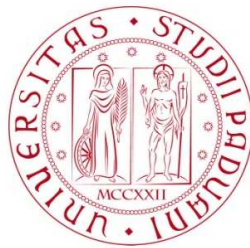


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Magistrale in
Scienze Statistiche



TESI DI LAUREA

**SORVEGLIANZA DELLA CONFORMITÀ DI
GEOMETRIE COMPLESSE**

Relatrice Prof.ssa Giovanna Capizzi
Dipartimento di Scienze Statistiche

Laureanda Sara Bonacina
Matricola n. 2052533

Anno Accademico 2022/2023

Indice

Introduzione	4
1 La sorveglianza statistica di forme geometriche	7
1.1 <i>Additive Manufacturing</i> : opportunità e sfide	7
1.2 Sistemi di misurazione dei prodotti di <i>Additive Manufacturing</i>	11
2 Stato dell'arte sulla sorveglianza statistica di forme geometriche	15
2.1 Metodi basati sulla sorveglianza di superfici	16
2.1.1 Metodi parametrici	16
2.1.2 Metodi non parametrici	20
2.2 Metodi basati sull'analisi di profili di deviazione	22
3 Procedure di sorveglianza per geometrie complesse	27
3.1 Descrizione del caso di studio	28
3.2 Metodo basato sull'analisi delle componenti principali funzionale simpliciale (SFPCA)	34
3.3 Metodo basato sullo spettro dell'operatore di Laplace-Beltrami (LB)	39
4 Implementazione delle procedure di sorveglianza per geometrie complesse	47
4.1 Sorveglianza dei profili di deviazione tramite SFPCA	48
4.2 Sorveglianza non parametrica dello spettro dell'operatore LB discreto	51

Conclusioni	55
A Codice utilizzato	61
A.1 Sorveglianza dei profili di deviazione tramite SFPCA	61
A.1.1 Funzioni ausiliarie	61
A.1.2 Procedura di sorveglianza basata sulla SFPCA	63
A.1.3 Grafici	70
A.2 Sorveglianza non parametrica dello spettro dell'operatore LB discreto	73
A.2.1 Calcolo dell'operatore LB discreto	73
A.2.2 Procedura di sorveglianza non parametrica	77
A.2.3 Analisi post-diagnostica	78
Bibliografia	79

Introduzione

Grazie al recente sviluppo dell'Industria 4.0 e di processi produttivi come l'*Additive Manufacturing* (AM), è possibile realizzare nuovi oggetti caratterizzati da forme irregolari e complesse, che superano le restrizioni geometriche tipiche dei metodi di stampaggio tradizionali. Nell'ambito del controllo statistico di processo (SPC), questa maggiore flessibilità di progettazione si traduce nella necessità di definire nuovi approcci per la modellazione e la sorveglianza di geometrie complesse. La complessità non riguarda solo le forme sempre più libere che i prodotti possono assumere, ma anche la tipologia di dati resi disponibili dai più recenti strumenti di misurazione. Tra questi, i sistemi di sensori senza contatto permettono di ricostruire in formato digitale i prodotti delle tecnologie di stampa 3D sotto forma di nuvole di punti 'non strutturate' e a elevata dimensionalità. Sebbene negli ultimi anni tali argomenti stiano guadagnando sempre più interesse, la letteratura a riguardo è piuttosto contenuta e relativamente recente.

La seguente trattazione si propone di presentare e implementare, mettendoli a confronto, due approcci di sorveglianza che, a differenza degli altri metodi presenti in letteratura, sono applicabili a geometrie effettivamente complesse, rappresentate, in questo elaborato, da gusci trabecolari a forma di uovo, prodotti da un processo di AM e scansionati tramite un sistema di sensori senza contatto. Il primo approccio si basa sulla sorveglianza di profili che descrivono le deviazioni tra la geometria dei prodotti e quella del modello nominale. Il secondo approccio sfrutta, invece, le proprietà geometriche intrinseche e locali degli oggetti, che vengono sorvegliate tramite una procedura non parametrica. Questi approcci meritano una particolare attenzione, poiché si distinguono dai metodi più tradizionali, in quanto sono in grado di

gestire in modo opportuno la complessità geometrica dei dati e la loro elevata dimensionalità.

Il seguito del presente elaborato è così strutturato. Nel Capitolo 1 viene fornita una breve introduzione alle opportunità e alle sfide dell'AM, in particolare nell'ambito dell'SPC, e ai sistemi di misurazione che permettono l'acquisizione di dati relativi alla geometria di prodotti realizzati da processi di AM. Nel Capitolo 2 vengono discussi i principali approcci esistenti in letteratura per la sorveglianza statistica di forme geometriche. Nel Capitolo 3 viene presentato il caso di studio affrontato nella presente trattazione, che costituisce un esempio paradigmatico delle geometrie irregolari e complesse derivanti da processi di AM, e vengono illustrate nello specifico due procedure di SPC per la sorveglianza della conformità di tali prodotti. Nel Capitolo 4 vengono descritti i passaggi necessari per l'implementazione delle due carte di controllo, la loro applicazione al caso di studio e i risultati ottenuti. Infine, nelle Conclusioni sono riportate le osservazioni finali in merito ai metodi presentati e alle analisi effettuate nel presente elaborato. Inoltre, in Appendice A viene fornito il codice R e Matlab necessario per implementare le due carte di controllo e riprodurre le analisi effettuate.

