

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

Scuola di Scienze  
Dipartimento di Fisica e Astronomia  
Corso di Laurea in Fisica

SEGMENTAZIONE DI IMMAGINI  
TOMOGRAFICHE DI UN  
ANTICO DOCUMENTO

Relatore:

Dott.ssa Maria Pia Morigi

Presentata da:

Simone Fumagalli

Correlatori:

Dott.ssa Fauzia Albertin

Dott.ssa Rosa Brancaccio

Anno Accademico 2017/2018



# Abstract

Questo progetto di tesi è stato svolto all'interno di un più ampio progetto di studio dei testi antichi, i quali sono fragili e deteriorati a causa della loro età. L'interesse di questo studio riguarda l'utilizzo di tecniche non invasive per indagare approcci di digitalizzazione innovativi e meno invasivi per questi preziosi - e spesso fragili - oggetti. In questo ambito, presso il Dipartimento di Fisica e Astronomia dell'Università di Bologna, si stanno indagando tecniche di digitalizzazione tramite tomografia a raggi X per leggere il contenuto di documenti antichi senza aprirli, evitando così di danneggiarli ulteriormente.

La tomografia a raggi X permette infatti di poter ottenere un'immagine tridimensionale dell'intero documento - scritte comprese - senza la necessità di interventi potenzialmente invasivi/dannosi per l'oggetto. La lettura delle scritte è resa possibile dalla differenza di contrasto radiografico fra i materiali poco radio-opachi dei supporti - come carta e/o pergamena - e l'elevato assorbimento X degli inchiostri antichi a base ferro - utilizzati per la maggior parte dei manoscritti europei e per svariati secoli.

Per poter però realmente leggere un manoscritto a più pagine è necessario poterle estrarre dal volume e visualizzarle come immagini bidimensionali.

In questa tesi verrà presentata l'analisi e il test di un innovativo software di segmentazione per l'estrazione semi-automatica delle pagine di una scrittura risalente al 1679 ripiegata su se stessa al fine di simulare un documento composto da più pagine.

Attraverso il software sviluppato presso il Laboratorio di Tomografia a Raggi X di Bologna, le immagini tomografiche sono state sottoposte ad una prima fase di *pre-processing* per minimizzare il rumore ed aumentare così il contrasto fra il segnale delle pagine e quello del *background*, ad una seconda fase di identificazione automatica delle pagine all'interno del volume 3D ed infine all'estrazione delle pagine sotto forma di immagini.

# Indice

<b>Indice</b>	<b>3</b>
<b>Introduzione</b>	<b>4</b>
<b>1 La tomografia computerizzata con raggi X</b>	<b>6</b>
1.1 I raggi X . . . . .	6
1.2 Interazione dei raggi X con la materia . . . . .	9
1.3 I Principi alla base della tecnica radiografica . . . . .	12
1.4 Sistema tomografico . . . . .	13
1.5 La tecnica tomografica . . . . .	14
<b>2 Algoritmo di elaborazione dei dati</b>	<b>19</b>
2.1 Preparazione dei dati . . . . .	22
2.2 Caratteristiche geometriche delle pagine . . . . .	24
2.3 Riempimento delle pagine . . . . .	26
2.4 Separazione delle pagine . . . . .	27
2.5 Propagazione . . . . .	29
2.6 Estrazione delle pagine separate . . . . .	30
<b>3 Risultati</b>	<b>31</b>
<b>Conclusione</b>	<b>36</b>
<b>Bibliografia</b>	<b>37</b>

# Introduzione

Lo scopo principale di questo lavoro di tesi è quello di studiare una procedura non invasiva per leggere le pagine di documenti chiusi, come alcuni testi antichi, i quali, data la loro età, risultano troppo fragili e degradati per essere aperti o sfogliati. A tal fine, è stata analizzata una scrittura risalente al 1679 in buone condizioni di conservazione, che è stata ripiegata più volte così da simulare un documento che contenga più pagine.

L'elevata penetrazione dei raggi X permette di acquisire immagini tomografiche di un documento con il contrasto necessario alla lettura del contenuto grazie alla composizione chimica dell'inchiostro che per la maggior parte dei manoscritti europei è a base di ferro. L'inchiostro infatti assorbe maggiormente i raggi X rispetto al supporto del documento, permettendo la distinzione fra scrittura e pagina.

La composizione chimica dell'inchiostro dei documenti è un'informazione cruciale per quanto riguarda l'assorbimento dei raggi X, ed esistono diversi studi sulla composizione generale degli inchiostri usati nei manoscritti in Europa [1, 2]. In questo campo di ricerca vale la pena citare la ricostruzione tomografica, impiegando luce di sincrotrone, di un papiro di Ercolano completamente carbonizzato e quindi impossibile da aprire [3]. Un altro esempio interessante è la ricostruzione virtuale completa di antico rotolo di pergamena rinvenuto a En-Ghedi, anch'esso completamente carbonizzato, utilizzando la micro-tomografia con strumentazione da laboratorio - e quindi senza la necessità di dover spostare il prezioso oggetto presso grandi facilities [4].

Rispetto alle tecniche più innovative ancora in fase di sviluppo, la digitalizzazione comune presenta molteplici problemi legati principalmente alla fragilità dei documenti e ai tempi di acquisizione dei dati. I documenti che vengono digitalizzati devono poter essere maneggiati a lungo poiché ogni pagina viene scansionata singolarmente: questo, oltre all'enorme quantità di tempo necessaria, rischia di danneggiare i documenti più fragili in modo irreparabile.

Il primo capitolo contiene una breve descrizione della fisica dei raggi X e i fondamenti dell'acquisizione e della ricostruzione tomografica.

Il secondo capitolo tratta l'algoritmo di elaborazione delle immagini digitali, composto da tre fasi: diminuzione del segnale relativo al fondo con conseguente riduzione del rumore presente nelle immagini (segmentazione), identificazione e separazione delle pagine (classificazione) ed, infine, estrazione delle pagine sotto forma di immagini (estrazione).

Per finire, nel terzo capitolo vengono analizzati i risultati ottenuti attraverso le ricostruzioni tomografiche e le elaborazioni effettuate con il software.

# Capitolo 1

## La tomografia computerizzata con raggi X

Nel presente capitolo si illustra cosa sono e come vengono prodotti i raggi X, gli strumenti impiegati, l'interazione che essi hanno con la materia e i processi di ricostruzione tomografica.

### 1.1 I raggi X

I raggi X sono radiazioni elettromagnetiche con lunghezze d'onda comprese tra circa  $10^{-8}$  e  $10^{-12}$  m (Figura 1.1). Queste radiazioni possono essere emesse da qualsiasi elemento chimico colpito da un fascio di elettroni ad alta energia e quindi possono essere prodotte artificialmente bombardando con elettroni accelerati un *target*.

La capacità di penetrazione dei raggi X dipende dalla loro lunghezza d'onda: i raggi X caratterizzati da lunghezze d'onda maggiori, ossia più vicini alla banda ultravioletta dello spettro elettromagnetico, sono comunemente chiamati "mollì" mentre quelli con lunghezze d'onda minori sono detti "duri".

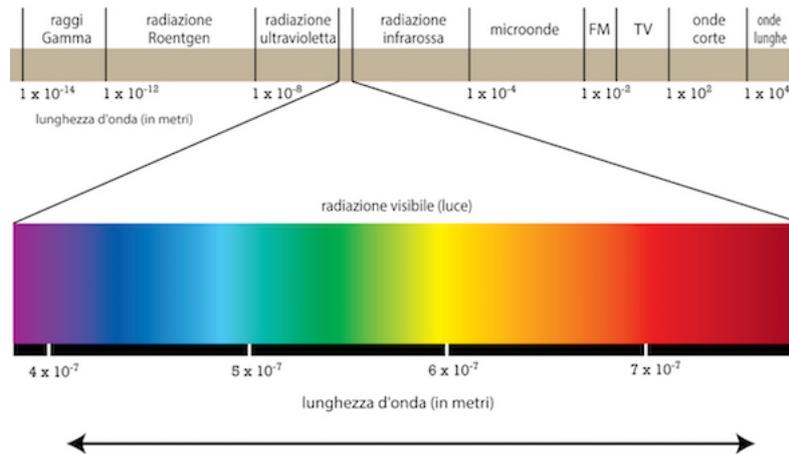


Figura 1.1: Spettro elettromagnetico.

Uno dei metodi più utilizzati nella produzione dei fasci di raggi X è l'utilizzo del "tubo radiogeno" (Figura 1.2), anche chiamato "tubo a raggi X". Questo strumento è formato da un'ampolla di vetro in cui viene fatto un vuoto spinto, contenente due elettrodi disposti frontalmente: un catodo (elettrodo negativo) e un anodo (elettrodo positivo). Sulla superficie del catodo vi sono uno o due filamenti metallici, i quali vengono alimentati a bassa tensione, ma ad alto amperaggio, in modo da provocare l'emissione dal catodo di una nuvola di elettroni per effetto termoionico. A causa della differenza di potenziale elevata tra i due elettrodi, la nuvola di elettroni verrà accelerata verso l'anodo, formato da un disco di materiale ad alto numero atomico, acquistando così energia cinetica.

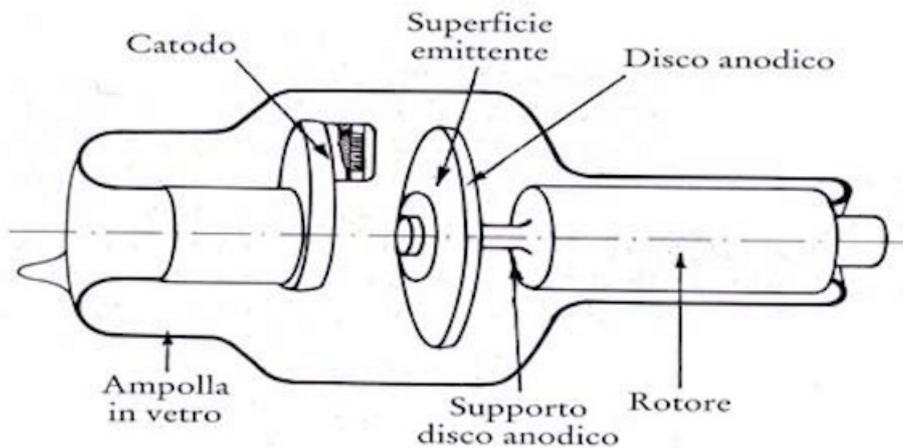


Figura 1.2: Tubo a raggi X.

L'impatto degli elettroni con l'anodo causa una loro decelerazione, con conseguente emissione di radiazioni, le quali sono prodotte attraverso due processi differenti di interazione (Figura 1.3):

- gli elettroni vengono deviati in seguito all'interazione con il nucleo, perdendo così energia cinetica, emessa sotto forma di radiazione X, detta anche radiazione di frenamento o di *Bremsstrahlung*;
- gli elettroni del fascio interagiscono con uno degli elettroni più interni, allontanandolo dalla sua orbita, creando così un atomo in uno stato eccitato o ionizzato. Saranno poi gli elettroni esterni ad occupare la lacuna creata precedentemente ed è proprio durante questa transizione che l'elettrone interessato libera energia sotto forma di un fotone, la cui energia corrisponde esattamente alla differenza tra l'energia dello stato eccitato e quella dello stato fondamentale (radiazione caratteristica).

