

1 Wstęp do formatu DICOM

DICOM - standard przechowywania, drukowania i przesyłania obrazowych danych medycznych. Opracowany w roku 1993 jako rozwinięcie standardu ACR/NEMA. Poza samymi danymi obrazowymi (takimi jak zdjęcia rentgenowskie, USG, zapisami EKG etc.) Zawiera informacje o pacjencie jak również kontekst i opis danych obrazowych: opis badania, format i struktura obrazu.

Dane opisowe zapisane są w specjalnie opracowanym *języku*. Format DICOM ma zdefiniowane 27 standardowych typów danych, określanych jako VR (value representation). Każdy VR określa długość danych (w bajtach), format zapisu (binarnie, znakowo) oraz dozwolone znaki. Format DICOM nie ma ustandaryzowanego kodowania znaków, co może utrudniać między narodową wymianę danych. Ustandaryzowany jest za to porządek bajtów w zapisie binarnym, domyślnie Little Endian.

Słownik DICOM definiuje stuktury danych określające ich przeznaczenie. Struktury te (nazywane atrybutami, elementami DICOM lub elementami danych DICOM) opisane są 4 bajtowym tagiem, gdzie dwa pierwsze bajty przyporządkowują go do grupy danych, a dwa kolejne dokładnie go określają. Grupy elementów miały dużo większe znaczenie w starszych wersjach standardu, teraz, gdy zdefiniowanych jest ponad 2000 elementów (i ponad to każdy może definiować swoje), grupa elementu bardzo luźno opisuje jego przynależność. Każdy element opórcz opisu danych które reprezentuje, ma też na stałe przypisany VR.

Elementy DICOM mogą być kodowane na dwa sposoby: Implicite i Explicite. Format Implicite jest krótszy w zapisie ponieważ nie zawiera VR elementów. Pozwala to zaoszczędzić miejsce/transfer, ale staje się uciążliwe przy wymienianiu danych które definiują własne niestandardowe elementy. Format Explicite służy do tego aby zapisać również VR co pozwala odczytać dane bez potrzeby otwierania słownika (choć i tak wtedy nie znamy przeznaczenia danych).

2 brain_001.dcm

Plik brain_001.dcm przedstawia jeden obraz wykonany podczas rezonansu magnetycznego głowy.

Plik jest okdowany w formacie explicite, dane obrazowe to aż 98% objętości pliku. Plik nie używa grupowania elementów.

To badanie MRI zostało wykonane 16 marca 2001 roku, o godzinie 14:30 na urządzeniu SIGNA firmy GE Medical Systems.

W pliku znajdują się częściowe dane o pacjencie. Była to kobieta w wieku 28 lat, o wadze 61 (prawdopodobnie) kilogramów.

Badanie zostało wykonane w technologii MRI 2D, grubość przekroju to

5cm, użyta siła pola magnetycznego, to 1.5. Pacjent podczas badania był w pozycji na wznak z głową do przodu.

Obraz ma rozmiar 256x256 pikseli, gdzie odstęp między nimi to 0.859375mm. Obraz jest zapisany w palecie monochromatycznej. Wartości mieszczą się w zakresie od 0 do 884. Element DICOM (0028,0103) *Pixel Representation* mówi o sposobie zapisu każdej próbki i może przyjmować 2 wartości: 0000H - zapis jako *unsigned integer* lub 0001H - zapis w kodzie uzupełnień do dwóch. W tym pliku używany jest zapis w kodzie uzupełnień do dwóch. Element (0028,0100) *Bits Allocated* mówi o szerokości bitowej pojedynczej próbki. W tym pliku każdy piksel (próbka) ma 16 bitów.

3 US-PAN-8-10x-echo.dcm

Plik US-PAN-8-10x-echo.dcm zawiera serię 10 obrazów z badania echokardiograficznego.

Dane w pliku zapisane są w formacie *explicite* z grupowaniem elementów.

Zapis badania został wykonany 30 stycznia 1994 roku, o godzinie 11.25 na sprzęcie firmy ACME Products.

W pliku brak informacji o pacjencie - dane zostały zanonimizowane.

Dane obrazowe to 10 kolejnych ramek zapisu echokardiografii. Każdy obraz ma rozmiar 430x600 pikseli. Obraz jest zapisany w palecie monochromatycznej. Każdy piksel ma 8 bitów, reprezentowany jest jako *unsigned integer* z kolejnością bitów *big endian*, najbardziej znaczący bit numer 7. Tablica przekodowań palety przekształca jedną z 256 wartości pikselna na 3 kolorowe 16 bitowe piksele. Minimalna wartość dla tablicy przekodowań to 0, wszystkie wartości mniejsze przekodowywane są na kolor odpowiadający wartości 0, choć w tym pliku to nie ma znaczenia, ponieważ dane są przechowywane w postaci liczb całkowitych bez znaku.

4 S-RGB-8-esopecho.dcm

Plik US-RGB-8-esopecho.dcm przedstawia jeden obraz z badania Echokardiograficznego.

Plik ten jest kodowany w formacie *explicit*. Jego wielkość to około 90kb, z czego danych obrazowych jest 99%. Plik używa grupowania elementów danych.

Jest to plik powstały podczas badania echokardiograficznego 5 listopada 1994 roku, o godzinie 14:30. Rejestracja została przeprowadzona przy pomocy urządzenia Acme Products.

W tym pliku dane pacjenta zostały usunięte, plik poddano anonimizacji.

Pobieranie danych zostało zrealizowane za pomocą protokołu Quad Capture. Jest to protokół rejestracji dźwięku.

Zawarty w pliku obraz jest obrazem numer 107 z 2 serii danych.

Parametry obrazu. Każdy piksel jest zapisany za pomocą trzech wartości. Pojedyncza wartość odpowiada jednemu z kolorów: czerwonego, niebieskiego lub zielonego. Piksele ułożone są obok siebie dla danego piksela. Rozmiar obrazu to 120 na 256 pikseli, z których każdy ma kształt prostokąta o proporcjach boków 4 do 3. Zapisane wartości pikseli są bez znaku.

W pliku piksele są zapisane w przedstawiony sposób. Trzy wartości każdego piksela są ułożone obok siebie w pliku.

5 Wnioski

W plikach największą częścią pod względem ilości danych są same obrazy, co oznacza, że dane opisowe o obrazie, dane o pacjencie, badaniu, nie są obciążające dany plik. Dzięki temu przechowywanie obrazów w formacie DICOM nie jest znacząco bardziej wymagające niż przechowywanie samych obrazów. Również przy transmisji obrazów najczęściej danych jest o samym obrazie i nie powoduje to spowolnienia transmisji przez dane opisujące obraz.

Obrazy w większości nie są skompresowane. Jest to spowodowane tym, że przy użyciu kompresji stratnej można by było stracić ważne informacje medyczne. W jednym z trzech plików zastosowano kompresję z pomocą użycia mapy kolorów.

Standard DICOM udostępnia możliwość dokładnego opisu jak powinien wyglądać obraz. Służą do tego specjalne parametry jak rozmiar piksela czy odległość pomiędzy pikselami.