Capitolo 1

La sorveglianza statistica di forme geometriche

1.1 *Additive Manufacturing*: opportunità e sfide

Gli ultimi dieci anni sono stati interessati dalla rivoluzione tecnologica portata dalla cosiddetta ‘Industria 4.0’ nei più svariati ambiti di ricerca e nei settori produttivi e manifatturieri. L’Industria 4.0 è un concetto che combina l’automazione tradizionale con l’Internet delle cose (IoT), l’intelligenza artificiale, l’analisi dei dati, la stampa 3D e altre tecnologie emergenti. L’obiettivo principale dell’Industria 4.0 è applicare, ottimizzare e integrare la conoscenza derivata dai dati, raccolti in tempo reale da sensori e dispositivi avanzati, nei settori industriali e commerciali, al fine di raggiungere nuovi livelli di efficienza produttiva e migliorare la qualità del prodotto. Questa trasformazione digitale consente alle imprese di adottare modelli di produzione flessibili e personalizzati, ridurre i tempi di produzione, minimizzare gli sprechi e rispondere rapidamente alle esigenze del mercato (Pratola, 2022).

Proprio in questo contesto si assiste alla rapida espansione dell’*Additive Manufacturing* (AM), comunemente nota come stampa 3D, una delle tecnologie più innovative e promettenti nell’ambito della produzione industriale. Essa sta rivoluzionando il modo in cui vengono progettati e prodotti gli oggetti in vari settori, dall’industria aerospaziale all’automazione industriale,

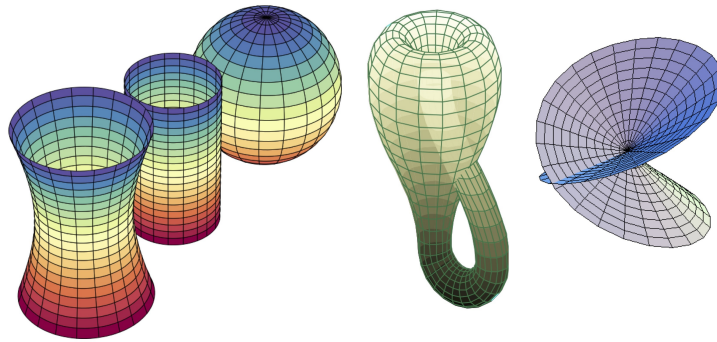


Figura 1.1: Esempi di *manifold*

dalla medicina alla moda (Colosimo et al., 2018a). I prodotti di AM vengono progettati tramite sistemi di CAD (*Computer-Aided Design*), un insieme di tecnologie e software utilizzati per creare disegni tecnici e modelli tridimensionali. A partire da un modello CAD è possibile fabbricare direttamente un prodotto fisico. A differenza dei tradizionali metodi di stampaggio, come quello a iniezione, che prevede anche passaggi di lavorazione, fissaggio e rimozione del materiale in eccesso, un processo di AM aggiunge materiale strato per strato (*layer by layer*) solo dove necessario, in base a quanto definito nel modello CAD. Pertanto l'AM permette di ridurre gli scarti di produzione ed elimina la necessità di realizzare attrezzature specifiche per la creazione di ciascun prodotto (Colosimo et al., 2018b).

Per garantire una corretta generazione del percorso di stampa e prevenire errori durante il processo di produzione, è necessario assicurarsi che le geometrie degli oggetti disegnati via CAD siano solidi 'a tenuta stagna', ossia dal volume chiuso. Si intende, cioè, che le facce o superfici esterne che delimitano un solido non siano difettose, sovrapposte o mancanti. In letteratura per indicare i solidi 'a tenuta stagna' si utilizza il termine *manifold*, che per brevità verrà adottato anche nel seguito della presente trattazione. Più formalmente, un *manifold* è una struttura geometrica astratta che generalizza il concetto di spazio curvo, in quanto localmente si comporta come uno spazio euclideo, ma globalmente può assumere forme tridimensionali note (come cilindri, sfere, etc.), oppure complesse, come mostrato in Figura 1.1. Dunque, un *manifold* rappresenta una prospettiva analitica per descrivere la struttura



Figura 1.2: Esempi di geometrie *free-form*

o l'organizzazione dei dati nel loro spazio multidimensionale. In particolare, nell'ambito dell'AM, il concetto di *manifold* viene utilizzato per descrivere le geometrie, siano esse volumi o superfici, che possono essere realizzate utilizzando la stampa 3D. Questa tecnologia consente, perciò, una maggiore libertà e flessibilità di progettazione, che si traduce nella possibilità di creare oggetti senza le restrizioni geometriche tipiche dei processi di produzione tradizionali: si parla, in questo caso, di strutture irregolari e complesse (*free-form*) (Figura 1.2).

