

A novel keycode for mono-hierarchical and multi-axial categorization of radiological images

Introduction

Modern communication standards such as DICOM include non-image data for a standardized description of study, patient, body region examined, and technical parameters related to the imaging modality in use. However, DICOM tables have been found to be insufficiently de-tailed for order entry systems [1] or even invalid, if values are set automatically by the system [2]. Concerning image retrieval in medical applications (IRMA), a detailed coding scheme is required to describe the imaging modality including technical parameters, the orientation of the image with respect to the body, the body region examined, and the biological system under investigation. Furthermore, the code must be strictly hierarchical in order to support semantic queries on a database. In this paper, we present a mono-hierarchical multi-axial classification code for medical images [3] in English language and emphasize its advantages for image retrieval in comparison to existing proposals supplementary to the DICOM standard.

Material and Method

Within the IRMA system [4], categorization is the first of seven successive analyzing steps extracting content information from medical images. It aims to establish intelligent processing strategies adapted to the current image under investigation. Therefore, categorization is required to be hierarchical. In other words, valid relations between code and sub-code elements are "is-a" and "part-of" only. Furthermore, causality is important for grouping of processing strategies. Therefore, a mono-hierarchical scheme is required, where each sub-code element is connected to only one code element. Since categorization of medical images must cover all aspects influencing the image content and structure, a multi-axial scheme was designed. This scheme, which has been presented previously in German language [2] is now available in English language, too.

Results

The IRMA coding system (IRMA: TTTT-DDD-AAA-BBB) consists of four axis with three to four positions, each in {0,...9,a,...,z}, where "0" denotes "not specified" to determine the end of a path along an axis:

1. T (technical): image modality
2. D (directional): body orientation
3. A (anatomical): body region examined
4. B (biological): biological system examined

The T-code describes within a maximum of four positions the technical method. The digits contain the physical source (e.g.: 1 x-ray, 2 ultrasound), the modality (e.g.: 11 plain film projection radiography, 12 fluoroscopy), the technique (e.g.: 111 digital, 112 analog), and the sub-techniques (e.g.: 1111 tomography, 1112 high energy). The D-code incorporates a two-step orientation description starting with the common orientation (e.g.: 1 coronal, 2 sagittal) and giving a more detailed specification in the second position (e.g.: 11 posterior-anterior (pa), 12 anterior-posterior (ap)). The third position describes functional orientation tasks of the examination (e.g.: 111 standing, 112 lying). The A-code unambiguously describes the anatomical region. For each of the nine major regions (e.g.: 1 total body, 2 head/skull), up to two hierarchical sub-codes are defined (e.g.: 3 spine, 31 cervical spine, 311 atlas (C1)). The B-code determines the organ system that is imaged. This axis is necessary because the body region examined insufficiently describes content and structure of images. For example, fluoroscopy of the abdominal region may access the vascular or the gastrointestinal system depending on the way the contrast agent is administered, which results in different image textures. On the top-level of this 3 digit IRMA-code, ten organ systems are specified (e.g.: 1 cerebrospinal system, 2 cardiovascular system) each having up to three positions to exactly identify the organ in question (e.g.: 1 cerebrospinal system, 11 central nervous system, 111 mesencephalon).

Discussion/Conclusions

A mono-hierarchical multi-axial code scheme is presented that enables unique classification of medical images. The entire code results in a character string of not more than 13 characters. The code can be easily extended by introducing characters in certain code positions, e.g., if new modalities are introduced. In contrast to other approaches, mixtures of one- and two-literal code positions are avoided which simplifies automatic code processing. Furthermore, the IRMA code steers clear of unambiguities resulting from overlapping code elements within the same level.

Existing nomenclatures describing all parameters influencing the image content and structure are not causal (e.g., the Mesh thesaurus is a poly-hierarchical structure where entities can be reached on different paths), non-hierarchical (e.g., "extremity", "arm", "hand" are valid entries of the DICOM tag 0018/0015 "body part examined") incomplete (e.g., the SNOMED/DICOM-Microglossary offers "breast" but not "mammary gland") or unambiguous (e.g., JJ1017 offers "chest", "chest/abdomen", and "abdomen" on the first level of its "large region code") and hence, not suitable for content-based image retrieval in medical applications. Likewise the JJ1017-approach, the IRMA-code enables image classification with respect to technical, directional, anatomical, and biological criterions. In contrast to JJ1017, IRMA is more detailed, offers four instead of three separated axis for the four parameters de-scribed above, and is strictly mono-hierarchical in each axis. Although this code was originally designed to be used in the IRMA-project, other use is welcome.

References

- [1] Kimura M, Kuranishi M, Sukenobu Y, Watanabe H, Nakajima T, Morimura S, Kabata S: JJ1017 image examination order codes - standardized codes supplementary to DICOM for image modality, region, and direction. Proceedings SPIE 2002; 4685: 307-315.
- [2] Güld MO, Kohnen M, Keyzers D, Schubert H, Wein BB, Bredno J, Lehmann TM: Quality of DICOM header information for image categorization. Proceedings SPIE 2002; 4685: 280-287.
- [3] Lehmann TM, Wein BB, Keyzers D, Kohnen M, Schubert H: A monohierarchical multiaxial classification code for medical images in content-based retrieval. Proceedings IEEE International Symposium on Biomedical Imaging, Washington, DC, July 7-10, 2002; 313-316.
- [4] Lehmann TM, Wein B, Dahmen J, Bredno J, Vogelsang F, Kohnen M: Content-based image retrieval in medical applications: A novel multi-step approach. Proceedings SPIE 2000; 3972: 312-320.