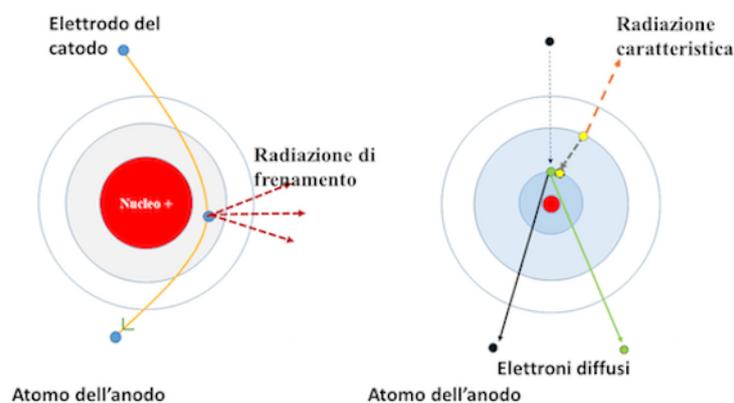


Figura 1.3: Rappresentazione dell'emissione delle due tipologie di radiazioni in seguito all'interazione elettrone-atomo anodico: radiazione di frenamento (a sinistra), radiazione caratteristica (a destra).

Lo spettro di un tubo a raggi X (Figura 1.4) è costituito perciò dalla somma delle due componenti: una parte continua, derivante dalla radiazione di *Bremsstrahlung*, la cui energia massima dipende dalla tensione applicata fra anodo e catodo, ed i picchi caratteristici la cui energia è peculiare del materiale anodico. Sperimentalmente sappiamo che in un tubo a raggi X solo una piccola percentuale dell'energia cinetica degli elettroni viene trasformata in radiazione utilizzabile, mentre il restante viene convertito in calore.

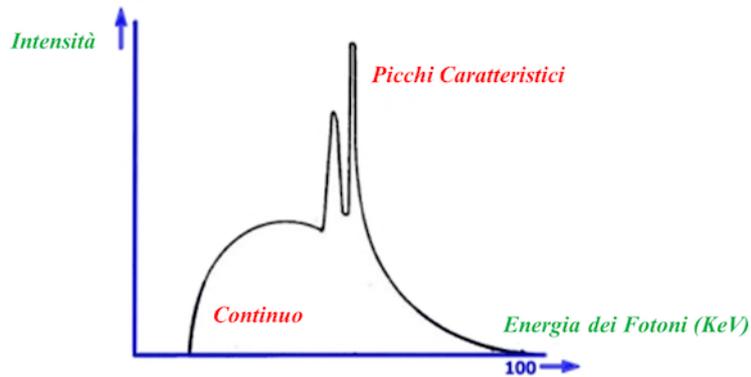


Figura 1.4: Spettro di un tubo a raggi composto da una parte continua - Bremsstrahlung - e dai picchi caratteristici.

## 1.2 Interazione dei raggi X con la materia

L'interazione dei raggi X con la materia può avvenire attraverso tre tipi di processi: effetto Compton (Figura 1.5), effetto fotoelettrico (Figura 1.6) e produzione di coppia (Figura 1.7).

- Effetto Compton: consiste nell'urto anelastico tra il fotone e uno degli elettroni presenti negli orbitali esterni, dove il primo cambia direzione e cede parte della sua energia al secondo. Considerando le leggi di conservazione dell'energia e della quantità di moto, si può notare che l'energia finale del fotone  $E_f$  è in relazione all'energia iniziale  $E_i$  e all'angolo di scattering  $\theta$ :

$$E_f = \frac{E_i}{1 + \gamma(1 - \cos \theta)} \quad (1.1)$$

dove

$$\gamma = \frac{h\nu}{mc^2} \quad (1.2)$$

mentre l'energia cinetica dell'elettrone  $E_{e^-}$  :

$$E_{e^-} = E_i \cdot \frac{\gamma(1 - \cos \theta)}{1 + \gamma(1 - \cos \theta)} \quad (1.3)$$

Alle basse energie la distribuzione angolare dei fotoni diffusi è abbastanza isotropa, ma più l'energia cresce più gli angoli di scattering sono piccoli. In particolare, l'effetto Compton è tanto più probabile quanto maggiore è l'energia associata al fascio incidente.

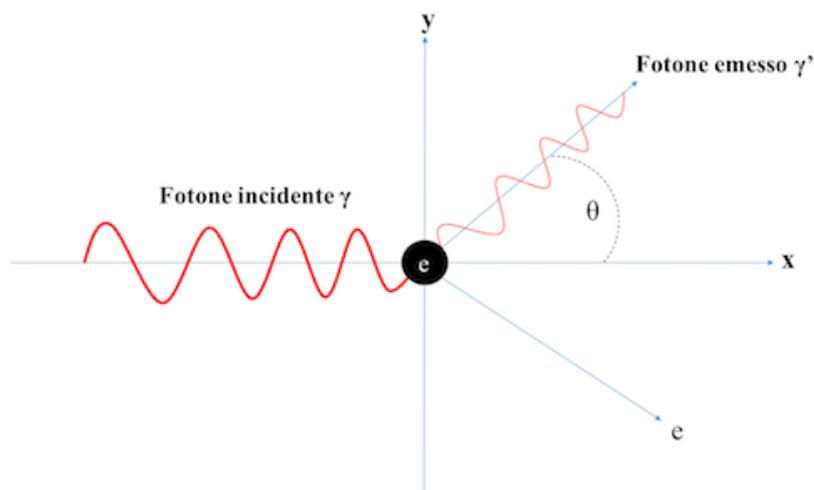


Figura 1.5: Effetto Compton.

- Effetto fotoelettrico: avviene quando un fotone di medio-bassa energia interagisce con un elettrone fortemente legato all'atomo del materiale. In tale interazione il fotone cede tutta la sua energia all'elettrone con cui interagisce e viene assorbito, determinando l'emissione di un fotoelettrone con un'energia cinetica pari alla differenza tra l'energia del fotone e l'energia di legame dell'elettrone;

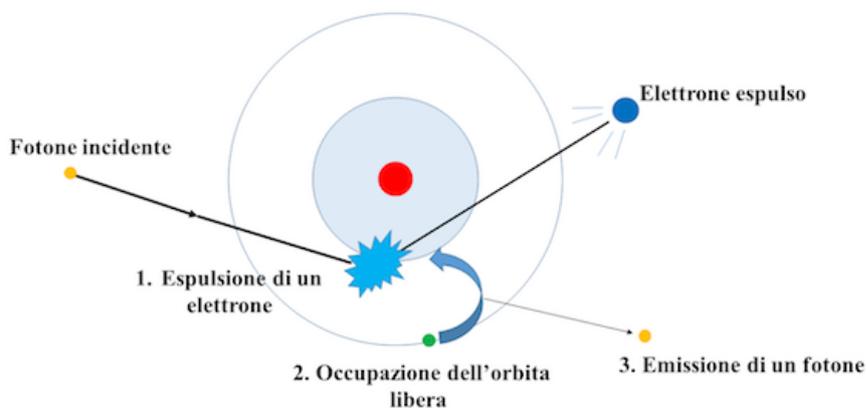


Figura 1.6: Effetto fotoelettrico.

- Produzione di coppia: è un processo in cui il fotone interagendo con la materia converte la sua energia in una coppia elettrone-positrone. Tale interazione può avere luogo quando l'energia del fotone è superiore a 1.022 MeV, in quanto pari alla somma delle masse delle particelle create. L'energia del fotone viene suddivisa tra la massa in quiete della coppia che si forma e l'energia cinetica delle due particelle, dove però il positrone tende a riceverne una maggiore quantità. Essendo generalmente basse le energie in gioco nell'analisi tomografica, questo effetto ha una bassa probabilità di verificarsi.

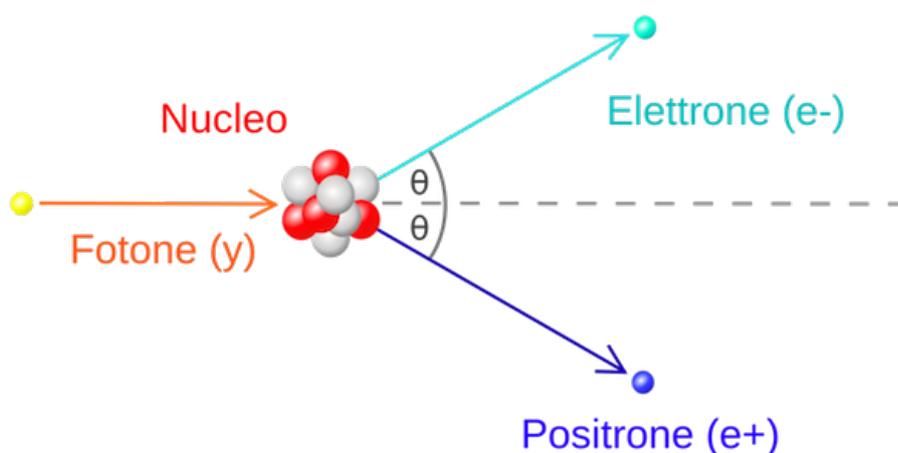


Figura 1.7: Produzione di coppie.

La probabilità che questi tre tipi di interazione possano verificarsi, ossia la sezione d'urto legata ad ognuno di essi, è strettamente connessa al numero atomico dell'elemento costituente il materiale e all'energia dei fotoni incidenti. La sezione d'urto totale nel caso in cui avvenga più di un processo è data dalla somma delle singole sezioni d'urto. L'assorbimento dei raggi X da parte dei vari materiali dipende dalla probabilità di interazione dei tre processi precedentemente esaminati (Figura 1.8) ed è questo che determina la radio-opacità di un materiale a determinate energie. L'assorbimento è introdotto attraverso il coefficiente di attenuazione  $\mu$ , dato dalla somma dei coefficienti di attenuazione delle singole modalità di interazione:

$$\mu = \mu_{\text{Compton}} + \mu_{\text{fotoel}} + \mu_{\text{coppie}} \quad (1.4)$$

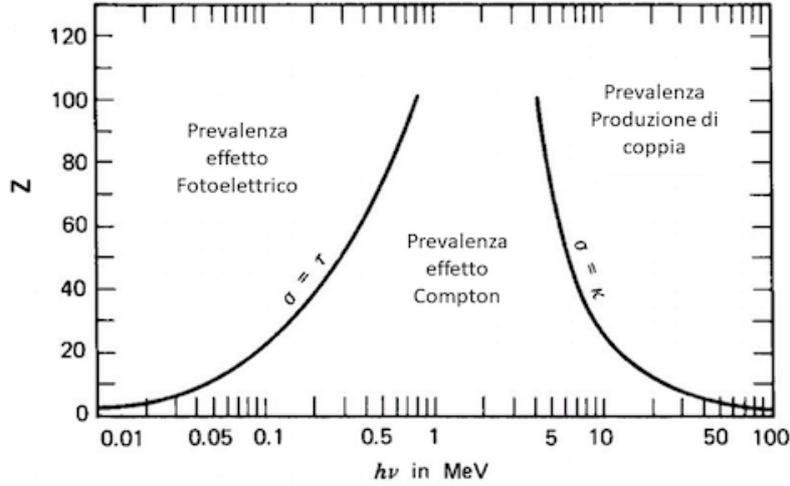


Figura 1.8: Prevalenza dei tre tipi di interazione dei fotoni X con la materia in funzione dell'energia e del numero atomico  $Z$  del materiale. La curva identificata come  $\sigma = \tau$  (sulla sinistra) indica l'equivalenza fra le sezioni d'urto dell'effetto fotoelettrico e dell'effetto Compton. Allo stesso modo, la curva identificata come  $\sigma = \kappa$  (sulla destra) indica l'equivalenza delle sezioni d'urto dell'effetto Compton e della produzione di coppia.

### 1.3 I Principi alla base della tecnica radiografica

La radiografia è una delle tecniche più adatte per analizzare la struttura interna di un oggetto senza danneggiarlo, ottenendo così un'immagine bidimensionale. L'immagine radiografica è generata dall'interazione tra il flusso di raggi X e l'oggetto considerato, nello specifico è ottenuta grazie al contrasto tra l'assorbimento e la trasmissione dei raggi X attraverso l'oggetto. L'assorbimento dei raggi X viene regolato dalla legge di Lambert-Beer, che, nel caso di un fascio monocromatico di raggi X che attraversa un oggetto omogeneo e di densità costante, assume la forma:

$$I = I_0 e^{-\mu S} \quad (1.5)$$

dove  $I_0$  è l'intensità del fascio iniziale senza l'oggetto. L'intensità del fascio  $I$  dipenderà, una volta fissata l'energia, dal coefficiente di attenuazione lineare  $\mu$  e dallo spessore dell'oggetto attraversato  $S$ .