Un esempio rappresentativo dei vantaggi derivanti dall'AM è dato dall'iniettore di carburante rappresentato in Figura 1.3, un componente del *Leap jet Engine* della *General Electric*. Grazie ai nuovi metodi di progettazione e di produzione, è stato possibile fabbricare questo pezzo in modo che, rispetto allo stesso pezzo prodotto con metodi tradizionali, fosse il 25% più leggero (permettendo di risparmiare il 15% di carburante), avesse una durabilità cinque volte maggiore e fosse costituito da un solo componente anziché venti, mantenendo inalterate le sue funzionalità (Kellner, 2017).

Tuttavia, l'AM presenta anche numerose sfide ancora aperte, relative in particolare al controllo statistico di processo (SPC). L'SPC è un insieme di tecniche statistiche finalizzate alla sorveglianza della stabilità dei processi produttivi. Solitamente si sceglie di sorvegliare uno o più parametri del processo di interesse, che dovrebbero contenere quanta più informazione possibile riguardante le caratteristiche di qualità del processo stesso. Attraverso il calcolo di statistiche e limiti di controllo e l'utilizzo di strumenti grafici, che



Figura 1.3: Iniettore di carburante della *General Electric*

insieme costituiscono una carta di controllo, l'SPC si propone di individuare nel più breve tempo possibile se un processo sia in-controllo (IC) o fuori controllo (OC), segnalando di conseguenza un allarme (Montgomery, 2019; Qiu, 2014).

Solitamente si possono individuare due fasi. Nella Fase I, che avviene *offline*, si cerca di ottenere un campione di osservazioni stabili, in modo da poter caratterizzare la distribuzione generatrice delle osservazioni in-controllo. Nella Fase II si effettua, invece, una sorveglianza *online* dei parametri del processo, con lo scopo di verificare in modo sequenziale che le nuove osservazioni del processo, raccolte in tempo reale, siano stabili. Le metriche utilizzate per valutare la *performance* delle carte di Fase I e Fase II sono rispettivamente la probabilità di falso allarme (FAP) e l'*Average Run Length* (ARL). I limiti di controllo vengono usualmente calcolati in modo tale da garantire una certa *performance* in-controllo desiderata.

Nell'ambito dell'AM si possono distinguere due tipologie di sorveglianza. La prima, per *process data*, consiste nel monitoraggio *in situ* della stabilità del processo, sfruttando dati provenienti da immagini, video e sensori presenti sui macchinari di stampa 3D. La seconda, per *product data*, si basa sulla misurazione di caratteristiche riguardanti le proprietà geometriche del prodotto finito, al fine di identificare *ex situ* eventuali anomalie o difetti. La presente trattazione si inserisce in quest'ultimo contesto, caratterizzato da diverse criticità.

Innanzitutto, la complessità geometrica di alcuni prodotti, derivante non solo dalle forme sempre più libere che questi possono assumere ma anche dalla presenza di componenti o superfici interne, fa sì che i pezzi ottenuti via AM siano più soggetti a difetti di produzione che difficilmente possono essere misurati e segnalati dalle carte tradizionali. Per l'SPC si pone, dunque, la necessità di definire nuovi metodi per la modellazione e il controllo della qualità di *manifold* complessi e strutture *free-form*.

Inoltre, non è semplice sorvegliare i parametri dei processi di AM, poiché questi solitamente implicano bassi volumi di produzione. Infatti, data la maggiore flessibilità di progettazione e la possibilità di creare un prodotto strato per strato, semplicemente a partire da un modello CAD, grazie all'AM è possibile fabbricare pezzi personalizzati, cioè adattati a pochi o addirittura a singoli clienti, senza particolari costi aggiuntivi in termini di materiale e risorse. Tuttavia, questo nuovo tipo di produzione, caratterizzata da un numero ridotto di pezzi, rende difficile l'usuale distinzione tra Fase I e Fase II, poiché non è disponibile un numero sufficiente di osservazioni per poter prima stimare i parametri in-controllo del processo e poi sorvegliarli *online*. In questo caso, per l'SPC si tratta di affrontare la sfida della cosiddetta produzione di tipo *short run* (Colosimo et al., 2018b).

1.2 Sistemi di misurazione dei prodotti di *Additive Manufacturing*

I processi di AM possono essere influenzati da diversi fattori, come la temperatura del laser, la velocità di stampa, la composizione chimica del materiale e la configurazione della macchina. Di solito, questi cambiamenti possono causare il deterioramento delle prestazioni del processo, cioè il processo può produrre un numero maggiore di pezzi che presentano geometrie non conformi. Pertanto, è necessario eseguire controlli non solo sul processo produttivo (SPC per *process data*), ma anche sugli oggetti, una volta stampati, attraverso opportuni strumenti di misurazione (SPC per *product data*).

