

Entwurf eines DICOM Structured Report am Beispiel Content-Based Image Retrieval

Petra Welter¹, Ralph Gülpers¹, Thomas M. Deserno¹, Jörg Riesmeier²
Marco Eichelberg³, Michael Onken³, Christoph Grouls⁴, Rolf W. Günther⁴

¹Institut für Medizinische Informatik, RWTH Aachen

²ICSMED AG, Oldenburg

³OFFIS - Institut für Informatik, Oldenburg

⁴Klinik für Radiologische Diagnostik, Universitätsklinikum Aachen

`pwelter@mi.rwth-aachen.de`

Kurzfassung. Der Einsatz von Digital Imaging and Communications in Medicine (DICOM) Structured Reporting (SR) ermöglicht einen standardisierten Austausch zwischen dem Picture Archiving and Communication System (PACS) und anderen klinischen Systemkomponenten. Ein strukturiertes Standardformat erlaubt die automatische Interpretation und eine Auswertung mittels Data Mining. Jedoch decken die im DICOM Standard festgelegten Structured Reporting Templates nicht alle Domänen ab. Wir zeigen beispielhaft das Design eines privaten SR-Templates für den Einsatz im Content-Based Image Retrieval (CBIR). Unser Lösungsansatz basiert auf dem Standard Computer-Aided Detection/Diagnosis (CAD) Template für Mammographie. Die Wiederverwendung bewährter SR-Module verspricht ein zuverlässiges Design. Wir analysieren die speziellen CBIR-Anforderungen und integrieren das neue Konzept von ähnlichen Bildern in unser Template. Allgemein anerkannte Templates für die Darstellung und den Austausch von Ergebnissen in einem standardisierten Format fördern den verbreiteten Einsatz neuer Verfahren in der klinischen Routine.

1 Einleitung

Der Nutzen neuer Verfahren in der medizinischen Informatik hängt wesentlich davon ab, wie gut sie in den Alltag eines Arztes integriert und angewendet werden können. Geht es um Ergebnisse, die ausgetauscht, präsentiert und dauerhaft gespeichert werden sollen, so ist die Verwendung eines Standardformats unabdingbar. Im klinischen Betrieb ist dies Digital Imaging and Communications in Medicine (DICOM). Für die Repräsentation von strukturierten Daten bietet DICOM das Structured Reporting-Format (SR).¹ Standard Templates² definieren zulässige Inhalte und Werttypen. Sie sind notwendig, um SR-Dokumente automatisch auswerten und systematisch nutzen zu können [1]. Neben z. Zt. ca. 35

¹ ftp://medical.nema.org/medical/dicom/2008/08_03pu.pdf

² ftp://medical.nema.org/medical/dicom/2008/08_16pu.pdf

allgemeinen Templates existieren spezielle für z.B. EKG und Röntgenstrahlendosis. Einige Templates können erweitert werden, falls dies für eine bestimmte Anwendung erforderlich ist. Ansonsten können neben den Standard-Templates auch private Templates definiert werden. Dies wird im Folgenden für die inhaltsbasierte Bildsuche (CBIR) durchgeführt. Der Nutzen von CBIR im medizinischen Bereich ist vielfach belegt [2]. Die Suche nach inhaltlich ähnlichen Bildern, ungeachtet ihrer textuellen Kategorisierung, und die damit ermittelten ähnlichen Patientenfälle, sollen als Diagnoseunterstützung eines Radiologen dienen.

2 Material und Methoden

2.1 Inhaltliche Anforderungen an das Template

Bei den inhaltlichen Anforderungen an ein CBIR-Template bedarf es u. a. der Angabe von Region of Interest (ROI). Um das Beispiel zu vereinfachen, beschränken wir uns auf folgende: a) ein Anfragebild, zu dem ähnliche Bilder zu suchen sind, b) Ergebnismenge der ermittelten ähnlichen Bilder, c) ihre jeweilige Ähnlichkeitsgüte und d) die Angabe des CBIR-Verfahrens.

2.2 Evaluierung existierender Templates

Im Bereich CAD finden sich unter den Standard Templates z. B. folgende:

- TID 2000 - Basic Diagnostic Imaging Report: Kann weder inkludiert noch erweitert werden. Es sieht selbst keine Beschreibung des CBIR/CAD-Verfahrens (d) vor. Die einzige Datenstruktur, die Anfragebild (a), Ergebnismenge (b) und Ähnlichkeitsgüte (c) aufnehmen könnte, ist *Diagnostic Imaging Report Elements*. Die Semantik für CBIR wird hier jedoch nicht deutlich.
- TID 4014 - CAD Image Quality: Hier ist nicht eine Ähnlichkeitsgüte (d) gemeint, sondern die Bildqualität, z. B. Stärke des Bildrauschens.
- TID 4000 - Mammography CAD Document Root: In dem inkludierten TID 4017 bzw. 4018 wird die Bildmenge genannt, auf der eine CAD Detektion bzw. Analyse durchgeführt wurde. Eine Ergebnismenge aller ähnlichen Bilder (b) kann nicht spezifiziert werden. TID 4017 und 4018 beinhalten Kennzeichnungen für bedeutsame Bildbereiche. Die Angabe einer ermittelten Ähnlichkeitsgüte (c) fehlt. Jedoch ist mit TID 4019 die Beschreibung zum CAD-Verfahren (d) aufgeführt.