Nel caso in cui l'oggetto non sia omogeneo la legge di attenuazione diventa:

$$I(x, y) = I_0(x, y) e^{-\int_L \mu(x, y, l) dl} \quad (1.6)$$

dove  $l$  è la direzione dei raggi X.

## 1.4 Sistema tomografico

Dal momento che la radiografia è una proiezione bidimensionale di un oggetto tridimensionale, essa non permette di dare un'interpretazione sulla struttura interna dell'oggetto. Una tecnica diagnostica non invasiva in grado di visualizzare in 3D la struttura interna dell'oggetto in esame è la tomografia computerizzata con raggi X, la quale si basa sull'acquisizione di più radiografie dell'oggetto analizzato, a diverse angolazioni, e nella sua successiva ricostruzione tridimensionale attraverso algoritmi matematici.

I principali componenti di un sistema tomografico sono:

- la sorgente di raggi X, generalmente un tubo radiogeno, il quale è il più utilizzato per le applicazioni nei beni culturali;
- un sistema di movimentazione per l'oggetto, rappresentato da una base rotativa, oppure, in alternativa, un sistema di movimentazione in grado di ruotare la sorgente e il rivelatore, lasciando quindi l'oggetto in esame immobile, come nel caso medico;
- un rivelatore digitale, capace di registrare la radiazione che non è stata assorbita dall'oggetto.

In base alla sorgente di raggi X, al rivelatore e alla distanza tra di essi si possono ottenere setup tomografici con geometrie diverse:

- Tomografia con fascio parallelo (*Parallel Beam*), dove la sorgente ed il rivelatore sono lineari;

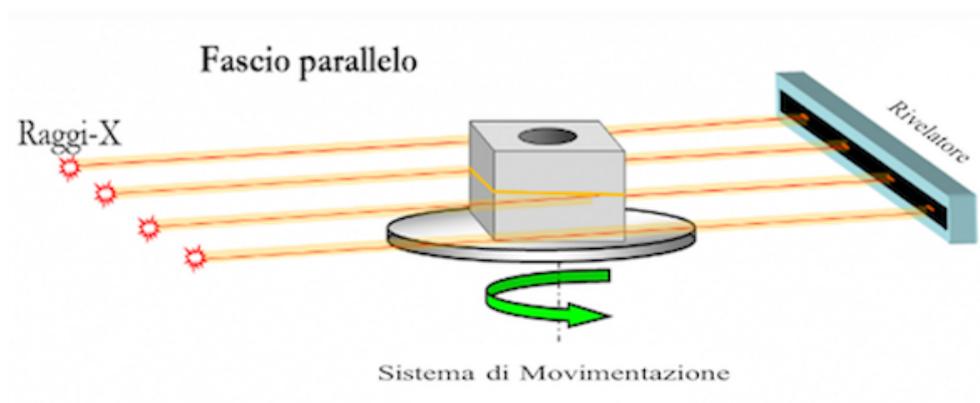


Figura 1.9: Tomografia con fascio parallelo, dove sorgente e rivelatore sono lineari.

- Tomografia con fascio a ventaglio (*Fan Beam*), dove la sorgente ha un fascio collimato ed il rivelatore è composto da una serie di rivelatori in linea;

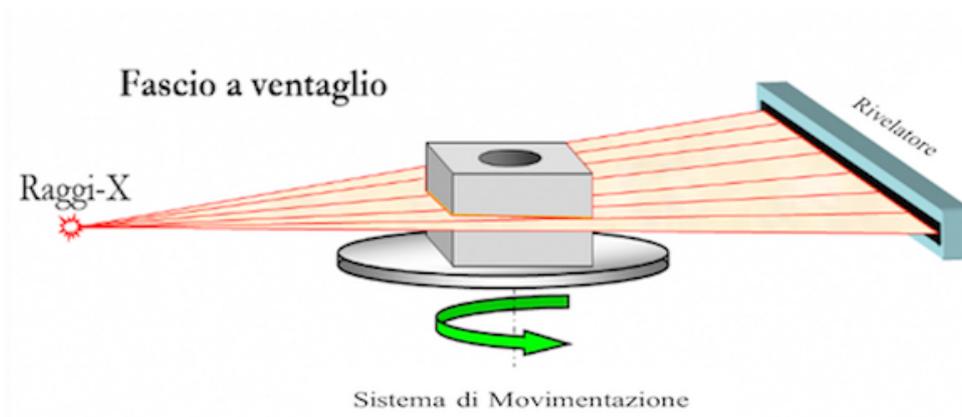


Figura 1.10: Tomografia con fascio a ventaglio, dove la sorgente ha un fascio collimato e il rivelatore è composto da una serie di rivelatori in linea.

- Tomografia con fascio conico (*Cone Beam*) dove la sorgente, non collimata, ha fascio conico per irradiare un rivelatore bidimensionale.

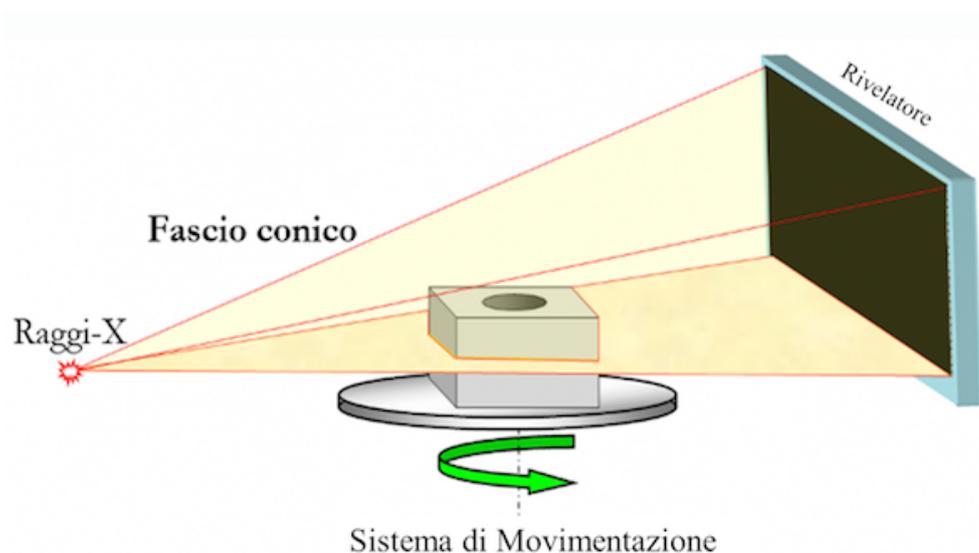


Figura 1.11: Tomografia con fascio conico, dove la sorgente, non collimata ha un fascio conico e un rivelatore bidimensionale.

## 1.5 La tecnica tomografica

Una volta terminata la scansione a diversi angoli, inizia la fase di ricostruzione vera e propria, alla fine della quale, attraverso l'elaborazione delle immagini, si ottengono le sezioni interne dell'oggetto, anche dette *slices*.

Lavorando sulle immagini tomografiche ottenute grazie alla scansione, si considera il

caso ideale in cui un fascio di raggi X paralleli tra loro passi attraverso un solo piano dell'oggetto in esame, terminando sul rivelatore suddiviso in  $n$  elementi.

L'immagine digitale che vogliamo ricostruire è rappresentata da una griglia di  $n \times n$  pixel, ciascuno con il proprio coefficiente di attenuazione  $\mu_{ij}$  e dimensione  $w \times w$  pari a quelle di un singolo elemento che costituisce il rivelatore (Figura 1.12).

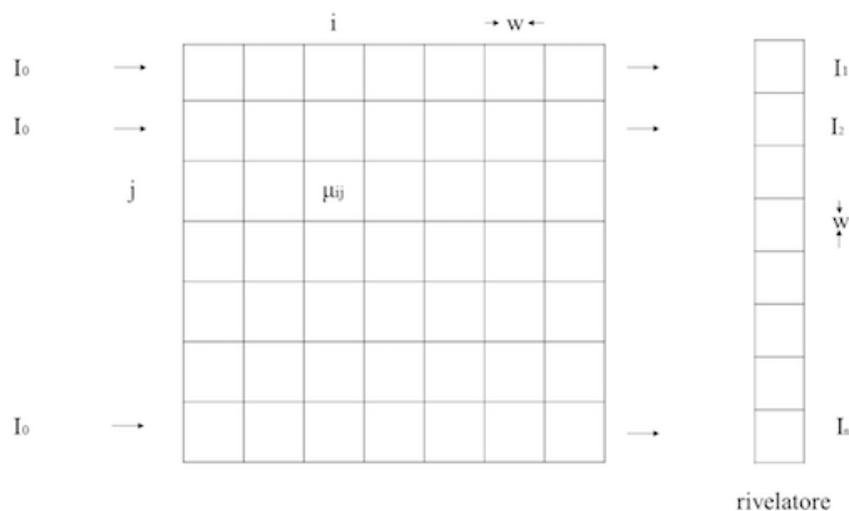


Figura 1.12: Schema della slice suddivisa in pixel con indicazione dei coefficienti di attenuazione per ogni pixel e della relativa intensità che raggiunge il rivelatore (a destra).

Allora l'intensità del fascio misurata da ogni elemento del rivelatore sarà data da:

$$I_j = I_0 \cdot e^{-\mu_{1j} \cdot w - \mu_{2j} \cdot w - \dots - \mu_{nj} \cdot w} \quad (1.7)$$

dove  $j=1,2,\dots,n$

da cui risulta quindi:

$$\ln \frac{I_0}{I_j} = \sum_{i=1}^n \mu_{ij} \cdot w \quad (1.8)$$

La ricostruzione nella geometria *Parallel Beam* prevede che i raggi X siano paralleli. Sia  $f(x,y)$  l'oggetto che vogliamo ricostruire, allora l'ultima equazione rappresenta la proiezione  $p_j$  ed è la radiografia di  $f(x,y)$ . Ogni proiezione sarà quindi unidimensionale ed

esisteranno  $T$  proiezioni ciascuna per un diverso angolo  $\theta$ . La ricostruzione è possibile grazie al teorema della *slice* di Fourier (Figura 1.13) il quale afferma che applicando la trasformata di Fourier unidimensionale alla proiezione lineare  $p_j$  ottenuta ad un certo angolo  $\theta$ , si ottiene una retta  $P(u, v)$  nello spazio di Fourier passante per l'origine con un'inclinazione pari a  $\theta$ . La retta  $P(u, v)$  nello spazio delle frequenze rappresenta una porzione dei punti della  $F(u, v)$  che è la trasformata 2D della funzione  $f(x, y)$  che è proprio la *slice* che stiamo cercando di ricostruire.

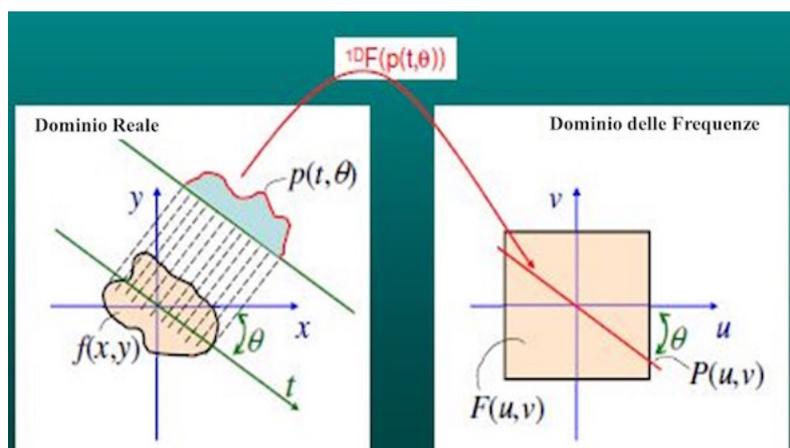


Figura 1.13: Applicazione della trasformata di Fourier 1D alla proiezione lineare  $p_j$ , ottenuta ad un certo angolo  $\theta$  (sinistra), da cui si ottiene una retta  $P(u, v)$  passante per l'origine e inclinata sempre di  $\theta$ , nello spazio delle frequenze (destra).