Le Macchine di Misura a Coordinate (*Coordinate Measurement Machines*, CMM) rappresentano i sistemi di misurazione più tradizionali nel contesto ingegneristico e sono utilizzati per determinare le dimensioni e le posizioni di oggetti tridimensionali. Le CMM sono costituite da una piattaforma di misurazione su cui viene posizionato l'oggetto da misurare. Questa tavola è dotata di un sistema di movimento che consente di spostare l'oggetto lungo tre assi: x , y e z . Le CMM sono equipaggiate con sensori di misurazione di contatto (*contact sensors*), come sonde ottiche o tattili, che vengono utilizzati per rilevare punti specifici dell'oggetto. Questi punti vengono quindi registrati e analizzati da un software dedicato, per determinare le coordinate tridimensionali e le caratteristiche geometriche dell'oggetto, come lunghezze, angoli, forme e posizioni (Sładek, 2016). Le CMM consentono, dunque, di ispezionare solo le superfici esterne degli oggetti e di acquisire un numero piuttosto limitato di punti.

Strumenti di misurazione più avanzati, costituiti da sistemi che non prevedono un contatto (*non-contact sensors*), permettono, invece, di acquisire rapidamente milioni di punti, in modo da ricostruire interamente la geometria della superficie di un prodotto. Un primo esempio di tali strumenti è dato dagli *scanner laser*, i quali attraverso un raggio laser scansionano un oggetto e generano una nuvola di punti tridimensionale a elevata dimensionalità. Con gli *scanner laser* è perciò possibile individuare quei difetti che le CMM non sono in grado di rilevare a causa della grande distanza tra un punto di misurazione e l'altro. Nemmeno con questi strumenti è, però, possibile ispezionare superfici interne agli oggetti.

Data la complessità geometrica dei prodotti di AM, che possono presentare superfici o cavità interne o altre caratteristiche non accessibili tramite metodi di ispezione presentati in precedenza, si è reso, dunque, necessario l'utilizzo di un altro sistema di rilevazione senza contatto (*non-contact sensors*), la Tomografia Computerizzata a Raggi X (X-Ray CT). Si tratta di una tecnologia di acquisizione e rappresentazione di immagini avanzata, che utilizza i raggi X per ricostruire sezioni trasversali di un oggetto e ottenere immagini tridimensionali dettagliate delle sue caratteristiche interne senza distruggerlo. Questa tecnologia è ampiamente impiegata in campo medico, ma recentemente sta trovando applicazione anche in ambiti ingegneristici e indu-

striali. Attraverso l'X-Ray CT una serie di immagini radiografiche vengono acquisite da diverse angolazioni intorno all'oggetto da esaminare. Durante la scansione, un tubo emette una sottile fascia di raggi X che attraversa l'oggetto. Un rilevatore posizionato sul lato opposto dell'oggetto rileva la quantità di raggi che emerge dopo aver interagito con l'oggetto. Queste informazioni vengono quindi elaborate da un computer per generare una serie di sezioni trasversali (chiamate 'immagini a sezione') dell'oggetto da diverse prospettive. Le immagini a sezione risultanti possono essere combinate per creare un modello virtuale tridimensionale dell'oggetto esaminato. Questo modello consente agli operatori di visualizzare e analizzare l'oggetto da diverse angolazioni e di esplorare dettagli specifici all'interno dell'oggetto stesso (Kalender, 2011).

Sia gli *scanner laser* sia l'X-Ray CT producono ricostruzioni basate su 'voxel'. Un voxel è l'equivalente tridimensionale di un pixel e viene impiegato nella rappresentazione di dati volumetrici. La densità o il valore del voxel possono essere utilizzati per descrivere varie caratteristiche, come la presenza di materiale, la luminosità, il colore o altre proprietà specifiche dell'oggetto. Entrambi i sistemi di rilevazione senza contatto producono, inoltre, dati e insiemi di dati 'non strutturati', nel senso che i) le posizioni in cui vengono acquisiti i punti sono distribuite in modo irregolare sulla superficie degli oggetti, ii) i pezzi potrebbero presentare orientazioni differenti rispetto al laser o al fascio di raggi X e iii) il numero di punti rilevati può variare tra un pezzo e l'altro.

Capitolo 2

Stato dell'arte sulla sorveglianza statistica di forme geometriche

In questo capitolo vengono brevemente presentati i principali approcci esistenti nella letteratura dell'SPC proposti per trattare problemi che riguardano la sorveglianza di *manifold*. La pratica industriale standard prevede l'utilizzo di caratteristiche chiave del prodotto (KPC), derivanti dalla conoscenza ingegneristica del prodotto, per il controllo della qualità di forme geometriche, anche complesse. Si tratta di approcci basati sul calcolo di semplici indicatori di sintesi legati a prestazioni funzionali o estetiche dei prodotti (approcci *feature-based*), come diametri, rugosità, porosità, lunghezze o distanze. Tali indici descrittivi vengono monitorati nel tempo attraverso carte di controllo univariate o multivariate tradizionali. Tuttavia, questi metodi si rivelano spesso non efficaci, in quanto le KPC non sono in grado di catturare tutte le possibili fonti di variazione relative alla geometria del prodotto, soprattutto se il *manifold* è molto complesso o i difetti sono difficili da individuare (per esempio, difetti di piccola entità o locali). Infatti, in questi casi è possibile che i difetti di produzione non alterino le KPC.

