

Konverter DICOM von/nach PGM und PPM

by Thomas Hacklaender, MD, MSc (hacklaender@iftm.de)
IFTM Institut für Telematik in der Medizin GmbH

Dieses Kurzhandbuch beschreibt zwei Programme, die DICOM und PPM/PGM ineinander konvertieren.

Die Programme unterliegen der GNU General Public License. Die Bibliotheksfunktionen unterliegen der GNU Lesser General Public License. Siehe den Source-Code für Details.

Die Programme basieren auf der *dcm4che* DICOM Bibliothek von Gunter Zeilinger. Die Homepage dieses Projektes ist <http://sourceforge.net/projects/dcm4che/>.

Inhaltsverzeichnis

1	Installation	3
2	Beschreibung der Programme.....	4
2.1	Allgemeine Informationen	4
2.2	DICOM Secondary Capture Bilder.....	5
2.3	Metadatenformat in den PGM/PPM Dateien	6
2.4	dcm2pgm	7
2.5	pgm2dcm.....	9
	References	11

1 Installation

Die Programme benötigen eine zu Java 1.4. kompatible Runtime-Umgebung.

Die Beispiel-Skripte setzen voraus, dass die Umgebungsvariablen des Betriebssystems durch die Installation von Java so gesetzt worden sind, dass die Virtuelle Maschine im Suchpfad liegt.

Das Installationsverzeichnis wird im Folgenden mit <HOME> bezeichnet.

Alternativ kann auch eine Java-Runtime-Umgebung (JRE) in das Installationsverzeichnis kopiert werden. Dann müssen die Referenzen auf die Virtuelle Maschine in den Beispielen auf relative Pfadangaben auf diese virtuelle Maschine ersetzt werden. Befindet sich die JRE beispielsweise in dem Verzeichnis <HOME>/j2re1.4.0, dann lautet die absolute Referenz

```
./j2re1.4.0/bin/java
```

Die hier beschriebenen Programme sind dem ZIP Archiv `pgm.zip` zusammengefasst. Das Archiv muss ins Verzeichnis <HOME> entpackt werden. Dabei entsteht das Verzeichnis `/pgm` mit folgendem Inhalt:

File/Directory	Beschreibung
dcm2pgm.bat	Ein Windows Startskript für das Programm dcm2pgm.
dcm2pgm.jar	Das Programm dcm2pgm.
default_comments.txt	Die Default-Metadaten in der Form von Comments der PGM Dateien
gpl.txt	Die GNU General Public Lizenz.
/in/ CT.dcm	Ein Beispiel DICOM Bild (Computertomographie)
/in/RGB.ppm	Ein Beispiel Farb-Bild im PPM Forma
lgpl.txt	Die GNU Lesser General Public Lizenz.
log4j.properties	Property file für den Logger log4j.
pgm2dcm.bat	Ein Windows Startskript für das Programm pgm2dcm.
pgm2dcm.jar	Das Programm pgm2dcm.
src.zip	Der source code der Programme.

Um die Programme nutzen zu können, sind einige Zusatzbibliotheken notwendig. Es wird erwartet, dass sich diese in dem Verzeichnis <HOME>/lib befinden:

File/Directory	Beschreibung
dcm4che.jar	Die low -level DICOM Bibliothek.
getopt.jar	Eine Bibliothek für command-line Optionen.
links.txt	Links zum SourceCode und weiteren Informationen zu diesen Bibliotheken.
log4j.jar	Die Bibliothek für den logger von dcm4che.

Diese Bibliotheken können als gesondertes ZIP Archiv `lib.zip` heruntergeladen werden. Es wird allerdings empfohlen, statt dessen das ImageJ-DICOM Paket http://www.iftm.de/dicom/dicom_projekte/dcmie.htm einschließlich der Netzwerkkomponenten zu installieren. Entsprechende Manuals sind online verfügbar.

2 Beschreibung der Programme

2.1 Allgemeine Informationen

Innerhalb von DICOM wird ein Objekt (Bild) durch die hierarchische Angabe von Größen zum Patienten, zur Studie, zur Serie und zum Objekt selbst gekennzeichnet. Dabei ist zu berücksichtigen, dass zu einem Patient eine oder mehrere Studien existieren können, die ihrerseits aus einer oder mehreren Serien bestehen. Diese beinhalten ebenfalls ein oder mehrere Objekte.