Calcolando tutte le proiezioni a tutti i possibili angoli  $\theta$  si può ricoprire quindi la zona dello spazio di Fourier che rappresenta la funzione  $F(u, v)$ . Effettuando l'anti-trasformata di questi punti si ottiene la  $f(x, y)$ , sezione ricostruita dell'oggetto.

Purtroppo i punti noti della  $F(u, v)$  non sono omogeneamente distribuiti nello spazio di Fourier (Figura 1.14): infatti la  $F(u, v)$  è densamente popolata al centro, in corrispondenza delle basse frequenze, presentando una migliore definizione di questo tipo di segnale, mentre la  $F(u, v)$  all'esterno, in corrispondenza delle alte frequenze, bordi dell'immagine, è poco definita. Quindi ci si aspetta che la ricostruzione con l'anti-trasformata 2D della  $F(u, v)$  così com'è restituisca come risultato un'immagine del tutto sfuocata.

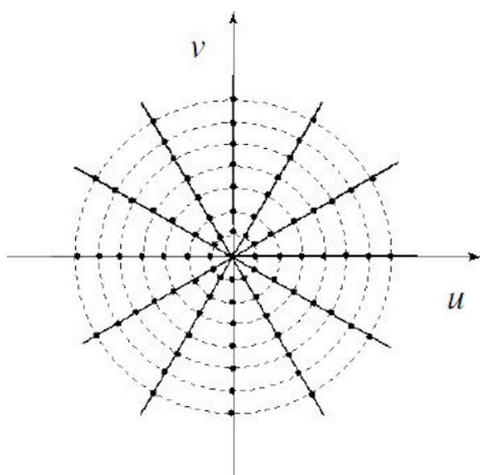


Figura 1.14: Rappresentazione del riempimento dello spazio delle frequenze con i punti delle rette corrispondenti alle proiezioni prese ad angoli diversi. I punti più vicini al centro dell'immagine sono ben definiti, mentre i punti più esterni (i bordi dell'immagine) sono sfuocati.

Per ovviare a questo problema si utilizza il filtro, il quale associa ai punti centrali un valore più basso e ai punti esterni un valore più alto. Non applicando il filtro si otterrebbe un'immagine ricostruita correttamente ma sfuocata, mentre, applicando il filtro, si ottiene un'immagine ben ricostruita e con i bordi definiti (Figura 1.15).

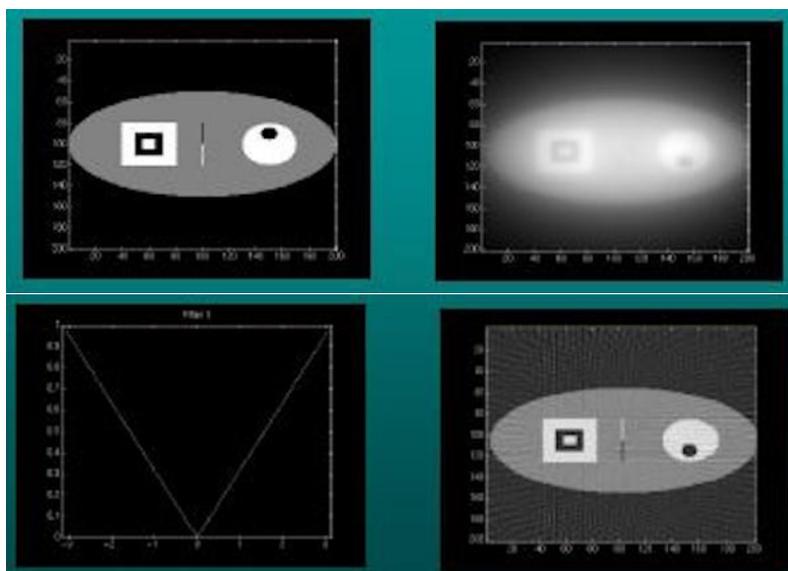


Figura 1.15: Rappresentazione dell'applicazione del filtro: in alto, non viene applicato il filtro e l'immagine risulta sfuocata, mentre, in basso, il filtro viene applicato e i bordi dell'immagine sono ben definiti.

L'algoritmo di ricostruzione tomografico, chiamato *Filtered Back Projection (FBP)*, assume la seguente forma:

$$f(x, y) = \int_0^\pi d\theta \int_{-\infty}^{+\infty} P(\omega, \theta) |\omega| e^{j2\pi\omega \cdot t} dt$$

dove  $P(\omega, \theta)$  rappresenta la trasformata di Fourier unidimensionale della proiezione all'angolo  $\theta$ ,  $|\omega|$  è il filtro nello spazio di Fourier, l'integrale in  $dt$  è l'anti-trasformata 1D della proiezione filtrata e l'integrale in  $d\theta$  la somma su tutti gli angoli considerati.

La trasformata 1D di Fourier viene moltiplicata per il filtro e, successivamente, viene eseguita l'anti-trasformata 1D della proiezione filtrata, eseguendo infine una sommatoria su tutti gli angoli considerati.

La ricostruzione tridimensionale per raggi non paralleli ha una formulazione simile, attraverso l'impiego di alcuni fattori correttivi che tengono in considerazione la diversa geometria e permettono la ricostruzione per il *Fan Beam*[4]. Esiste poi un'applicazione in tre dimensioni più generale, data dall'algoritmo di *Feldkamp-Davis-Kress (FDK)*[4].

## Capitolo 2

# Algoritmo di elaborazione dei dati

Molti testi antichi, a causa della fragilità dei supporti e di fattori esterni, sono deteriorati e dunque molto fragili, motivo per cui aprirli o anche solo sfogliarne le pagine rischia di danneggiarli in modo irreparabile. Per evitare di arrecare danni a questi documenti, si cercano tecniche non invasive per digitalizzarli che, come la tomografia computerizzata, permettano di poterne leggere il contenuto senza aprirli o maneggiarli.

Affinché le scritte siano leggibili utilizzando i raggi X, l'inchiostro deve permettere un contrasto radiografico così che il testo sia distinguibile dai supporti come carta, papiro o pergamena. Questo è possibile per inchiostri che presentano una componente ad alto numero atomico come gli inchiostri ferro-gallici, largamente utilizzati in tutta Europa e per numerosi secoli. Per individuare le scritte, però, bisogna prima individuare le pagine dalle immagini digitali ottenute.

Il processo di individuazione ed estrazione delle pagine è reso più complicato da due fattori: da un lato -la separazione fra una pagina e l'altra è di soli pochi pixel - e in alcuni punti queste entrano in contatto - dall'altro - le pagine non sono piane e parallele. Presso l' *X-Ray Imaging Group*, del Dipartimento di Fisica e Astronomia dell'Università di Bologna, sono stati sviluppati diversi algoritmi *ad hoc* per l'estrazione delle pagine e per risolvere le principali criticità di questo processo. Nel dettaglio, si è deciso un approccio in tre fasi: la preparazione del set di dati tomografici per eliminare il segnale relativo al *background*; la classificazione, in cui si cerca di identificare e separare ogni singola pagina ed, infine, l'estrazione delle pagine stesse .

In questo progetto di tesi per testare il corretto funzionamento dell'algoritmo è stata presa in esame una scrittura privata risalente al 1679 (Figura 2.1), aperta e in buono stato. La scansione tomografica è stata eseguita sul documento ripiegato più volte su se stesso al fine di renderlo simile a un testo contenente più pagine.

in Londra li 29 giugno 1679  
 Rodolfo ~~...~~ ~~...~~  
 del ~~...~~ ~~...~~  
~~...~~ ~~...~~  
 Londra li 29 giugno 1676 & 1677  
 Condi \_\_\_\_\_  
 Giovanni ~~...~~ 1677 \_\_\_\_\_ 3  
 Matteo ~~...~~ 1677 \_\_\_\_\_ 1  
 Per il ~~...~~ 1677 \_\_\_\_\_ 31  
 In ~~...~~ ~~...~~ \_\_\_\_\_ 7  
 In tutto \_\_\_\_\_ 94  
 Biggino ~~...~~

Figura 2.1: Scrittura privata risalente al 1679.

Ciò che si è ottenuto dalla ricostruzione tomografica, effettuata in Svizzera presso EP-FL - Ecole Polytechnique Fédérale de Lausanne con hardware e software per tomografia veterinaria, sono le varie *slices* del documento, le quali possono essere visualizzate o da una prospettiva parallela al documento (Figura 2.2) o perpendicolare ad esso (Figura 2.3).

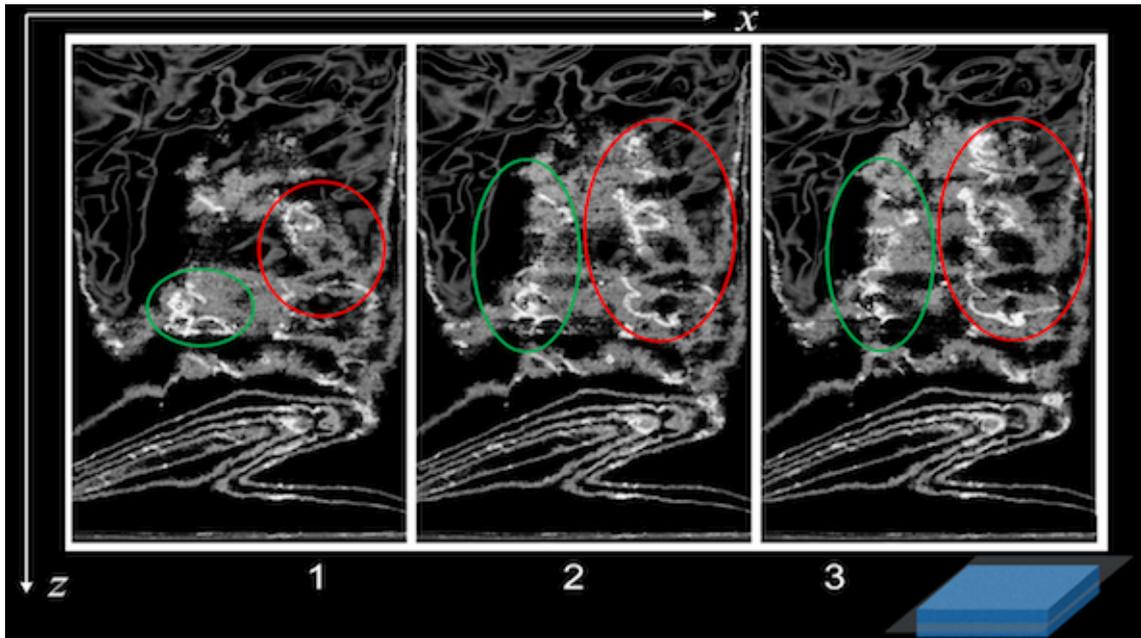


Figura 2.2: Visualizzazione delle scritte da una prospettiva parallela al piano delle pagine. Si noti come una stessa pagina sia presente in tre slices differenti (1, 2 e 3).

La prospettiva parallela al documento non è utile allo scopo in quanto scritte appartenenti alla stessa pagina si possono ritrovare su piani diversi (ovvero *slices*), motivo per cui si preferisce utilizzare la prospettiva perpendicolare alle sezioni. Le immagini sono state analizzate e processate utilizzando algoritmi appositamente sviluppati e contenuti nel software sviluppato dall' *X-Ray Imaging Group*.

