

## An MPEG-7 Extension for Describing Visual Impairments

Werner Bailer, Peter Schallauer

JOANNEUM RESEARCH, Institute of Information Systems & Information Management  
{werner.bailer, peter.schallauer}@joanneum.at

### Abstract

*Analysing the condition of audiovisual essence is an important step in audiovisual production and preservation. Standardised impairment description of audiovisual media is a pre-requisite for system interoperability between content digitisation, documentation, management, restoration, production and delivery systems.*

*We analyse existing capabilities for describing impairment in audiovisual metadata standards. Because of its unique detailed spatiotemporally structured description capabilities we have selected MPEG-7 as the basis for visual impairment description. Following the approach for audio quality description, we define a general description scheme for visual impairments, which allows representing defect events and statistical quality measures. For certain defects, more specialised descriptors are proposed. In addition, we define a comprehensive classification scheme for visual impairments.*

### 1. Introduction

Quality assurance of audiovisual content is a crucial task in several steps of the media production, delivery and archiving processes. Broadcasters are checking audio and video quality during the ingest process, after editing, after encoding and before distribution, such as play-out for terrestrial, satellite and cable broadcast or for delivery to internet and video-on-demand services. Archives are checking for content integrity and content impairments at archive ingest and delivery. Content providers are checking content for correct encoding and conformance to the required audiovisual quality before dispatching to the broadcasters or other service providers.

In all these tasks a wide range of impairments are checked, e.g. noise level, blocking, luma and chroma violations, dropouts, freeze frames, image instability and flickering. Standardised impairment description of

audiovisual media is a pre-requisite for system interoperability between content digitisation, documentation, management, restoration, production and delivery systems.

In the following section, we review the capabilities of existing multimedia content description standards for describing visual impairments. We then discuss the requirements for impairment description. Based on this analysis, MPEG-7 extensions for impairment description as well as a corresponding classification scheme are proposed.

### 2. State of the art

While most multimedia metadata standards allow the description of technical signal parameters of the audio and visual content, capabilities for describing impairments are rare.

The SMPTE Metadata Dictionary [1] contains elements for describing overall assessments of the technical quality, some video test parameters, the audio and visual signal to noise ratio and quality events and parameters of audio data as defined by the Broadcast Wave Format (BWF [2]).

In MPEG-7 already some impairment descriptors and description schemes have been standardised. The *MediaQuality* descriptor [4] is an element in the *MediaProfileDS* and thus applicable to all media types. It contains (i) quality rating, expressed as a floating point value, (ii) a rating source and a reference to the rating information and (iii) a list of perceptible defects, discriminated into visual and audio defects. Each defect is a reference to a term in a classification scheme. It is not possible to describe the defect in more detail or its exact (spatio-) temporal location. The *AudioSignalQualityDS* has been introduced in amendment 1 to part 4 [3]. It can be added to each audio segment and contains the following elements: balance, noise level, DC offset, cross channel correlation, delay, a list of error events. Each of these error events is described by the error class (a reference to a term in a classification scheme), time stamp and

channel number, detection method (manual, automatic), relevance, status and optional text annotation.

Classification schemes are MPEG-7 description schemes for defining hierarchies of controlled vocabulary. The following classification schemes exist for the description of impairment information: *AudioDefectsCS* and *VisualDefectsCS* [4] and *ErrorClassCS* [3].

### 3. Requirements

In order to define description tools for the condition of audiovisual essence, i.e. its technical quality and the defects it may have, we review the requirements to such a description in this section. The impairment description shall allow getting an overview of the condition of the audiovisual material. It shall thus be a compact description and contain details only if absolutely necessary.

The description shall not include intermediate results of specific restoration algorithms, configurations of analysis or restoration and a history of applied restoration steps.

The description is mainly produced by automatic tools, and it shall also be possible to process the description automatically. Therefore,

- the time point or range for which a description is valid must be specified,
- quality has to be quantified numerically or by sets of defined terms,
- defects need to be unambiguously identifiable, and
- optionally properties of defects may be further described numerically or by sets of defined terms.

Additionally it must be possible to annotate and extend the descriptions manually. Preferable, the descriptions are stored in a human readable format.

As the descriptions shall support a user in getting a quick overview of the condition of the material, they shall be defined in a way that they are easy to visualize. Especially quality measures and defect descriptors that represent a larger time range shall allow condensed visualization over time.

Quantitative descriptions of impairments shall correspond to the perceived severity of the defect.

Impairment descriptions shall be efficiently searchable. Whenever possible, the descriptors shall be defined in a way that they are easy to index.

### 4. Description of visual impairments

Based on an analysis of the state of the art and the requirements defined above it becomes clear that

MPEG-7 is a suitable standard to serve as a basis for the description of visual impairments. MPEG-7 allows to structure descriptions on different levels of granularity and already offers some tools for quality description, especially in the audio domain. We thus propose the following steps for extending MPEG-7:

- Define a description scheme for visual impairments, similar to that for audio quality and defects defined in [3].
- In addition, allow the extension by detailed descriptors for specific quality measures and defect descriptors.
- Define these specific descriptors for some common quality measures and defects.
- Extend the existing MPEG-7 controlled vocabularies (classification schemes) for both visual and audio impairments.