2.3 Form und Aufbau einer Template Definition

Ein DICOM SR-Dokument enthält einen Dokumentenkopf (Header), wie ihn auch DICOM-Bilder besitzen. Er wird um einen Dokumentenbaum erweitert, der Struktur und Inhalt des eigentlichen Reports beinhaltet [3]. Templates beziehen sich auf den Dokumentenbaum. Sie werden als Tabelle spezifiziert. Jede Tabellenzeile definiert einen Baumknoten, der die Beziehung zum Vaterknoten (Rel)

Tabelle 1. Übersicht der Knotenbeziehungen in einer Template-Definition [1].

Nr.	Beziehung	Vaterknoten...
1	CONTAINS	enthält Komponente
2	HAS OBS CONTEXT	hat Beobachtungskontext
3	HAS CONCEPT MOD	hat Konzeptänderung
4	HAS PROPERTIES	hat Eigenschaft
5	HAS ACQ CONTEXT	hat Aufnahmekontext
6	INFERRED FROM	ist schlussgefolgert aus
7	SELECTED FROM	ist ausgewählt von

angibt, sowie Werttyp (VT), Konzeptnamen zur Erläuterung der Inhaltskomponente (Concept Name), Häufigkeit des Vorkommens (VM), Baumtiefe (NL) und ob der Knoten erforderlich ist (RT). In Tab. 1 sind mögliche Knotenbeziehungen erläutert. Ein SR-Dokument kann, beginnend ab dem Wurzelknoten des Dokumentenbaums, ein Root Template enthalten, welches wiederum durch Angabe von Template-Bezeichnern (TID) weitere Templates inkludiert.

2.4 Erzeugen eines SR-Dokuments

Bei der Erzeugung eines SR-Dokuments empfiehlt sich die Verwendung eines DICOM Toolkits mit entsprechender Unterstützung für die SR-Datenstrukturen. Wir verwenden DCMTK von OFFIS³. Die C/C++ Programmquellen lassen sich unter Windows, verschiedenen Unix-Betriebssystemen, sowie Mac OS X übersetzen. Ein alternatives bewährtes Open Source Toolkit [4] für Java ist z.B. dcm4che.⁴ DCMTK bietet mit dem Modul dcmsr Methoden für Zugriff, Lesen, Erstellen, Schreiben, Modifizieren, Drucken und Darstellen. Das folgende Beispiel für TID 4019 zeigt, wie ein neuer Knoten mit tieferer Verschachtelung hinzugefügt und der Name des Algorithmus angegeben wird:

```
srDoc->getTree().addContentItem(DSRTypes::RT_contains, \
    DSRTypes::VT_Text, DSRTypes::AM_belowCurrent);
srDoc->getTree().getCurrentContentItem().setConceptName \
    (DSRCodedEntryValue("111001", "DCM", "Algorithm Name"));
srDoc->getTree().getCurrentContentItem().setStringValue("IRMA");
```

3 Ergebnisse

Unser Template *CBIR_Root* (Tab. 2) beinhaltet in Zeile 2 das Anfragebild (a). Es inkludiert in Zeile 3 das Standard Template TID 4019 zur Angabe des CBIR-Systems (d) und in Zeile 3 das private Template *CBIR_Results* zur Angabe der

³ <http://dicom.offis.de>

⁴ <http://www.dcm4che.org>

Tabelle 2. SR-Template *CBIR_Root*.

Line	NL	Rel	VT	Concept Name	VM	RT
1			CONTAINER	EV(, , „CBIR Report“)	1	M
2	>	1	IMAGE	EV(, , „Query Image“)	1	M
3	>	1	INCLUDE	DTID(4019) „CAD Algorithm Identification“	1	M
4	>	1	INCLUDE	DTID(CBIR_Result) „CBIR Ergebnisse“	1	M

Tabelle 3. SR-Template *CBIR_Results*.

Line	NL	Rel	VT	Concept Name	VM	RT
1			CONTAINER	EV(, , „CBIR Results“)	1	M
2	>	1	CONTAINER	EV(, , „Scored Images“)	1-n	U
3	>>	1	IMAGE	EV(, , „Image“)	1	M
4	>>	1	NUM	EV(, , „Similarity Score“)	1	M

CBIR-Ergebnisse. *CBIR_Results* (Tab. 3) enthält in Zeile 2 einen Container für die Ergebnismenge, die sich in Zeile 3 aus einem Bild (b) und in Zeile 4 aus der zugehörigen Ähnlichkeitsgüte (c) zusammensetzt.