Pertanto, negli ultimi anni si è reso necessario sviluppare metodologie più opportune, che considerassero le effettive caratteristiche geometriche del prodotto piuttosto che indicatori di qualità. In questo contesto si possono individuare principalmente due tipologie di approcci di sorveglianza: uno basato sulla modellazione di superfici tramite metodi parametrici o non parametrici

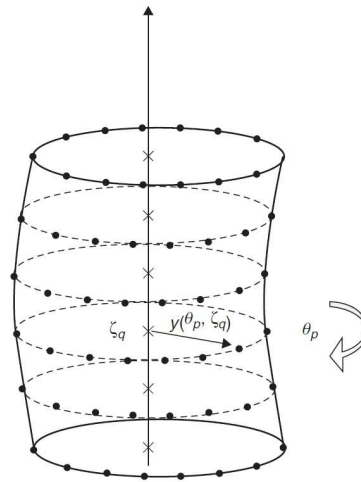


Figura 2.1: Rappresentazione in coordinate cilindriche

e un altro basato sull'analisi dei profili di deviazione tra l'oggetto prodotto e il modello nominale ottenuto via CAD.

2.1 Metodi basati sulla sorveglianza di superfici

2.1.1 Metodi parametrici

I primi contributi nell'ambito dell'SPC di specificazioni geometriche si possono ritrovare in Colosimo e Pacella (2007) e Colosimo et al. (2008), in cui si mostra che la sorveglianza di superfici può essere considerata una generalizzazione della sorveglianza di profili. In questi lavori è d'interesse monitorare la rotondità di cilindri derivanti da un processo di tornitura. Attraverso una CMM per ogni cilindro viene rilevato un certo numero di misurazioni equispaziate del raggio, dove ogni misurazione avviene a una data posizione angolare (rappresentazione in coordinate cilindriche), come mostrato in Figura 2.1. Le misurazioni originarie sono riscalate sottraendo la stima ai minimi quadrati del raggio nominale. Pertanto ogni cilindro può essere visto come un profilo definito da un insieme di punti che costituiscono le variabili esplicative e rappresentano la deviazione del raggio di tale cilindro dal raggio nominale in corrispondenza di un dato angolo (Figura 2.2).

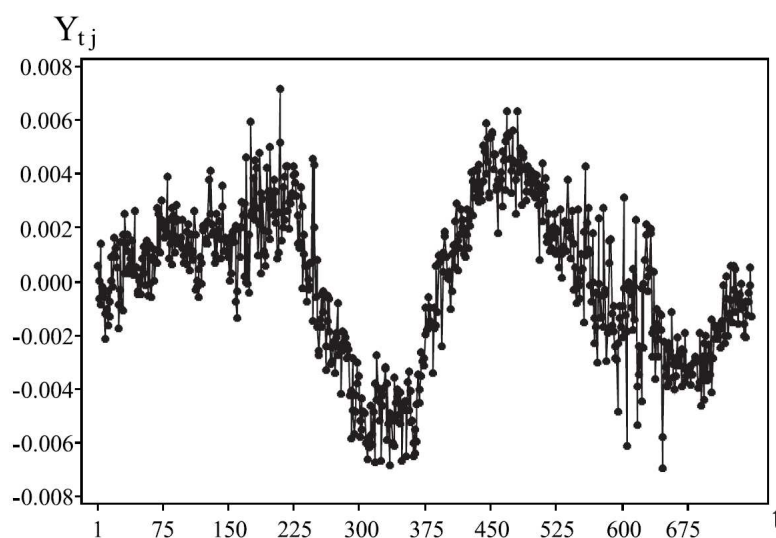


Figura 2.2: Profilo di rotondità di un cilindro j in funzione dell'angolo t

L'approccio proposto da Colosimo e Pacella (2007) consiste nell'effettuare un'analisi delle componenti principali funzionale (FPCA) e nel sorvegliare il processo con una carta di tipo Shewhart multivariata (si veda Qiu, 2014). In particolare, si calcolano la statistica T^2 di Hotelling per i punteggi lungo le componenti principali (PC), selezionate in base alla proporzione di varianza spiegata, riscalati per i loro autovalori, e la statistica Q per la componente residua. Questo metodo consente di individuare la *signature* del processo, ossia delle regolarità sistematiche (*pattern*) che caratterizzano i profili di rotondità dei cilindri, e di studiare, a partire dalle autofunzioni relative alle PC, le principali fonti di variabilità. Colosimo et al. (2008), invece, adattano un modello di regressione con errori autocorrelati (SARX), selezionando il numero di regressori con la procedura iterativa proposta da Box et al. (2015), e si sorvegliano con le statistiche T^2 di Hotelling e Q , rispettivamente, i coefficienti stimati del modello e il termine di errore. In questo modo è possibile tenere conto in modo opportuno della correlazione spaziale esistente tra le misurazioni relative al medesimo cilindro.