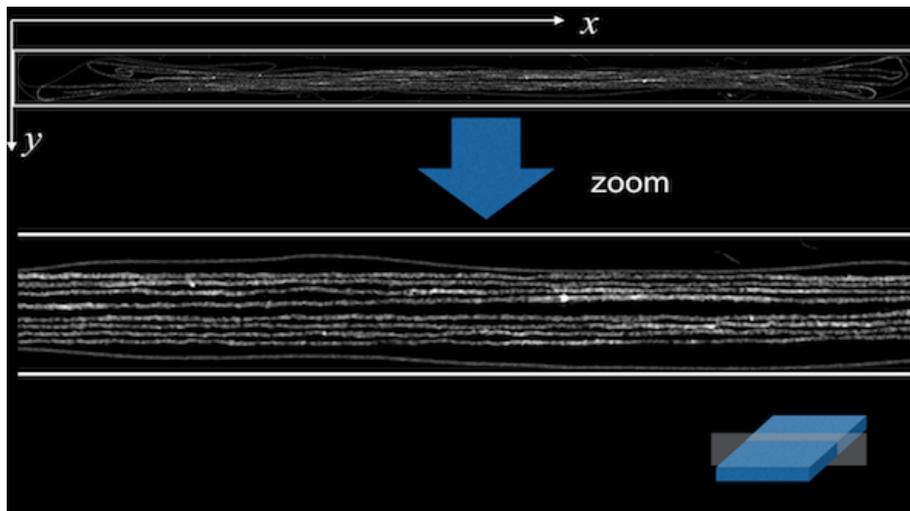


Figura 2.3: Rappresentazione della visualizzazione delle pagine attraverso una prospettiva perpendicolare alla direzione delle pagine.

## 2.1 Preparazione dei dati

Al fine di evidenziare le pagine rispetto al fondo, diminuendo il segnale del fondo e il rumore presente nelle immagini digitali (Figura 2.4) è stato impiegato un algoritmo di segmentazione che si basa su sistemi di filtrazione delle immagini. L'algoritmo si divide in due passi: filtrazione dell'immagine e analisi dell'immagine filtrata al fine di separare al meglio le pagine dal fondo.



Figura 2.4: Immagine non filtrata con evidente presenza di rumore.

I filtri svolgono operazioni differenti in base al *Kernel* scelto, il quale rappresenta una matrice di dimensione dispari che viene traslata su tutta la matrice rappresentante l'immagine digitale (Figura 2.5) operando una convoluzione. Il *Kernel* viene traslato sulla matrice relativa all'immagine digitale in modo da analizzare tutti i pixel e attribuire ad essi un valore differente in base alle caratteristiche del filtro. I valori che vengono attribuiti ai pixel dell'immagine sono calcolati come somma di tutti i prodotti ottenuti moltiplicando i coefficienti della prima matrice per i corrispondenti valori della seconda matrice (*Kernel*), il tutto diviso eventualmente da un fattore di normalizzazione:

$$O_{57} = I_{46}K_{11} + I_{47}K_{12} + I_{48}K_{13} + I_{56}K_{21} + I_{57}K_{22} + I_{58}K_{23} + I_{66}K_{31} + I_{67}K_{32} + I_{68}K_{33}$$

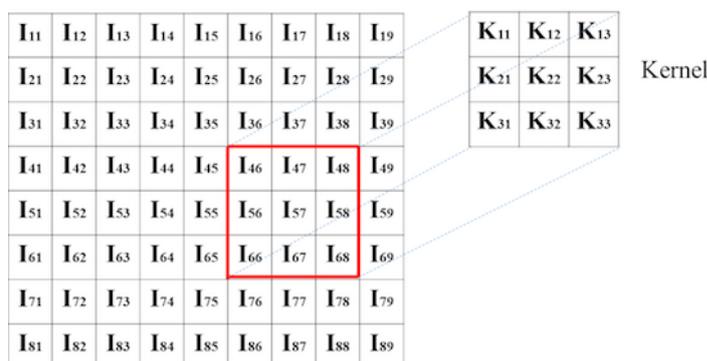


Figura 2.5: Il *Kernel* viene traslato su tutta l'immagine in modo da calcolare per ogni posizione un singolo valore associato al pixel centrale dell'immagine.

Nella prima fase dell'algoritmo i filtri scelti per il nostro problema sono stati il filtro *Maximum* (Massimo) e il *Median* (Mediano). Questi filtri hanno permesso di ottenere delle immagini in cui si riescano a distinguere le pagine dal fondo e in cui è stata ridotta al minimo la presenza del rumore (Figura 2.6).

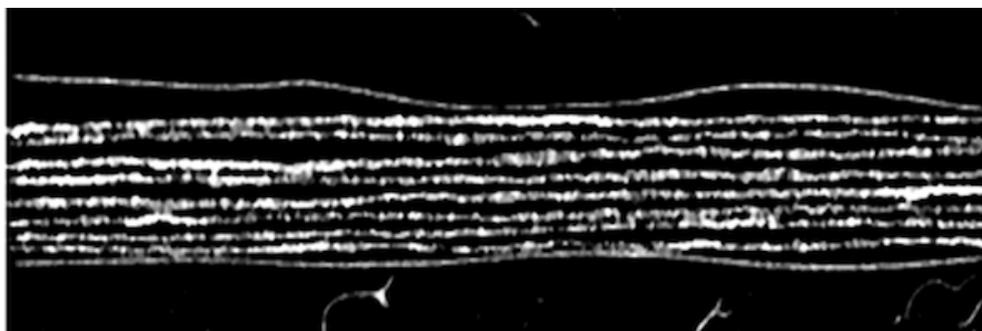


Figura 2.6: Immagine filtrata con una separazione migliore tra il segnale relativo alle pagine e il fondo con conseguente riduzione della presenza del rumore.

Nella seconda fase dell'algoritmo le immagini filtrate sono state impiegate per separare le zone dell'immagine che si vogliono conservare (segnale relativo alle pagine) dal fondo (rumore). A questo punto si calcola l'istogramma dell'immagine e si suddivide l'intero intervallo di livelli di grigio in cinque parti che possono contenere una determinata informazione del segnale messa in evidenza dal filtro utilizzato. Ad ogni parte dell'intervallo è stata, inoltre, associata una soglia modificabile così che si possa scegliere quale informazione del segnale evidenziare maggiormente. Un esempio di istogramma dell'immagine filtrata con le diverse parti in evidenza è in Figura 2.7.

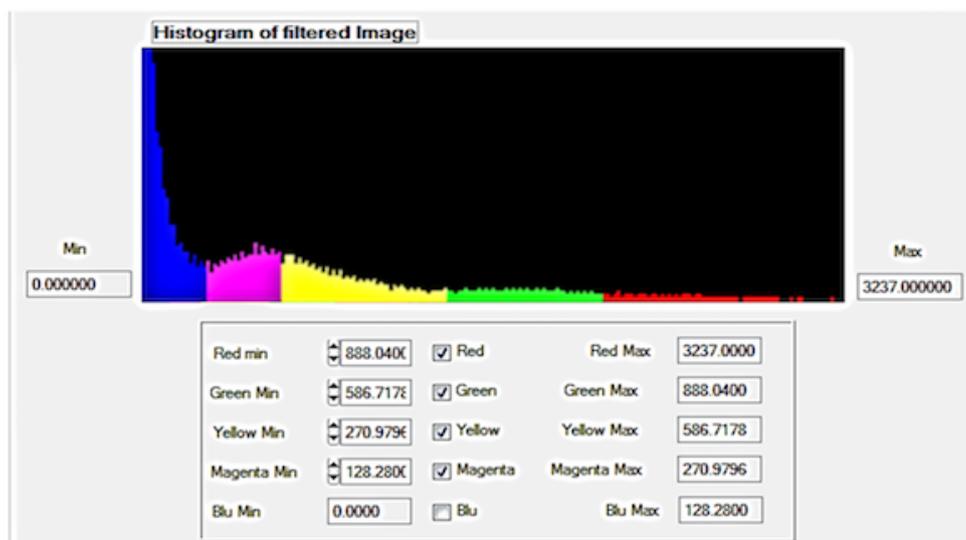


Figura 2.7: Istogramma dell'immagine filtrata, con le varie zone evidenziate da diversi colori, a diverse soglie.

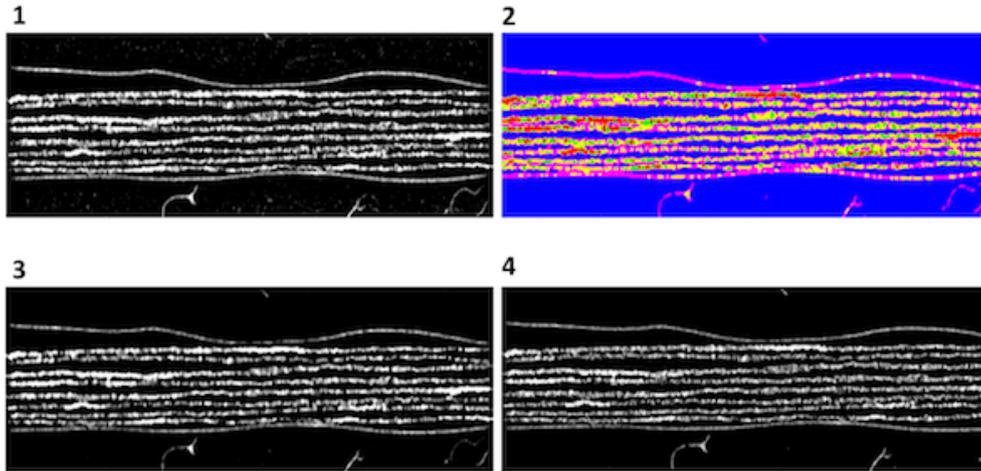


Figura 2.8: Preparazione dei dati: (1) immagine originale, (2) classificazione delle parti (in alto a destra), (3) immagine filtrata, (4) immagine finale con le parti evidenziate dal blu eliminate.

Ora che vi è una buona separazione tra ciò che è il segnale relativo alle pagine e il fondo, si possono studiare le caratteristiche delle pagine. Nonostante le pagine non siano perfettamente piane, esse presentano un andamento generalmente orizzontale, l'algoritmo prosegue quindi con la definizione delle caratteristiche geometriche.

## 2.2 Caratteristiche geometriche delle pagine

Per iniziare, viene stabilita una soglia di separazione tra ciò che è il segnale relativo alle pagine e ciò che invece è fondo, chiamata *threshold*.

