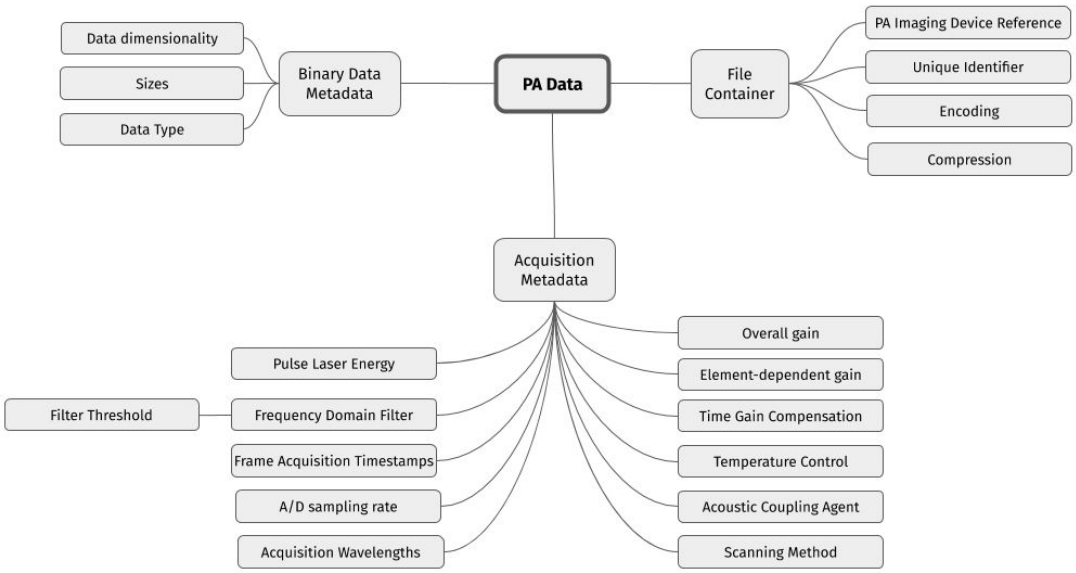

Figure 2: Graphical overview of the proposed four main elements of PA raw time series data and the specific information that these elements hold. The main elements are (1) file container format metadata, (2) acquisition metadata and (3) PA binary metadata.

File Container Format

The container format metadata refers to the inherent features of the file format which specify how the different elements of metadata are combined in a computer file.

Encoding

Necessity: Minimal
dtype: String
Description: The *encoding* field is representative of the character set that was used to encode the binary data and the metadata. E.g. one of ‘UTF-8’, ‘ASCII’, ‘CP-1252’, ...

Compression

Necessity: Minimal
dtype: String
Description: The *compression* field is representative of the compression method that was used to compress the binary data. E.g. one of ‘raw’, ‘gzip’, ...

UUID

Necessity: Minimal
dtype: String
Description: The *UUID* (Universally Unique Identifier) is a unique identifier to the data that can be referenced.
Condition: 128-bit Integer displayed as a hexadecimal string in 5 groups separated by hyphens, in the form 8-4-4-4-12 for a total of 36 characters. The UUID is randomly generated using the UUID Version 4 standard.

Photoacoustic Imaging Device

Necessity: Report if present
dtype: String
Description: A reference to the UUID of the PA imaging device description as defined in part 1.

Binary Data Metadata

Data Type

Necessity: Minimal
dtype: String
Description: The *data type* field represents the datatype of the binary data. This field is given in the C++ data type naming convention. E.g. ‘short’, ‘unsigned short’, ‘int’, ‘unsigned int’, ‘long’, ‘unsigned long’, ‘long long’, ‘float’, ‘double’, ‘long double’.

Dimensionality

Necessity: Minimal
dtype: String

Description: The *dimensionality* field represents the acquisition format of the binary data and specifies the number of spatiotemporal dimensions of the data that is comprised of one or more frames. E.g. '1D', '2D', '3D', '1D+t', 2D+t', '3D+t'. In this notion, the time series sampling of one transducer would count as a "spatial" dimension. These are defined as 1D = [$\boldsymbol{\tau}$], 2D = [x_1, $\boldsymbol{\tau}$], 3D = [x_1, $\boldsymbol{\tau}$, x_2]. The "+t" will then add a time dimension for multiple of these frames.