In diesem Ordnungsprinzip gibt es zwei unabhängige Bezeichnungsebenen: Einmal die dem Anwender präsentierte mit semantisch bedeutungsvolle Inhalt

Kürzel	VR	Beschreibung
PatientName	PN	Name des Patienten
StudyID	SH	Identifikationsnummer der Studie
SeriesNumber	IS	Nummer der Serie
InstanceNumber	IS	Nummer des Bildes

Und eine zweite Ebene, die die (weltweite) Eindeutigkeit der Objekte gewährleistet:

Kürzel	VR	Beschreibung
PatientID	LO	Identifikationsnummer des Patienten (meist aus Krankenhausverwaltungssystem)
StudyUID	UI	Unique Identifier (UID) der Studie.
SeriesUID	UI	Unique Identifier (UID) der Serie.
SOPInstanceUID	UI	Unique Identifier (UID) der Bilddatei.

Der Inhalt der zur Bezeichnung verwendeten Tags unterliegt ebenfalls einer Normierung, die über die sogenannte Value Representation VR geregelt ist.

VR	Name	Beschreibung
IS	Integer String	A string of characters representing an Integer in base-10 (decimal), shall contain only the characters 0 - 9, with an optional leading "+" or "-". It may be padded with leading and/or trailing spaces. Embedded spaces are not allowed. The integer, n, represented shall be in the range: $-2^{31} \leq n \leq (2^{31} - 1)$. 12 Bytes maximum.
LO	Long String	A character string that may be padded with leading and/or trailing spaces. The character code 5CH (the BACKSLASH) shall not be present, as it is used as the delimiter between values in multiple valued Default Character Repertoire and/or as defined by (0008,0005). 64 chars maximum.
PN	Person Name	A character string encoded using a 5 component convention. The character code 5CH (the BACKSLASH) shall not be present, as it is used as the delimiter between values in multiple valued data elements. The string may be padded with trailing spaces. The five components in their order of occurrence are: family name complex, given name complex, middle name, name prefix, name suffix. Any of the five components may be an empty string. The component delimiter shall be the caret '^' character (5EH). Delimiters are required for interior null components. Trailing null components and their delimiters may be omitted. Multiple entries are permitted in each component and are encoded as natural text strings, in the format preferred by the named person. This conforms to the ANSI HISPP MSDS Person Name common data type. 64 character maximum per component group.
SH	Short String	A character string that may be padded with leading and/or trailing spaces. The character code 05CH (the BACKSLASH) shall not be present, as it is used as the delimiter between values for multiple data elements. The string

		string shall not have Control Characters except ESC. Default Character Repertoire and/or as defined by (0008,0005). 16 chars maximum.
UI	Unique Identifier (UID)	A character string containing a UID that is used to uniquely identify a wide variety of items. The UID is a series of numeric components separated by the period "." character. If a Value Field containing one or more UIDs is an odd number of bytes in length, the Value Field shall be padded with a single trailing NULL (00H) character to ensure that the Value Field is an even number of bytes in length. 64 chars maximum.

Ein Unique Identifier (UID) setzt sich aus einem root-Präfix und einem Postfix zusammen. Der Präfix ist weltweit eindeutig und lautet für diese verwendete low-level Bibliothek:

root-UID-Präfix = 1.2.40.0.13.1.1

Der Postfix kann beliebig gewählt werden. Der Anwender ist selbst dafür verantwortlich, dass der UID eindeutig bleibt. Die Ziffernfolgen innerhalb des UID haben keine semantische Bedeutung.