Die Erzeugung des vorgestellten Templates wurde als Applikation in C++ und DCMTK implementiert. IRMA [5], das CBIR-System unserer Wahl, übermittelt die CBIR-Ergebnisse an diese Applikation, welche das Template auf die Daten anwendet und ein DICOM SR-Dokument erstellt.

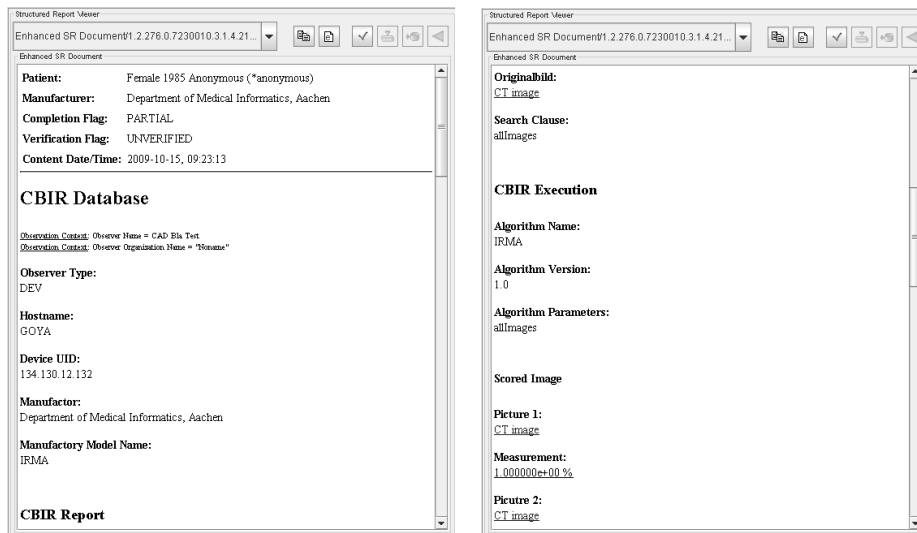


Abb. 1. Ausgabe des CBIR-SR in DICOMscope. Teile eins und zwei sind links und rechts dargestellt.

Zur Betrachtung des erzeugten SR-Dokuments verwenden wir DICOMscope von OFFIS. DICOMscope erzeugt wie die meisten SR-Viewer eine HTML-Ausgabe [6]. Das Layout ist allgemein gehalten (Abb. 1). Eine Konvertierung in das XML-Format mit dem Kommandozeilenprogramm *dsr2xml* erlaubt eine angepasste Darstellung über XSLT oder CSS. Alternativ ist mit dem Kommando *dsr2html* eine HTML-Konvertierung mit anschließender CSS-Anwendung möglich.

4 Diskussion

Dieser Beitrag bietet als Beispiel aus dem Bereich CBIR/CAD eine Hilfe bei der Erstellung privater SR-Dokumente. Die Verwendung strukturierter Reporte in einem standardisierten Format ermöglicht den zuverlässigen und einfachen Austausch von Ergebnissen. Dies fördert die Integration und damit den verbreiteten Einsatz von neuen Verfahren in der klinischen Routine.

SR-Dokumente werden u.a. in vordefinierten Arbeitsabläufen bzw. Profilen der Integrating the Healthcare Enterprise (IHE)⁵ eingesetzt. Für die Integration von CBIR in die radiologische Routine planen wir die Einbettung des vorgestellten SR-Templates im Post-Processing Profil, welches im Anschluss an die Bilderzeugung einsetzt.

Danksagung. Diese Arbeit wurde in Teilen von der Deutschen Forschungsgemeinschaft (DFG) gefördert, Le 1108/9

Literaturverzeichnis

1. Riesmeier J. Ein generisches Verfahren zur adaptiven Visualisierung von strukturierten medizinischen Befundberichten [PhD]. Universität Oldenburg. Oldenburg; 2006.
2. Müller H, Michoux N, Bandon D, et al. A review of content-based image retrieval systems in medical applications: clinical benefits and future directions. *Int J Med Inform.* 2004;73(1):1–23.
3. Eichelberg M, Aden T, Riesmeier J, et al. A survey and analysis of electronic healthcare record standards. *ACM Compute Surv.* 2005;37(4):227–315.
4. Vasquez A, Bohn S, Gessat M, et al. Evaluation of open source DICOM frameworks. *Proc MCIM Workshop*; 2007. p. C–10.
5. Lehmann T, Güld M, Thies C, et al. Content-based image retrieval in medical applications. *Methods Inf Med.* 2004;43(4):354–61.
6. Hussein R, Engelmann U, Schroeter A, et al. DICOM structured reporting part 2. problems and challenges in implementation for PACS workstations. *RadioGraphics.* 2004;24(3):897–909.

⁵ http://static.ihe.net/Technical_Framework/upload/ihe_tf_rev8.pdf