Una soluzione alternativa, presentata successivamente da Colosimo et al. (2014), consiste nel modellare le deviazioni dal raggio nominale, misurate in predeterminate posizioni, espresse in questo caso in coordinate cartesiane,

tramite un Processo Gaussiano (GP) di media costante e funzione di covarianza Matérn (Matérn, 2013). Tale funzione assume che la dipendenza tra le osservazioni diminuisca all'aumentare della distanza tra esse con un andamento parabolico con concavità rivolta verso l'alto o verso il basso, a seconda del parametro di liscio che regola la pendenza. Si costruisce poi, per ogni cilindro, la statistica T^2 di Hotelling, data dalla distanza al quadrato tra tali deviazioni e la media del GP, riscalata per la matrice di covarianza. Questo approccio, a differenza dei precedenti, consente di evitare il passaggio di selezione del numero di PC o di regressori, che talvolta può rivelarsi complicato, in quanto non esiste un criterio univoco per tale scelta. Per esempio, il numero di PC o di regressori può essere fissato a priori in modo da favorire l'interpretabilità degli stessi oppure può essere scelto in modo tale da spiegare una certa proporzione (più o meno arbitraria) di variabilità.

Il principale limite dei metodi sopra presentati è dato dal fatto che le geometrie considerate sono '2.5 dimensionali'. In questo contesto, con il termine superficie '2.5 dimensionale' si intende una superficie in cui una delle tre dimensioni può essere espressa in funzione delle altre due (come nel caso dei cilindri, appunto). Questi approcci sono particolarmente utili, quindi, quando i dati sono ottenuti tramite una CMM, che, come si è detto, consente di ispezionare gli oggetti lungo ciascuno dei tre assi, fissati gli altri.

Tuttavia, nell'ambito dell'AM, geometrie non standard (*free-form*), ottenibili dai processi di stampa 3D, e nuvole di punti non strutturate, derivanti dalle scansioni tramite laser o X-Ray CT, rappresentano ormai la norma piuttosto che l'eccezione. Un metodo che permette di modellare superfici effettivamente tridimensionali è presentato da Del Castillo et al. (2015). Nei contesti industriali in cui è possibile ottenere ricostruzioni di superfici attraverso sensori senza contatto, è infatti solitamente difficile scegliere una coordinata euclidea (per es. z) come 'risposta' in funzione delle altre due coordinate, dette 'di posizione', (per es. x e y). Per questo motivo Del Castillo et al. (2015) propongono di utilizzare una rappresentazione parametrica della superficie ricostruita, in cui ogni coordinata (x , y e z) è modellata attraverso un Processo Gaussiano sullo spazio parametrico definito dalle coordinate (u , v) (Figura 2.3). Il GP genera quindi $\hat{x}(u, v)$, $\hat{y}(u, v)$ e $\hat{z}(u, v)$, dove le coordinate (u , v) devono prima essere calcolate attraverso algoritmi di

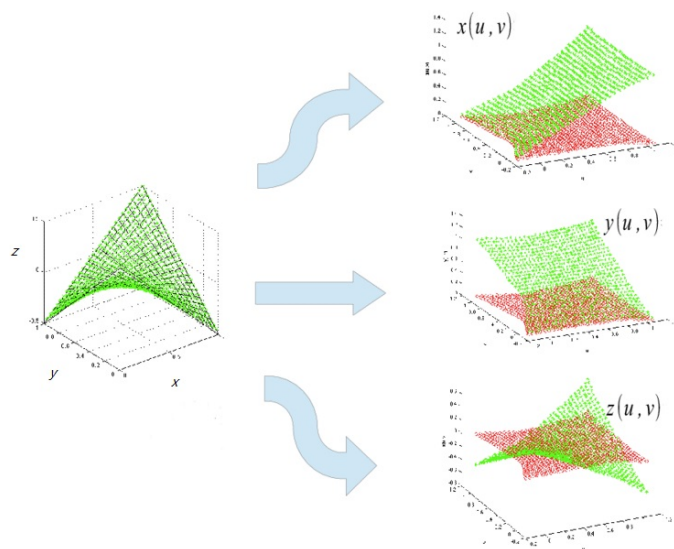


Figura 2.3: Riparametrizzazione di (x, y, z) in (u, v)

manifold learning, ovvero attraverso tecniche non lineari (a differenza della PCA) di riduzione della dimensionalità, spesso impiegate nell'apprendimento automatico per analizzare e rappresentare dati complessi (a elevata dimensionalità) in spazi di dimensioni ridotte. Tra le più utilizzate si evidenziano ISOMAP (*Isometric Feature Mapping*) (si veda Tenenbaum et al., 2000) e ARAP (*As-Rigid-As-Possible*) (si veda Liu et al., 2008).

I parametri di regressione dell'equazione di stato e della matrice di varianza e covarianza del GP vengono stimati tramite REML (*Restricted Maximum Likelihood*). In questo caso la correlazione spaziale è definita in funzione della distanza geodesica, la quale rappresenta la lunghezza del percorso più breve tra due punti lungo la superficie. Gli approcci parametrici precedentemente proposti per la sorveglianza di superfici, basati su modelli gaussiani di *kriging*, utilizzano, invece, la distanza euclidea (si veda, per esempio, Wang et al., 2014), che può differire significativamente dalla distanza geodesica, soprattutto quando si tratta di problemi che coinvolgono la geometria delle superfici. Dunque, la scelta della distanza geodesica appare ragionevole in questo contesto, in quanto essa segue la curvatura della superficie e riflette la reale separazione tra i punti.

