De-identification of Faces in 2D DICOM Images

David Rodriguez Gonzalez
DIR group – SBIRC
University of Edinburgh

Data Protection Act

- UK's Data Protection Act (1998). Implements the European Community Data Protection Directive 1995.
- Establish individuals' rights on data held about them and obligations for organisations or people processing personal data.
- Personal data must be processed in a fair and lawful manner.
  - 8 DPA principles.
- Other legislation pieces apply to medical data.
  - Common law: duty of confidentiality.
  - Human Rights Act 1998 (article 8).

DPA in research

- The DPA does not define the term "research purposes" apart from clarifying that it includes statistical or historical purposes.
- Data processing for research should be 'compatible' with the purpose for which the data were originally obtained.
- The data subjects should be aware that their personal information will be used for research purposes.

Face Recognition

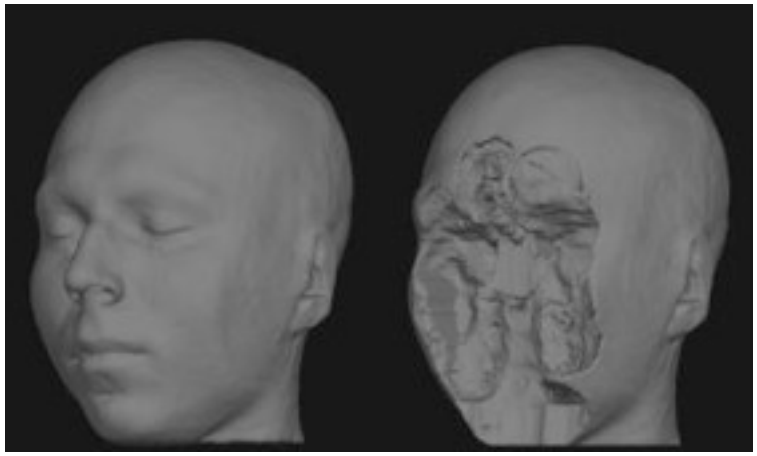
- It is possible to make 3D-reconstructions of the head from MRI or CT data
  - Risk of face recognition
  - Risk of automated face recognition?
- Several studies on face recognition from MR 3D reconstructions
  - In particular Chen et al (2007) & Prior et al (2009)
  - Risk of recognition even though face features missing (ears, eyebrows,...)

De-identification

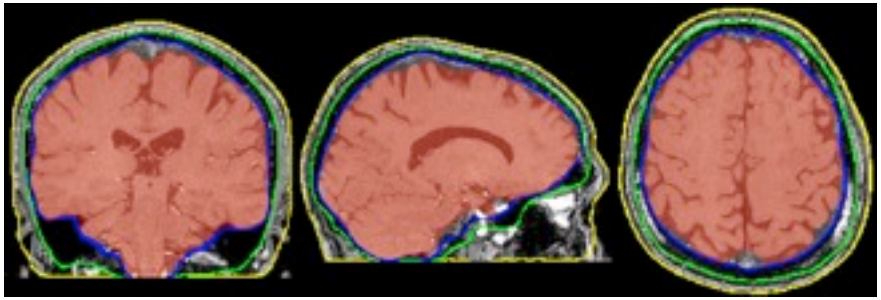
- Usually face de-identification in medical imaging is done after the reconstruction, i.e. in 3D
- Different techniques are used to this end including
  - brain extraction,
  - removal of facial features
  - and deformation of the face surface

Existing Tools

- MRI defacer (MBIRN)



- BET (Brain Extraction Tool) is part of the FMRIB software library.



Motivation

- We have a software (DICOM Confidential) for de-identification of DICOM headers
  - Researchers then use existing tools to remove the face after reconstruction
- In some scenarios we would prefer removing the face in the original DICOM 2D slices
  - DICOM receiver
  - Several potential users (avoid effort duplication)

Digital **I**maging and **Co**mmunications in **M**edicine (**DICOM**)

- Standard for handling, storing, printing and transmitting medical imaging information
  - Supports CT, MRI, PET, Nuclear medicine, ultrasound,…
  - Several types of objects: Images, Presentation States, Structured Reports, Encapsulated Objects
- Data format:
  - "Header": includes metadata
  - Pixel data
- Also defines communications, confidentiality profiles, …

Project

- Evaluate different possibilities and implement a prototype for 2D de-identification
  - Preferably in Java to ease the integration with existing software.
  - It would be desirable that the software handles both CT and MR images
- Define experiments for validation
  - As they require human intervention (and time) it is out of the scope their realisation

Possible Approach

- Literature review
  - Including which facial features are more relevant for face identification
- Machine learning to identify & locate those features
- Then remove or alter them without altering any of the brain pixels
- Nevertheless, other approaches might be explored