2.2 DICOM Secondary Capture Bilder

Normalerweise werden DICOM Bilder von Modalitäten, z.B. CT oder Kernspintomograph, erzeugt. Bilder, die in einem Nachverarbeitungsschritt entstehen, werden als Secondary Capture Bilder bezeichnet. Die entsprechende IOD (Information Object Definition) ist in PS 3.3 - A.8. definiert:

IE	Module Name	Reference PS 3.3	Option
Patient IE	Patient Module	C.7.1.1	Mandatory
Study IE	General Study Module	C.7.2.1	Mandatory
Study IE	Patient Study Module	C.7.2.2	User Option
Series IE	General Series Module	C.7.3.1	Mandatory
Equipment IE	General Equipment Module	C.7.5.1	User Option
Equipment IE	SC Equipment Module	C.8.6.1	Mandatory
Image IE	General Image Module	C.7.6.1	Mandatory
Image IE	Image Pixel Module	C.7.6.3	Mandatory
Image IE	SC Image Module	C.8.6.2	Mandatory
Image IE	Overlay Plane Module	C.9.2	User Option
Image IE	Modality LUT Module	C.11.1	User Option
Image IE	VOI LUT Module	C.11.2	User Option
Image IE	SOP Common Module	C.12.1	Mandatory

Die Angaben in der Image IE (Information Entity) werden aus dem zu exportierten Bild direkt abgeleitet. Der Anwender hat hierauf keinen Einfluss.

Als Default-Werte werden in den Programmen alle Attribute aufgenommen, deren Typ 1 (Required w/ value) or 2 (Required w/o value) ist. Für das Equipment Module werden auch einige Typ 3 (Optional) Attribute mit aufgenommen. Dies führt zu folgender Liste von Attributen für die Default-Metadaten:

Attribut Name	Gruppe, Element	Wert)* Werte werden zur Laufzeit berechnet
SOP Class UID	0008, 0016	1.2.840.10008.5.1.4.1.1.7
SOP Instance UID	0008, 0018	1.2.40.0.13.1.1.192.168.0.2.20020626183926558.32773)*
Study Date	0008, 0020	20020626)*
Content Date	0008, 0023	20020626)*
Study Time	0008, 0030	183926.598)*
Content Time	0008, 0033	183926.598)*
Accession Number	0008, 0050	0
Modality	0008, 0060	OT
Conversion Type	0008, 0064	WSD
Referring Physician's Name	0008, 0090	ReferringPhysicianName
Patient's Name	0010, 0010	PatientName
Patient ID	0010, 0020	1025109564004
Patient's Birth Date	0010, 0030	19501031
Patient's Sex	0010, 0040	O
Secondary Capture Device ID	0018, 1010	<no value>
Date of Secondary Capture	0018, 1012	20020626)*
Time of Secondary Capture	0018, 1014	183926.598)*
Secondary Capture Device Manufacturer	0018, 1016	dcm4cheri
Secondary Capture Device Software Version(s)	0018, 1019	1.0
Study Instance UID	0020, 000d	1.2.40.0.13.1.1.192.168.0.2.20020626183926558.32771)*
Series Instance UID	0020, 000e	1.2.40.0.13.1.1.192.168.0.2.20020626183926558.32772)*
Study ID	0020, 0010	1
Series Number	0020, 0011	1
Instance Number	0020, 0013	1
Patient Orientation	0020, 0020	<no value>
Laterality	0020, 0060	<no value>

2.3 Metadatenformat in den PGM/PPM Dateien

Jedes DICOM Attribut wird als eine Kommentarzeile in die Datei aufgenommen. Das Format ist

```
# Dicom/( <group> , <element> ) <name> : <value>
```

Leerstellen sind signifikant. Die einzelnen Abschnitte sind folgendermaßen definiert:

Abschnitt	Beschreibung
<group>	Die Gruppennummer des Attributes als 4 stellige Hexadezimalzahl. In DICOM genormt.
<element>	Die Gruppennummer des Attributes als 4 stellige Hexadezimalzahl. In DICOM genormt.
<name>	Eine Beschreibung des Attributes. Der Text ist nicht genormt und kann sich in Zukunft ändern. Der Text darf die Länge 0 haben. Er darf nicht das Zeichen „\“ enthalten.
<value>	Der Inhalt des Attributes als String. Der String darf die Länge 0 haben. Für Attribute mit einer Value Multiplicity größer als 1 werden die Einzelwerte durch den String „\“ voneinander getrennt.