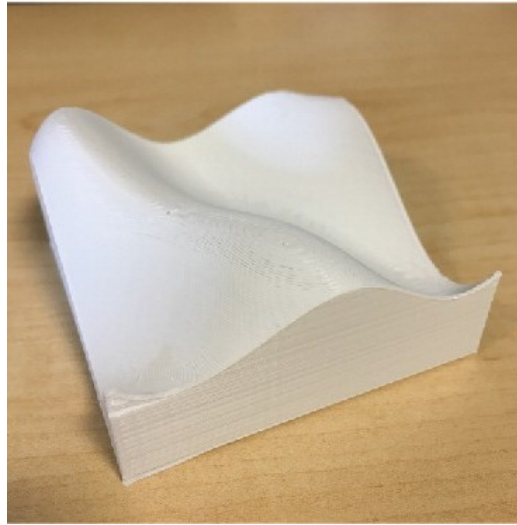


Figura 2.4: Superfici considerate in Zang e Qiu (2018)

Il metodo proposto da Del Castillo et al. (2015), tuttavia, presenta delle difficoltà nella definizione di opportune riparametrizzazioni in (u, v) quando si tratta di superfici molto irregolari o con curvature marcate e frequenti. Inoltre, tale approccio si rivela inefficiente se il numero di punti considerato per la ricostruzione delle superfici è elevato, poiché il calcolo delle REML richiederebbe l'inversione di una matrice molto grande. Infine, Del Castillo et al. (2015) non presentano una carta di controllo per la sorveglianza delle superfici ricostruite con tale metodo.

2.1.2 Metodi non parametrici

In letteratura sono state proposte anche carte di controllo non parametriche di Fase I (Zang e Qiu, 2018a) e di Fase II (Zang e Qiu, 2018b) per la sorveglianza di superfici complesse. È importante sottolineare che in questi due articoli è d'interesse monitorare solo la parte superiore dell'oggetto prodotto da un processo di AM; tutte le restanti facce sono assunte piatte, come mostrato in Figura 2.4. L'approccio non parametrico risulta particolarmente adatto nel caso di superfici irregolari e complesse (*free-form*), grazie alla sua maggiore flessibilità rispetto ai metodi parametrici. Infatti, i Processi Gaussiani solitamente assumono che la correlazione spaziale sia omogenea e

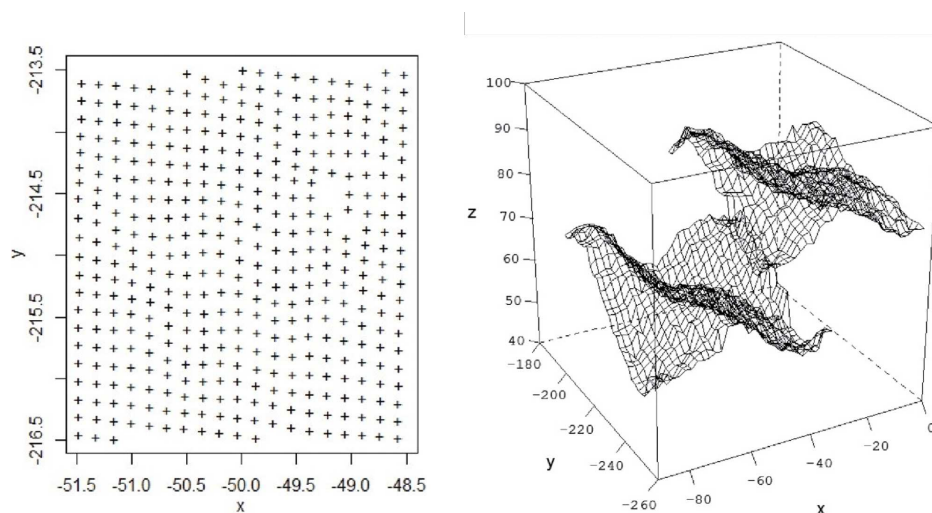


Figura 2.5: Griglia irregolare dei punti rilevati e superfici non allineate

che i punti osservati delle superfici siano normalmente distribuiti, ipotesi non semplici da verificare. Inoltre, tali punti, ottenuti spesso tramite le scansioni *laser*, non sono equispaziati, nel senso che le posizioni in cui lo *scanner* acquisisce i punti sono distribuite in modo irregolare. In aggiunta, la posizione relativa tra l'oggetto e il *laser* può variare leggermente tra un prodotto e l'altro, cioè diversi oggetti possono presentare diverse posizioni e orientazioni rispetto allo *scanner* (Figura 2.5). Rispetto ai metodi parametrici, gli approcci presentati da Zang e Qiu (2018a,b) affrontano, per la prima volta in letteratura, anche tale problema, assai critico, relativo all'allineamento di superfici ricostruite tramite sistemi di sensori senza contatto, i quali, come si è detto nel Paragrafo 1.2, producono dati 'non strutturati'.