Successivamente, si individua la distanza di ogni punto (pixel) dal bordo (l'ultimo punto che appartiene ancora al segnale della pagina) in tutte le direzioni possibili (Figura 2.9):

$X$  = distanza tra il punto e il bordo in tutte le direzioni

dove  $X$  = "alto", "basso", "destra", "sinistra", "diagonale Y" (con  $Y=1,2,3,4$ )

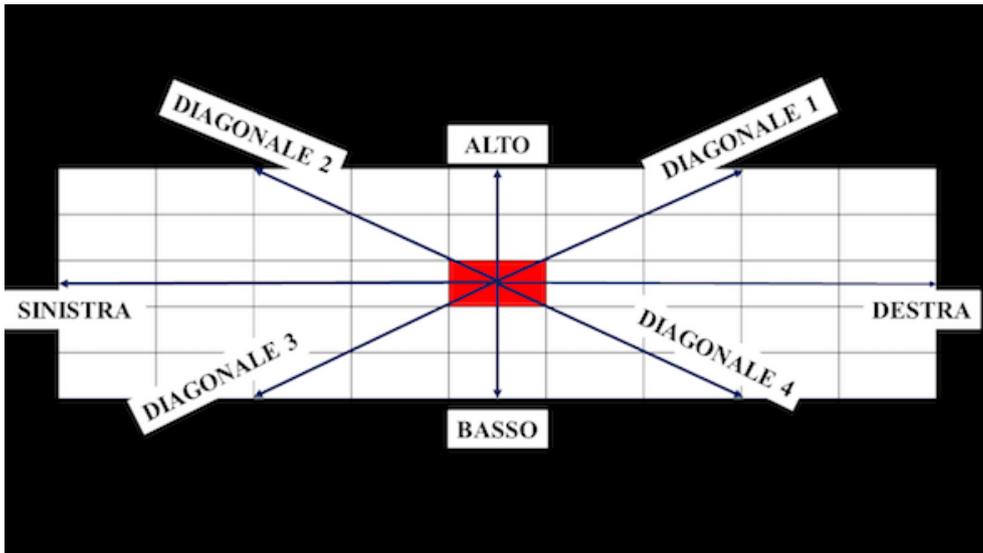
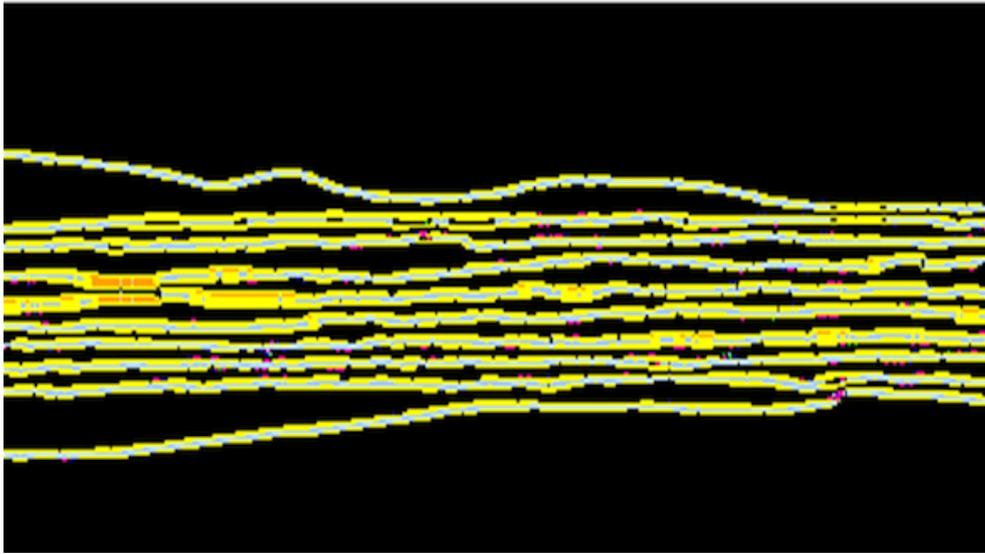


Figura 2.9: Distanza del pixel dal bordo delle pagine, in tutte le direzioni possibili.

Per individuare lo spessore della pagina, si considera lo spessore di dimensione maggiore, calcolato nel seguente modo:

- "orizzontale" = "destra" + "sinistra";
- "verticale" = "alto" + "basso";
- "diagonale positiva" = "diagonale 1" + "diagonale 3";
- "diagonale negativa" = "diagonale 2" + "diagonale 4".

Si osservi la Figura 2.10. Si vedono le varie aree dell'immagine digitale colorate in base alla direzione: giallo per l'orizzontale, ciano per la verticale, blu per la diagonale positiva, verde chiaro per la diagonale positiva e fucsia per quei punti che si trovano all'interno di aree con una direzione differente da tutti i suoi vicini.



*Figura 2.10: Pagine colorate in base alla direzione individuata: si osserva come la direzione orizzontale (giallo) è quella più diffusa.*

Visto che si studiano le pagine di un documento, vengono eliminati dalle immagini digitali tutti i pixel ai quali non sia stata attribuita la direzione orizzontale. A questo punto, vengono ricalcolati gli spessori di ogni area.

## 2.3 Riempimento delle pagine

Una volta ottenute le caratteristiche geometriche di ogni pixel, l'algoritmo prosegue con il riempimento dei pixel rimanenti, chiamato *filling*. Questo algoritmo semplicemente parte da un qualsiasi punto di segnale e classifica quel punto come zona 1. Poi ci si sposta in tutte le direzioni attorno a quel punto fino a riempire la zona 1 fermandosi solo ai bordi. L'algoritmo prosegue quindi cercando un nuovo pixel che sia segnale e che non appartenga alla zona 1, da qui si ripete l'operazione di riempimento e si definisce la zona 2. L'algoritmo prosegue fino alla definizione di tutte le zone nell'immagine (Figura 2.11). Ora ogni zona rappresenta un insieme di pixel che geometricamente è costituita da aree adiacenti.

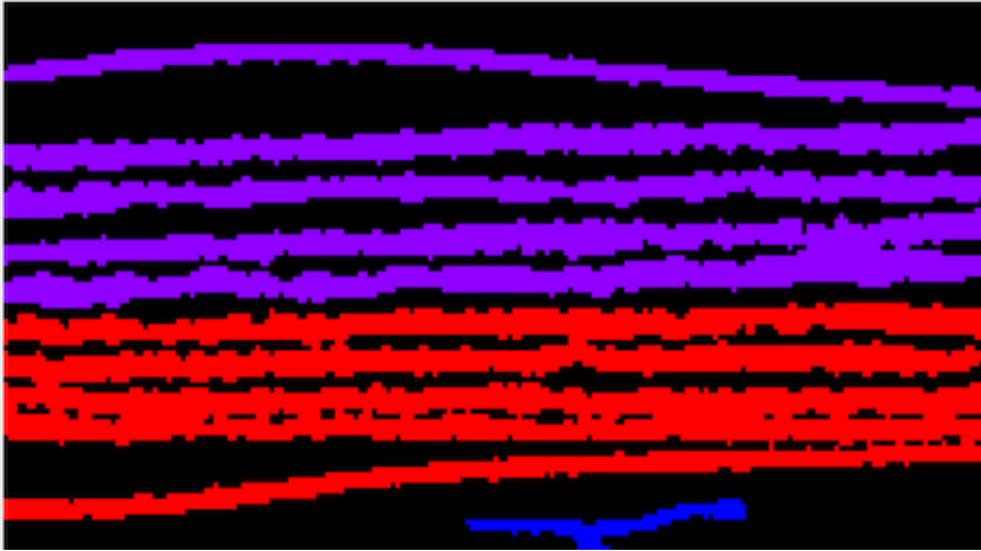


Figura 2.11: Riempimento delle zone del segnale relative alle pagine.

## 2.4 Separazione delle pagine

Uno dei problemi nella fase di riempimento delle zone riguarda quei punti in cui due pagine si toccano, ovvero i punti di contatto. Questi punti, infatti, appartengono a entrambe le pagine, motivo per cui l'algoritmo le riconosce come un'unica zona (Figura 2.12).

I punti di contatto sono caratterizzati dal fatto che si trovano in un punto tra le due zone tale che lo spessore in quei punti sia maggiore della media degli spessori delle pagine e tale per cui la differenza tra lo spessore "alto" e "basso" del pixel risulti molto piccola.

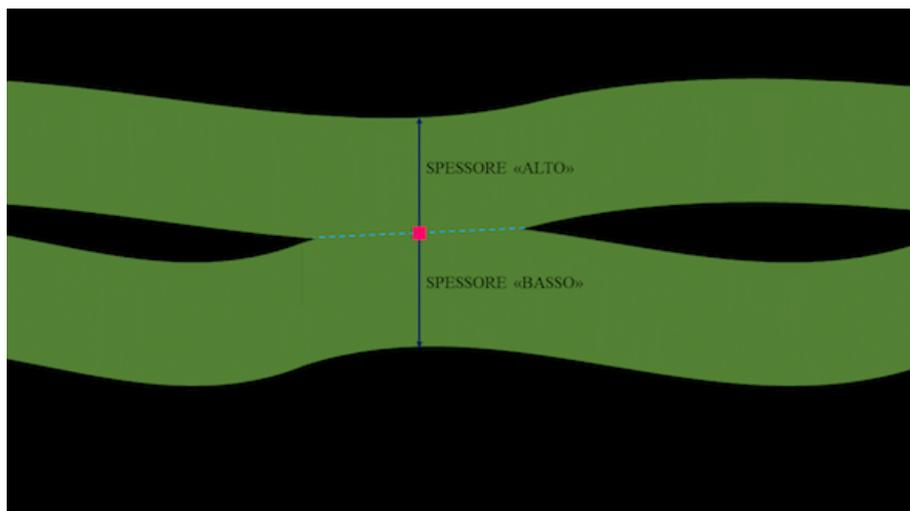


Figura 2.12: Punto di contatto tra due zone diverse.

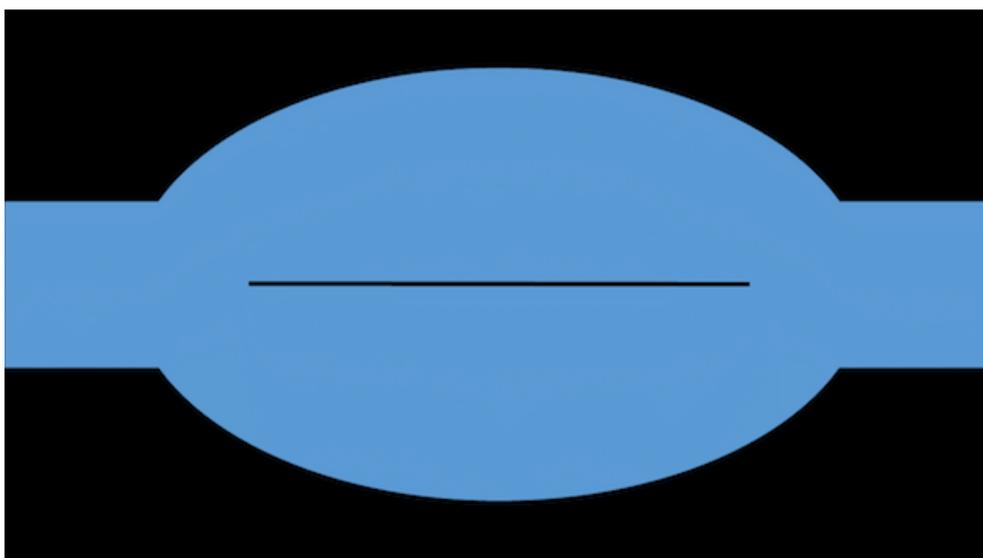
Quindi, prima vengono calcolati gli spessori di ogni zona e la media degli spessori, poi si vanno a cercare quei punti che superano lo spessore medio secondo la seguente formula:

$$\text{media} \times \text{depth threshold} = K$$

dove *depth threshold* rappresenta la soglia in forma di peso (è un valore vicino a 1). Tutti i pixel che si trovano all'interno di una zona con uno spessore verticale superiore al valore di K e che hanno "alto" circa uguale a "basso" vengono classificati come punti di contatto e per questo motivo vengono rimossi dall'immagine.

Ancora una volta, vengono calcolati gli spessori che assumeranno valori differenti da quelli calcolati precedentemente.

Attraverso il procedimento di rimozione dei punti di contatto, però, vengono rimossi anche quei pixel che si trovano al centro di zone ingrossate, i quali avranno dunque uno spessore "alto" circa uguale a "basso" e, per questo, considerati come punti di contatto (Figura 2.13). Dal momento che vengono considerati come punti di contatto, questi pixel vengono eliminati dall'immagine digitale, generando così una lacuna all'interno delle zone relative alle pagine.