2.4 dcm2pgm

Das Programm wandelt DICOM Dateien in PGM/PPM Dateien um. Für DICOM Dateien mit der Photometric Interpretation „MONOCHROME2“ wird eine PGM Datei mit der magic-number „P5“ erzeugt. Für Dateien mit der Photometric Interpretation „RGB“ wird eine PPM Datei mit der magic-number „P6“ erzeugt.

dcm2pgm version 1.0

Usage:

dcm2pgm [commands] <source>

Converts a DICOM file to a PGM or PPM file, depending on the contents of the source. By default only default metadata required for a Secondary Capture image are written besides the image data.

<source> the file or directory containing the DICOM file(s).

Commands:

-a <study_id> Set the StudyID tag (00200010) to <study_id>.
-A <study_uid> Set the StudyInstanceUID tag (0020000d) to <study_uid>.
-b <series_num> Set the SeriesNumber tag (00200011) to <series_num>.
-B <series_uid> Set the SeriesInstanceUID tag (0020000e) to <series_uid>.
-c <image_num> Set the InstanceNumber tag (00200013) to <image_num>.
-C <sop_uid> Set the SOPInstanceUID tag (00080018) to <sop_uid>.
-d <file-uri> The name of the destination file or a directory.
-F Include all DICOM tags as image comments.
-h Print this help message.
-n <name> Set the PatientName tag (00100010) to <name>.
-N <pat_id> Set the PatientID tag (00100020) to <pat_id>.
-x Include only DICOM tags required for a Secondary Capture image.
 If not enough information is given in the base image, default
 values are added.
-X Include no DICOM tags in the image.

<file-uri> Describes a file in a operating-system independent way. See the
 API-Doc of the URI class. For Windows-OS the absolute URI
 file:/c:/user/tom/foo.txt
 describes the file
 C:\user\tom\foo.txt
 Relative URI's, e.g. without the 'file:' schema-prefix, are
 relative to the user-directory, given by the system property
 user.dir. For example: If the user.dir is
 C:\user\tom\
 and the relative URI is
 /abc/foo.txt
 the referenced file is
 C:\user\tom\abc\foo.txt
 The abbreviations '.' for the current and '..' for the upper
 directory are valid to form a relative URI.

(C) 2002 Thomas Hacklaender, <mailto:hacklaender@iftm.de>

IFTM Institut fuer Telematik in der Medizin GmbH, www.iftm.de

This library is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

Wird als <source> eine einzelne Datei angegeben, wird nur diese konvertiert. Ist es eine Referenz auf ein Verzeichnis, werden alle Dateien in diesem Verzeichnis konvertiert. In diesem Fall entspricht der Name der Zieldatei dem der Source-Datei, wobei die Extension durch .pgm oder .ppm ersetzt wird. Wenn als Ziel eine Referenz auf ein Verzeichnis angegeben wurde, werden alle Dateien dort gespeichert. Falls nur eine Quelldatei angegeben wurde und das Ziel einen Dateinamen enthält, wird die Datei unter dem angegebenen Namen gespeichert.

Die Metadaten der DICOM Datei werden in Abhängigkeit von den Kommandos als Kommentare in die PxM Datei übernommen: Ohne besondere Angaben wird zur Runtime ein Satz von Default-Metadaten (siehe oben) erzeugt.

Mit dem Kommando `-F` werden alle Metadaten aus dem DICOM Bild übernommen (Ausnahme: Spezielle VRs. Siehe hierzu den Source-Code).

Mit den Kommandos `-aAbBcCnN` kann der Inhalt einzelner Tags überschrieben werden.

Mit dem Kommando `-X` wird der Einschluss von Metadaten vollständig unterdrückt.

2.5 pgm2dcm

Das Programm wandelt PGM/PPM in DICOM Dateien um. PPM Dateien mit 8-Bit RGB-Bilder werden zu der Photometric Interpretation „RGB“ konvertiert, PGM Dateien mit 16-Bit Grauwertbildern zu „MONOCHROME2“.

pgm2dcm version 1.0

Usage:
pgm2dcm [commands] <source>

Converts a PGM or PPM file to DICOM. By default only default metadata required for a Secondary Capture image are written besides the image data.