Sizes

Necessity: Minimal
dtype: Integer Array
Units: Dimensionless Quantity (the units can be inferred in combination with *Dimensionality* and the detection and illumination geometry).
Description: The *sizes* field quantifies the number of data points in each of the dimensions specified in the dimensionality field. e.g. [128, 2560, 26] with a "2D+t" dimensionality.

Acquisition Metadata

Pulse Laser Energy

Necessity: Report if present
dtype: Double Array
Units: Joule [J]
Description: The *pulse laser energy* field specifies the pulse-to-pulse laser energy that was measured for the acquisition of the raw time series data.
Condition: Array size must be the same as the size of 't' specified in the *sizes* field.

Frame Acquisition Timestamps

Necessity: Report if present
dtype: Double Array
Units: seconds [s]
Description: The *frame acquisition timestamps* field indicates the timestamp of the acquisition system.
Condition: Array size must be the same as the size of 't' specified in the *sizes* field. Timestamps are given in seconds with the elapsed time since epoch (Jan 1st 1970, 00:00).

Acquisition Optical Wavelengths

Necessity: Minimal
dtype: Double Array
Units: Meters [m]
Description: The *acquisition optical wavelengths* field is a 1D array that contains a list of all wavelengths used for the image acquisition.

Condition: The wavelengths must be listed in the order the images were taken in, starting with the first wavelength at the first frame. The number of recorded frames over time ('t' specified in the *sizes* field) must be divisible by the number of *Acquisition Wavelengths*.

Time Gain Compensation

Necessity: Report if present
dtype: Double Array
Units: Dimensionless Unit
Description: The *time gain compensation* field is a 1D array that contains the relative factors which have been used to modify the time series data to correct for the effect of attenuation.
Condition: The *time gain compensation* array has the same dimension as the time series dimension [$\boldsymbol{\tau}$].

Overall Gain

Necessity: Report if present
dtype: Double
Units: Dimensionless Unit
Description: The *overall gain* is a single value describing a factor that has been applied to all values of the raw time series data.

Element-dependent Gain

Necessity: Report if present
dtype: Double Array
Units: Dimensionless Unit
Description: The *element-dependent gain* field is a 2D array that contains the relative factors which have been used to perform apodization.
Condition: The *element-dependent gain* is a 2D array that has the same dimension as the time series dimension, with the spatial axis denoting the number of detection elements [x, $\boldsymbol{\tau}$].

Temperature Control

Necessity: Report if present
dtype: Double Array
Units: Kelvin [K]
Description: The *temperature control* field indicates the temperature during image acquisition.
Condition: The *temperature control* array either has the same dimension as the time dimension [t], or is a single value indicating a constant temperature over all imaged samples.

Acoustic Coupling Agent

Necessity: Report if present
dtype: String

Description: A string representation of the acoustic coupling agent that was used. For example, the following options are possible: D_2O, H_2O and US-gel.

Scanning Method

Necessity: Report if present
dtype: String
Description: A string representation of the *scanning method* that was used.

A/D (Analog/Digital) Sampling Rate

Necessity: Minimal
dtype: Double Array
Units: Hertz [Hz]
Description: The *A/D sampling rate* refers to the rate at which samples of the analog signal are taken to be converted into digital form.

Frequency Domain Filter

Necessity: Report if present
dtype: String
Description: The *frequency domain filter* refers to a signal processing step which aims to remove unwanted frequency components of the raw signal. The following options are possible: 'low pass', 'high pass', 'band pass' filter.
Nested Attributes:

Filter Threshold

Necessity: Report if present
dtype: Double Array
Units: Hertz [Hz]
Description: The *frequency threshold* levels that have been applied to filter the raw time series data.
Condition: This field is only mandatory if a filter type has been defined.

Version History

Version 0.3:

Document revised after first annual meeting. Addition of minimal vs. optional parameters. Consistent formatting.

Version 0.2:

Document revised mainly by dividing it into two main parts: The description of PA devices and the metadata that accompanies recorded data. Removed the ultrasound data in the current version. The first release should deal with PA raw time series data primarily.

Version 0.1:

Document revised after extensive feedback from the IPASC DAM theme members. Added ‘Necessity’ and ‘Nested Attributes’.

Version 0:

Document initially drafted by Janek Gröhl and Lina Hacker.