

La tomografia a raggi X per la digitalizzazione di manoscritti antichi

➤ Progetto di ricerca

L'obiettivo del progetto di ricerca nel quale si inserirà l'attività dell'assegnista è lo studio e lo sviluppo di una metodologia per la lettura e la digitalizzazione di antichi documenti manoscritti basata sulla tomografia a raggi X.

Le collezioni Italiane ed Europee di manoscritti antichi sono parte importante del nostro patrimonio culturale e contengono un'incredibile quantità di preziosi documenti ed informazioni. Oltre ai manoscritti consultabili dagli studiosi, una grossa parte di questi documenti non è disponibile in quanto molto fragili o, ad esempio, nel caso di numerosi testamenti, addirittura ancora sigillati. Per questo motivo negli ultimi anni si è accesa l'attenzione su progetti di digitalizzazione massiva, che permettano non solo la digitalizzazione di questi preziosi oggetti, ma anche la creazione di data-base indicizzati per la loro consultazione. Mentre per collezioni molto ben conservate e più recenti la tradizionale tecnica di digitalizzazione è sufficiente, per documenti molto fragili o ancora sigillati è necessario sviluppare procedure *ad hoc*.

La metodologia qui proposta, basata sulla tomografia a raggi X, acquisendo l'intero manoscritto senza la necessità di aprirlo, rende la procedura completamente non invasiva e minimizza l'interazione dell'operatore con l'oggetto. La lettura dei testi tramite raggi X è resa possibile grazie all'uso, in tutta Europa e per molti secoli, di inchiostri a base di ferro, denominati inchiostri ferro-gallici.

Negli ultimi anni, studi di fattibilità sono stati effettuati sia a livello nazionale che internazionale, soprattutto su pergamene, rotoli e oggetti estremamente preziosi, ma per poter rendere davvero applicabile questa tecnica su più ampia scala è necessaria la definizione di una metodologia *ad hoc*.

È fondamentale innanzitutto definire il design di un sistema tomografico a raggi X che sia utilizzabile *in situ* e che sia abbastanza flessibile da permettere l'analisi di differenti tipi di oggetti, da piccoli testamenti a documenti più voluminosi. Cruciale inoltre è il processing dei dati tomografici per l'estrazione virtuale delle singole pagine del documento al fine di renderne possibile la lettura. Il basso contrasto radiografico degli inchiostri, combinato con la non-omogeneità e non-planarità delle carte antiche, rende infatti i processi di segmentazione basati solamente sullo spessore delle pagine largamente insufficienti e richiede lo sviluppo e l'applicazione di algoritmi *ad hoc*.

Il progetto dovrebbe quindi essere sviluppato secondo due direttive parallele e strettamente connesse fra loro:

- il design di una facility *ad hoc*, ottimizzata per questo tipo di oggetti;
- lo sviluppo di algoritmi di segmentazione per poter separare ed "estrarre" le singole pagine di manoscritti e renderle così leggibili.

Cruciale inoltre è lo sviluppo di una metodologia flessibile, applicabile a differenti tipi di documenti, e disponibile *in situ*.

➤ Obiettivi

● *Design di una facility*

Per poter sviluppare un sistema tomografico ottimizzato per l'analisi *in situ* di antichi documenti manoscritti è necessario definire:

- le caratteristiche della sorgente, in termini di voltaggio e corrente, al fine di massimizzare il contrasto radiografico degli inchiostri;
- le caratteristiche del detector, in termini di dimensioni e risoluzione spaziale ottimali, necessarie per l'imaging di un ampio range di oggetti;
- la minima risoluzione volumetrica (voxel size) necessaria per l'estrazione delle singole pagine e per separarne il fronte dal retro.

A tal fine, verranno tomografati, variando i parametri di acquisizione, svariati oggetti creati *ad hoc* per simulare differenti tipi di manoscritti, da testamenti con poche pagine a oggetti più grandi come libri manoscritti. Verranno inoltre svolti test sperimentali anche con manoscritti storici.

● *Procedura di segmentazione*

Come per la facility, è necessario creare procedure *ad hoc* per la segmentazione dei manoscritti. In particolare, occorrerà sviluppare e/o ottimizzare:

- una metodologia di pre-processing dei dati, per l'ottimizzazione del contrasto fra le pagine ed il background;
- algoritmi di segmentazione semi-automatici o automatici, per poter estrarre le singole pagine.

A tal fine verranno utilizzate come immagini di input quelle ottenute dalle acquisizioni tomografiche sopramenzionate. I due obiettivi principali dovranno quindi procedere in parallelo e saranno strettamente interconnessi fra loro.

➤ ***Piano di attività***

Durante l'intera durata del progetto (mesi 1-12), l'assegnista parteciperà in prima persona alle acquisizioni tomografiche che verranno effettuate sia su oggetti creati per simulare manoscritti antichi sia su documenti storici. Per i primi test verrà utilizzato un sistema tomografico già disponibile presso il DIFA.

Alla luce dei risultati ottenuti nei primi mesi (1-3), si comincerà a definire meglio il design di un apparato tomografico ad hoc, concentrandosi in particolare sulle caratteristiche dei componenti chiave per le performance del sistema, ovvero la sorgente di raggi X e il detector. In parallelo e per i mesi seguenti (mesi 2-11) i risultati degli esperimenti tomografici saranno utilizzati per testare ed ottimizzare algoritmi per il pre-processing, la segmentazione e l'eventuale post-processing dei dati.

I risultati degli algoritmi di segmentazione verranno inoltre utilizzati per guidare le acquisizioni tomografiche e definire al meglio la risoluzione volumetrica minima necessaria per la separazione delle pagine e per la loro lettura.

A fine progetto, saranno quindi definiti sia il design della facility che la procedura di segmentazione dei manoscritti.