*Figura 2.13: Rappresentazione di una zona di fondo generata all'interno di una zona appartenente al segnale relativo alle pagine.*

Per evitare che questo succeda, si studiano i pixel di fondo che sono circondati solo da pixel appartenenti alla stessa zona: se il pixel esaminato ha come vicino al massimo un pixel di fondo a destra e uno a sinistra e e tutti gli altri pixel vicini sono solo pixel appartenenti alla pagina, allora quel pixel viene interpretato come effettivamente appartenente alla zona circostante e quindi la lacuna precedentemente creata verrebbe riempita. A questo punto viene eseguito nuovamente il *filling* e si calcolano nuovamente gli spessori, per un'ultima volta.

## 2.5 Propagazione

Tutto ciò che è stato fatto finora è stato lavorare solo sui pixel di una singola *slice*, ma è necessario che tutto il procedimento precedente venga eseguito su tutte le *slices* (Figura 2.14). Presupponendo di aver chiamato  $I$  l'immagine della sezione su cui si è lavorato precedentemente e della quale si conosce pixel per pixel la zona di appartenenza, si considera ora l'immagine successiva, ovvero  $I+1$ . Si suppone che in questa immagine non si abbiano degli spostamenti di zone consistenti rispetto all'immagine precedente. Dunque, come prima cosa, si classificano tutti i pixel dell'immagine  $I+1$  in base alla loro posizione nelle zone ricavate dall'immagine  $I$  e indipendentemente al loro valore. Poi si va ad analizzare il valore del pixel, studiando se esso sia fondo o segnale. Se il valore del pixel è superiore al valore del *threshold*, si conferma l'associazione alla zona corrispondente all'immagine  $I$ . Nell'immagine  $I+1$  può accadere però che i pixel, spostandosi leggermente dalla posizione precedente, non appartengano più alla zona relativa alla pagina, ma al fondo. Bisogna quindi togliere i pixel che sono diventati bordo ed aggiungere invece quelli che sono diventati parte delle zone del segnale (Figura 2.14).

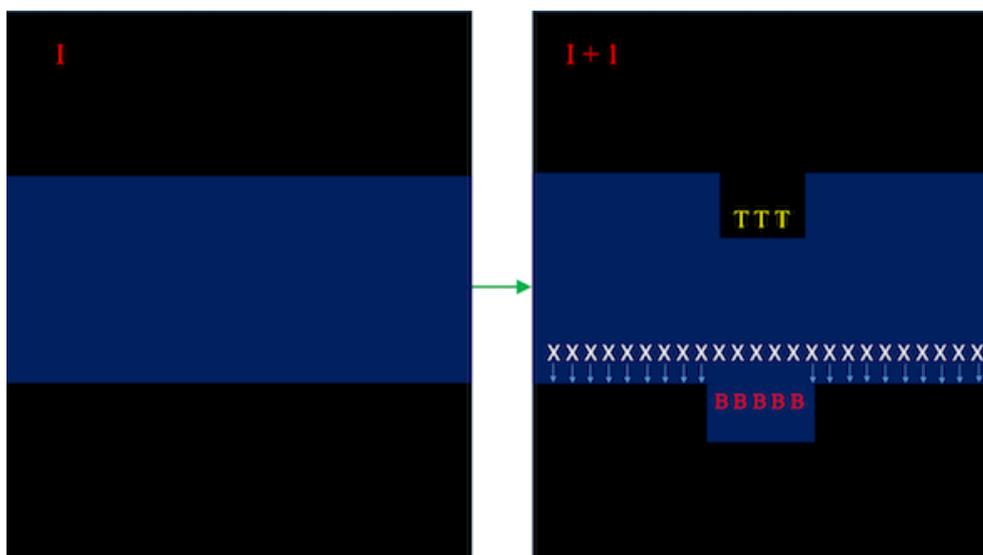


Figura 2.14: Illustrazione dell'algoritmo di propagazione: Immagine  $I$  a sinistra e immagine  $I+1$  a destra. I rettangoli blu rappresentano pixel che appartenevano a una certa zona e sono stati confermati in essa, con  $X$  si indicano i punti di bordo, con  $T$  i pixel rimossi perché diventati fondo, con  $B$  i punti di bordo aggiunti perché diventati segnale.

Nel caso in cui un pixel abbia un valore superiore al *threshold* nell'immagine  $I$  e uno minore del *threshold* nell'immagine  $I+1$ , questo pixel viene semplicemente rimosso. Può invece capitare che un pixel con un valore inferiore al *threshold* diventi un pixel con un valore maggiore. Questo accade solo nei pixel di bordo delle zone, i quali sono vicini sia alla zona che al fondo e quindi hanno una distanza verso l'alto, o verso il basso, pari a

0. Se questi pixel di bordo possiedono valori simili a quello posseduto dal pixel, allora esso viene inserito nell'area del segnale relativa a quella zona.

Questo stesso procedimento viene reiterato per le *slices* successive e, per ognuna di esse, si ottiene un'immagine contenente le informazioni alle varie zone in cui sono classificati i pixel.

In questo modo l'algoritmo propaga le informazioni ricavate dalla prima *slice* a tutte le immagini nella sequenza fino ad elaborarle tutte. Questa fase dell'algoritmo è detta propagazione.

## 2.6 Estrazione delle pagine separate

Dopo la fase di propagazione, per ogni immagine si ha la classificazione dei punti che appartengono ad ogni zona (pagina). Quindi la fase di salvataggio ed estrazione delle pagine prevede che per tutta la sequenza si vadano a prendere tutti i pixel appartenenti ad una certa zona e si effettui una media in verticale dei pixel che le appartengono. Si ottiene così una riga del documento analizzato, la quale viene salvata in un file che rappresenterà la pagina e sotto ad essa verranno via via salvate le righe della stessa zona. Si itera il procedimento per ogni zona/pagina. Nel caso in cui non si riesca a ricavare tutti i pixel di una riga, l'algoritmo introduce una serie di pixel neri all'interno del documento per preservare l'informazione spaziale.

# Capitolo 3

## Risultati

La scrittura privata analizzata è costituita da una pagina ripiegata su se stessa otto volte a formare diverse "pagine"(Figura 3.1).

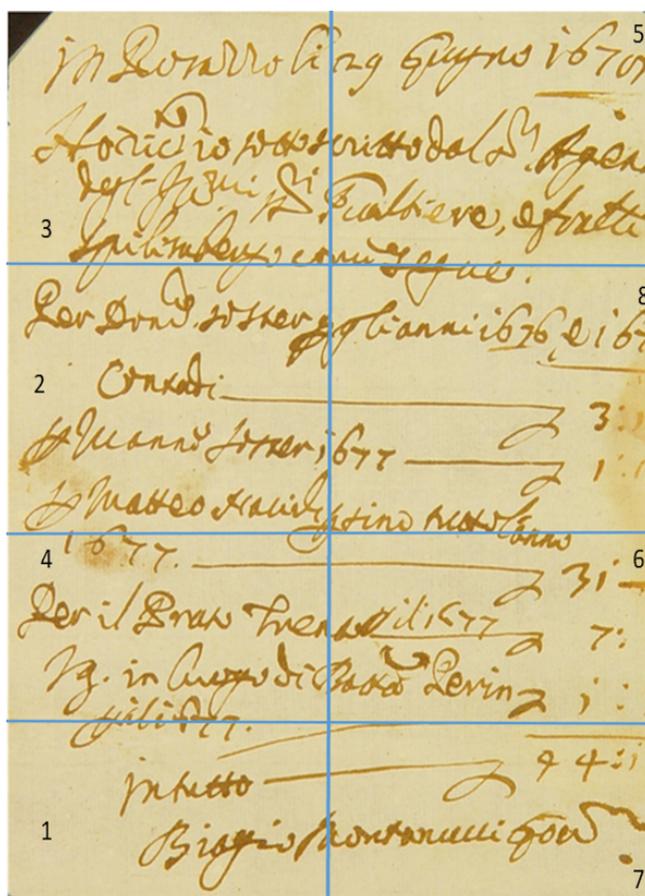
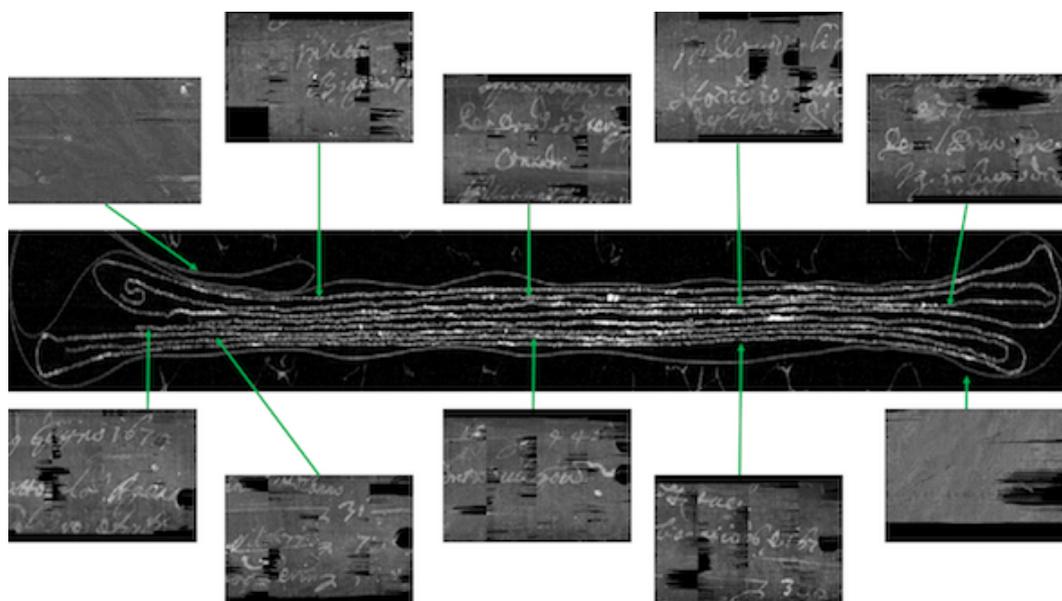


Figura 3.1: Documento originale, suddiviso nelle diverse aree in cui è stato piegato.

Il procedimento descritto nel dettaglio nel capitolo precedente ha portato alla separazione delle pagine del documento analizzato. In Figura 3.2 una *slice* dell'intera tomografia (al centro) e le otto pagine estratte. Da notare che la prima immagine in alto a sinistra e l'ultima in basso a destra non sono reali "pagine", ma l'involucro plastico in cui il manoscritto è stato inserito per procedere all'acquisizione tomografica.



*Figura 3.2: Risultati ottenuti dall'applicazione dell'algorithm: una slice tomografica dell'intero documento (centro) e le otto pagine estratte. Da notare che l'immagine in alto a sinistra e quella basso a destra non sono reali pagine del documento ma parte dell'involucro utilizzato in fase di acquisizione.*

Una delle problematiche riscontrate nell'utilizzo dei vari algoritmi è che questi si sono rivelati estremamente sensibili portando, alcune volte, ad una interpretazione errata e quindi alla necessità di un intervento manuale da parte dell'operatore. In particolare, questo è avvenuto in casi quali pagine non perfettamente separate - in contatto - o pixel rappresentanti un segnale molto debole ed erroneamente identificati come fondo (Figura 3.3).



*Figura 3.3: Elaborazione errata della sezione di una pagina, in cui sezioni diverse si sovrappongono nella stessa immagine.*

Infine, le diverse sezioni del documento sono state unite insieme verticalmente (Figura 3.4) e orizzontalmente (Figura 3.5) attraverso l'algoritmo *Collate*, così da restituire il documento originale (Figura 3.6).