The extension is based on the MPEG-7 Detailed Audiovisual Profile [7] which has been proposed for detailed description of audiovisual content in production and archiving.

#### 4.1 General approach

The description scheme for visual impairments describes a temporal segment of visual material and contains the following elements.

A **list of defects** with at least an identification of each defect (reference to a controlled defect vocabulary) or a specific defect descriptor that describes the occurrence of a defect in more detail. Usually a defect occurs at a time point or in a small time range (e.g. damaged frames). Depending on the defect, the spatial location may also be given. Examples: dropout (and the region of the image affected), line scratch (and its vertical position over time), etc.

A **list of quality measures**, each described by a specific quality measure descriptor that contains defect statistics of a temporal segment of the audiovisual material. Examples are the noise level, the amount of blocking, the graininess, the amount of dust, the amount and average frequency of flicker etc.

It is possible to define quality measures for all impairments by describing statistics of their occurrence in the material. Defect descriptors cannot be used for those properties, which cannot be expressed as an event with a specific spatial/temporal position (e.g. noise).

Similar to *AudioSignalQualityDS* [3] we propose a description scheme derived from *VisualDSType* which can be attached to any video segment. It contains:

- A list of impairments (quality measures and defects) for the segment. For each impairment, the class of impairment (reference to term in classification scheme), its strength (normalized numerical value), the relevance, the detection process (manual, automatic), a comment and its origin (a device or system from which the impairment originates can be specified).
- The tool used for generating description.
- Identification of the operator.
- Further textual annotation.

## 4.2 Specific defect and quality measure descriptors

As discussed above, each of the impairments in the description of a segment can be described in more detail by using a specific descriptor. The specified descriptors for visual impairments are presented in this section. Following the approach for the specification of MPEG-7 visual description, we propose only the description format, but not the extraction method. Analysis algorithms for some of the impairments below are discussed in [5].

**Dropout/partial frame damage.** Describes a dropout or other damage affecting a part of a single or a few frames. The lines or region and the channels affected can be described.

**Line scratches.** Describes vertical scratches on film material occurring in arbitrary temporal segments from a few to hundreds of frames and their properties (horizontal position, width, negative/positive).

**Number of line scratches.** Describes the number of scratches in a shot.

**Dust/dirt level.** Describes the level of dust or dirt spots level per shot or part of a shot. The average number and size of spots and the average intensity can be described.

**Noise/grain level.** Describes the noise or film grain level of a shot, using the properties PSNR, spatial frequency and brightness dependency.

**Flicker level.** The level of temporal intensity variations (flicker) in a shot, described by the average flicker intensity, its frequency distribution and the local variation.

**Image instability.** Describes the geometric position instability of the image by the average/maximum horizontal/vertical displacement.

**Blur.** Describes the spatial resolution of image in a shot by characteristic scale of the edges.

**Blocking level.** Describes amount of blocking artefacts level from lossy DCT based encoding per shot or part of a shot.

**Dropout level.** Describes the number and area of dropouts per shot or part of a shot.

**Channel misalignment.** Describes asynchronous colour channels in a some shots or the whole material.

**Colour range defects.** Describes high-/low contrast in one or more channels or saturation/clipping in a sequence of shots or the whole material by specifying the fraction of the intensity range used and/or the area affected by saturation.

## 4.3 Classification schemes

The classification schemes which are currently defined in MPEG-7 only specify a few defects and therefore need to be extended to cover all the impairments that occur in archive and restoration applications. Classification schemes can be easily extended without the need to adapt or extend the XML schema defining the standard.

It is crucial to define comprehensive impairment classification schemes, as they will serve as the basis for defect and quality measure description in MPEG-7. They can be used right away, even without any extension of the standard just by using the already existing description tools.

The general description scheme described above relies entirely on classification schemes for identifying defects. Applications that are not capable of understanding the more specialised visual impairment descriptors are still be able to get a rough description of the defect just from the temporal segment and the reference to a defect in the classification scheme.

The public BRAVA broadcast archive programme impairments dictionary [8] is very much oriented towards completeness and can thus serve as a good starting point for extending the audio and visual defect classification schemes of the MPEG-7 standard. The BRAVA dictionary is thus used as the basis of the PrestoSpace impairment classification scheme. The definition of a classification scheme from flat list of impairments in the BRAVA dictionary introduces a hierarchical structure of the defects, e.g. a dropout defect and several specialized kinds of dropouts. When available, the classification scheme contains multilingual defect names and textual descriptions (English and French).

The main organization criteria of the classification scheme are the visible and audible effects of defects. If the cause or origin of the defect is known, it may be annotated using the origin element of the description scheme or by selecting a narrower sub-term that specifies the defect in relation to a certain cause.