<source> the file or directory containing the DICOM file(s).

Commands:

-a <study_id> Set the StudyID tag (00200010) to <study_id>.
-A <study_uid> Set the StudyInstanceUID tag (0020000d) to <study_uid>.
-b <series_num> Set the SeriesNumber tag (00200011) to <series_num>.
-B <series_uid> Set the SeriesInstanceUID tag (0020000e) to <series_uid>.
-c <image_num> Set the InstanceNumber tag (00200013) to <image_num>.
-C <sop_uid> Set the SOPInstanceUID tag (00080018) to <sop_uid>.
-d <file-uri> The name of the destination file or a directory.
-f Write all DICOM tags found in the base image. Add default metadata required for a Secondary Capture image. Preserve the following tags:
 PatientName (00100010)
 PatientID (00100020)
 PatientBirthDate (00100030)
 PatientSex (00100040)
-F Write only DICOM tags found in the base image.
-h Print this help message.
-n <name> Set the PatientName tag (00100010) to <name>.
-N <pat_id> Set the PatientID tag (00100020) to <pat_id>.

<file-uri> Describes a file in a operating-system independend way. See the API-Doc of the URI class. For Windows-OS the absolute URI
 file:/c:/user/tom/foo.txt
 describes the file
 C:\user\tom\foo.txt
Relative URI's, e.g. without the 'file:' schema-prefix, are
relative to the user-directory, given by the system property
user.dir. For example: If the user.dir is
 C:\user\tom\
and the relative URI is
 /abc/foo.txt
the referenced file is
 C:\user\tom\abc\foo.txt
The abbreviations '.' for the current and '..' for the upper
directory are valid to form a relative URI.

(C) 2002 Thomas Hacklaender, <mailto:hacklaender@iftm.de>
IFTM Institut fuer Telematik in der Medizin GmbH, www.iftm.de

This library is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

Wird als <source> eine einzelne Datei angegeben, wird nur diese konvertiert. Ist es eine Referenz auf ein Verzeichnis, werden alle Dateien in diesem Verzeichnis konvertiert. In diesem Fall entspricht der Name der Zieldatei dem der Source-Datei, wobei die Extension durch .pgm oder .ppm ersetzt wird. Wenn als Ziel eine Referenz auf ein Verzeichnis angegeben wurde, werden

alle Dateien dort gespeichert. Falls nur eine Quelldatei angegeben wurde und das Ziel einen Dateinamen enthält, wird die Datei unter dem angegebenen Namen gespeichert.

Die Kommentare der PXM Datei werden in Abhängigkeit von den Kommandos in die DICOM Datei übernommen: Ohne besondere Angaben wird zur Runtime ein Satz von Default-Metadaten (siehe oben) erzeugt.

Mit dem Kommando `-f` werden zunächst alle Kommentare als Metadaten übernommen. In einem zweiten Schritt werden diejenigen Tags, die sich auf die Patientenidentifikation beziehen, gesichert. Dann werden die Metadaten mit einem Satz von Default-Metadaten überschrieben. Metadaten, die nicht in den Kommentaren definiert, aber im Default-Datensatz enthalten sind, werden ergänzt. Dann werden die Metadaten mit den gesicherten Patientendaten überschrieben. Mit dieser Option ist gewährleistet, dass die DICOM Datei eine Secondary-Capture Datei ist, gleichgültig, welche Kommentare in der PXM Datei definiert waren.

Mit dem Kommando `-F` werden alle Kommentare als Metadaten übernommen. Achtung: Es ist nicht gewährleistet, dass die DICOM Datei eine Secondary-Capture Datei wird!

Mit den Kommandos `-aAbBcCnN` kann der Inhalt einzelner Tags überschrieben werden.

References

Zusammenstellung von ImageJ Plugins und Netzwerkfunktionalität:
http://www.iftm.de/dicom/dicom_projekte/dcmie.htm

Das ImageJ Framework:
<http://rsb.info.nih.gov/ij/>

Die dcm4che project Homepage:
<http://sourceforge.net/projects/dcm4che/>

Die DICOM Norm:
<http://www.dclunie.com/>