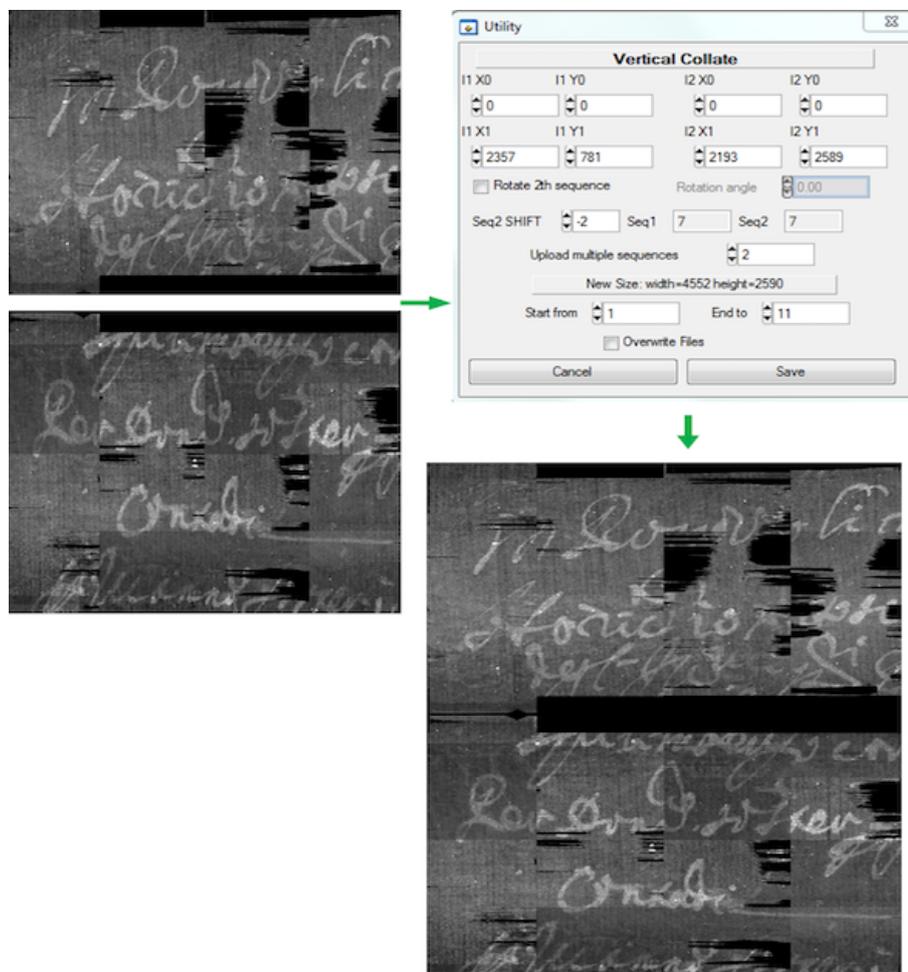


Figura 3.4: Unione verticale di sezioni diverse ottenute dall'elaborazione dei dati.

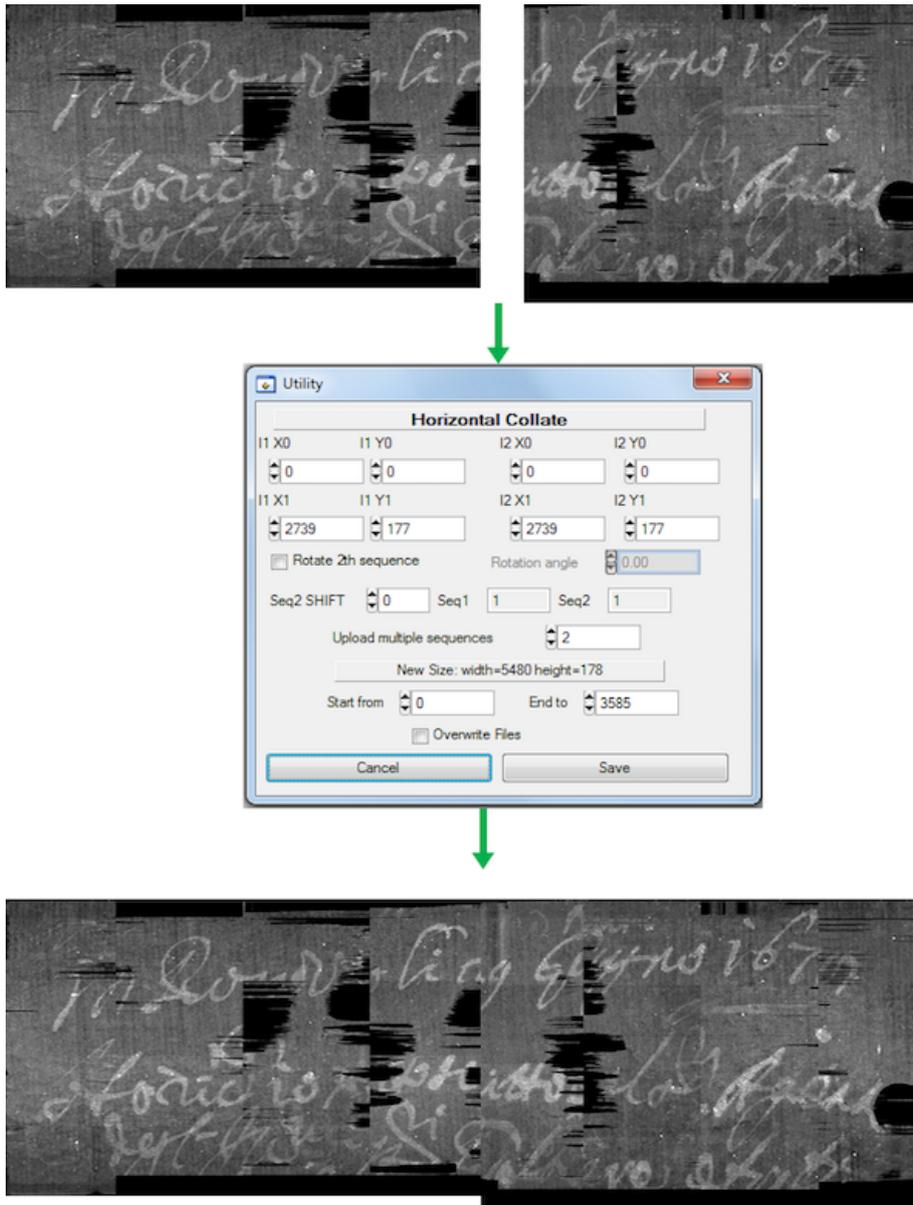


Figura 3.5: Unione orizzontale di sezioni diverse ottenute dall'elaborazione dei dati.

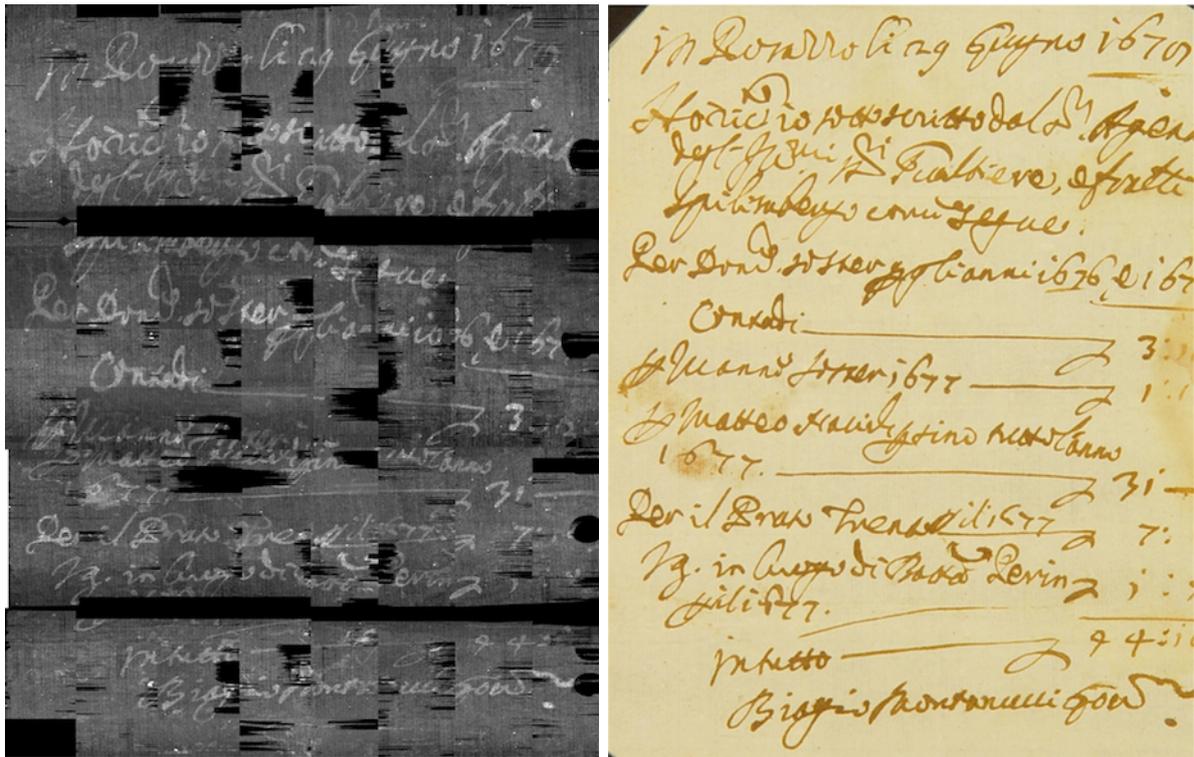


Figura 3.6: Confronto fra l'immagine tomografica ottenuta e la scrittura privata originale.

# Conclusioni

Il progetto di tesi aveva lo scopo di testare gli algoritmi appositamente sviluppati per la segmentazione di immagini tomografiche di documenti antichi. Per questo scopo è stata utilizzata la tomografia a raggi X di un documento manoscritto del 1679 e si è proceduto a testare i singoli algoritmi - segmentazione, analisi geometrica e classificazione ed infine estrazione delle singole pagine.

L'applicazione dei filtri ha permesso una buona separazione del segnale relativo al documento dal *background*, eliminando buona parte del rumore presente e portando, infine, ad una buona identificazione e separazione delle pagine. I risultati ottenuti hanno dimostrato come gli algoritmi siano adeguati alle fasi di classificazione, riempimento e propagazione.

Le maggiori problematiche incontrate sono state due. La prima è stata la difficoltà di rimuovere il rumore di fondo che spesso ha portato all'eliminazione anche di parte del segnale del documento determinando così alcune lacune nella ricostruzione delle aree del documento. La seconda invece è legata alla troppa sensibilità dell'algoritmo in fase di classificazione: anche un singolo pixel classificato erroneamente porta alla propagazione di questo errore per l'intero documento. Per ovviare a tali problemi, si è ritenuto opportuno eliminare meno rumore per non rimuovere troppi pixel del segnale e si è suddivisa la fase di propagazione in intervalli tali da non dover propagare per troppe immagini la classificazione.

In conclusione il presente lavoro ha mostrato l'adeguatezza dell'algoritmo di estrazione delle pagine, pur mettendone in evidenza alcuni limiti su cui in futuro sarà possibile agire per migliorarlo.

# Bibliografia

- [1] V. Mocella, E. Brun, C. Ferrero, & D. Delattre (2015). *Revealing letters in rolled Herculaneum papyri by X-ray phase-contrast imaging*. Nat. Commun. 6;
- [2] F. Albertin, A. Astolfo, M. Stampanoni, E. Peccenini, Y. Hwu, F. Kaplan, G. Margaritondo (2015). *Ancient administrative handwritten documents: X-ray analysis and imaging*. J. Synchrotron Radiat. 22 (2015) 446-451;
- [3] F. Albertin, M. Romito, E. Peccenini, M. Bettuzzi, R. Brancaccio, M. P. Morigi, M. del Rio, D. Raynes, G. Margaritondo & D. Psaltis, (2017, May). *From closed testaments to books: Virtual X-ray reading as an alternate digitization technology for fragile documents*. In Archiving Conference Proceedings (Vol. 2017, No. 1, pp. 14-18);
- [4] A. C. Kak & M. Slaney, *Principles of computerized tomographic imaging* (1988). IEEE, New York.