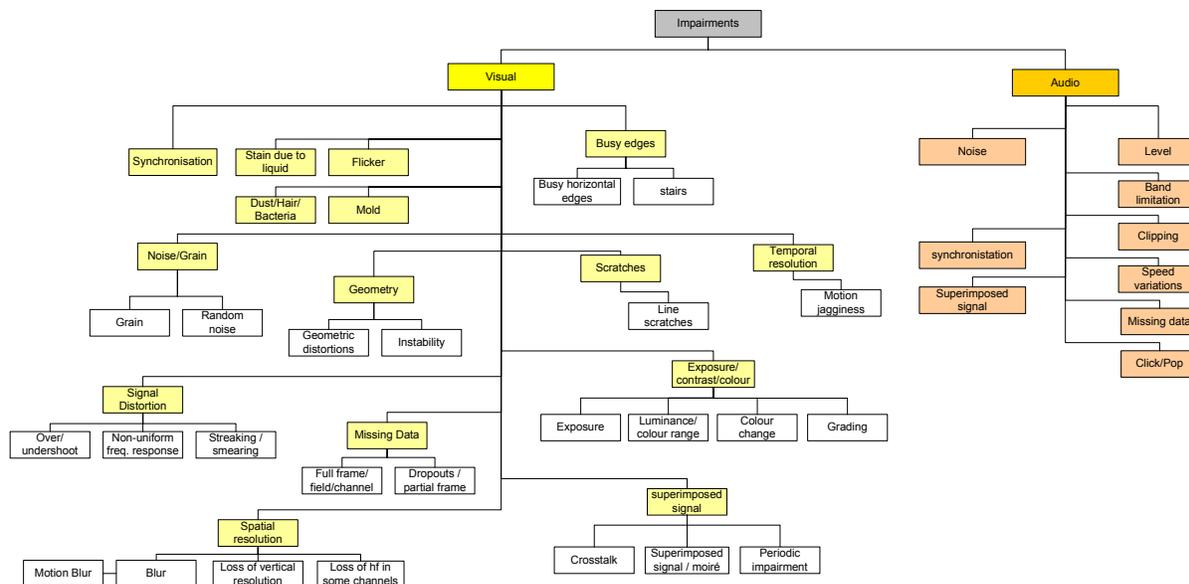


Figure 1: Top level elements of the impairment classification scheme.

## 5. Conclusion

The analysis of video quality is of growing importance in audiovisual media production and archiving. Defect and quality analysis tools are able to gather relevant information about the technical quality of audiovisual material. However, as no suitable format for exchanging impairment related information is available, interoperability between these systems and content management, documentation and restoration systems is very limited.

Starting from existing work in the audio domain we have defined a general MPEG-7 description scheme for the description of impairments of visual content. In addition, a number of specialised descriptors for common visual defects as well as a comprehensive impairment classification scheme have been defined.

The description of the proposed extensions, the XML Schema defining the new description schemes and descriptors and the PrestoSpace Impairment Classification Scheme are available for download at [6]. In order to ensure the practical usability of the proposed extensions, we have implemented a C++ API for the new description schemes and descriptors.

## 6. Acknowledgements

The authors would like to thank Jean-Etienne Noiré (Institute National de l'Audiovisuel) and Przemyslaw Maziewski (Gdansk University of Technology) for their input. The research leading to this paper was partially supported by the European Commission under

the project "PrestoSpace" (FP6-2002-IST-1, <http://www.prestospace.org>).

## 7. References

- [1] Metadata Dictionary Registry of Metadata Element Descriptions. SMPTE RP210.8, 2004..
- [2] BWF – a format for audio data files in broadcasting EBU Tech 3285, 2001.
- [3] ISO/IEC, Information Technology – Multimedia Content Description Interface, Part 4: Audio, ISO/IEC 15938-4:2002/Amd 1:2004.
- [4] ISO/IEC, Information Technology – Multimedia Content Description Interface, Part 5: Multimedia Description Schemes, ISO/IEC 15938-5:2001.
- [5] Peter Schallauer, Werner Bailer, Roland Mörzinger, Hermann Fürntratt and Georg Thallinger, "Automatic Quality Analysis for Film and Video Restoration," Proceedings of IEEE International Conference on Image Processing, San Antonio, TX, 2007.
- [6] Audiovisual Defect and Quality Description. URL: <http://mpeg7.joanneum.at>.
- [7] Werner Bailer and Peter Schallauer, "The Detailed Audiovisual Profile: Enabling Interoperability between MPEG-7 Based Systems," Proc. of 12th International Multi-Media Modelling Conference, Beijing, CN, Jan. 2006.
- [8] The Brava broadcast archive programme impairments dictionary. URL: [http://brava.ina.fr/brava\\_public\\_impairments\\_list\\_en.html](http://brava.ina.fr/brava_public_impairments_list_en.html).