Pertanto, il primo passaggio delle procedure non parametriche, sia di Fase I (Zang e Qiu, 2018a) sia di Fase II (Zang e Qiu, 2018b), consiste nella registrazione e nella stima delle superfici osservate. A tale scopo si utilizza una trasformazione rigida, cioè una trasformazione geometrica che conserva le distanze, e di conseguenza le ampiezze degli angoli, tra i punti nello spazio, mantenendo la forma e la dimensione degli oggetti senza distorcerli o deformati. Le trasformazioni rigide includono traslazioni, rotazioni e riflessioni. La scelta di una trasformazione rigida risulta particolarmente adatta in questo contesto, in cui il disallineamento è dovuto solo alla posizione del-

l'oggetto rispetto allo *scanner*. Le matrici di traslazione e di rotazione sono stimate minimizzando la distanza elevata al quadrato tra la superficie osservata e quella registrata tramite la trasformazione rigida. Per le superfici osservate si assume poi un modello non parametrico, che si stima attraverso una procedura di lisciamento locale (*local constant kernel*, LCK).

Nella procedura di Fase I, il secondo passaggio consiste nel calcolo di una statistica di controllo basata sulla somma delle differenze in valore assoluto tra le superfici registrate e la superficie media registrata. Tale statistica, di tipo Shewhart, segnala un allarme quando questa supera un dato limite di controllo, determinato via *bootstrap* in modo tale da garantire una pre-specificata probabilità di falso allarme (FAP). Nella procedura di Fase II, invece, si utilizza una carta tipo *Cumulative Sum* (CUSUM) (si veda Qiu, 2014), costruita sulla differenza in valore assoluto tra l'*i*-esima superficie osservata sequenzialmente *online* e la superficie in controllo stimata in Fase I. Viene segnalato un allarme quando la statistica di controllo CUSUM supera un dato limite, determinato via *bootstrap* a blocchi in modo tale da garantire l'ARL (*Average Run Length*) in controllo desiderata.

Uno dei principali svantaggi delle metodologie non parametriche presentate da Zang e Qiu (2018a,b) è dato dal fatto che viene considerata solo la superficie superiore degli oggetti prodotti, mentre potrebbe essere d'interesse monitorare l'intera superficie (anche i lati o la parte inferiore). Inoltre la fase di registrazione e stima delle superfici è particolarmente onerosa da un punto di vista computazionale, specialmente quando il numero di punti rilevati è elevato.

2.2 Metodi basati sull'analisi di profili di deviazione

Una seconda tipologia di approcci proposta in letteratura evita la fase di stima parametrica o non parametrica della superficie, prevista dai metodi presentati in precedenza, in quanto questa potrebbe risultare difficile o addirittura inapplicabile, soprattutto in caso di superfici complesse. L'idea principale consiste nel calcolare direttamente la distanza tra ogni elemento della nuvola di punti, che costituisce l'oggetto scansionato tramite sistemi di rilevazione senza contatto, e la forma nominale data dal modello CAD.

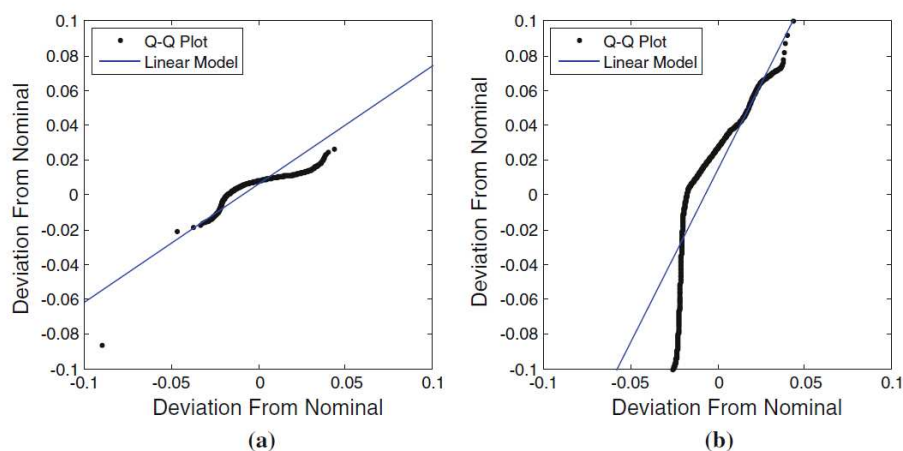


Figura 2.6: Q - Q plot di un pezzo IC (a) e di un pezzo OC (b)

Ogni pezzo prodotto è, dunque, descritto da un insieme di deviazioni che rappresentano una misura complessiva di dissimilarità rispetto al modello nominale.

Un primo approccio di Fase II in questo ambito è presentato da Wells et al. (2013), in cui la distribuzione di tali deviazioni è sintetizzata da un *quantile-quantile (Q-Q) plot*. Questo approccio semplifica notevolmente la questione della sorveglianza di superfici, anche complesse, in quanto si riconduce a un problema di sorveglianza di profili, già ampiamente trattato in letteratura (si vedano Saghaei et al., 2013, Woodall, 2007 e Maleki et al., 2018), e permette di gestire un numero molto elevato di punti con un costo computazionale sostenibile. Quando il processo è in controllo, il Q - Q plot dovrebbe avere un andamento lineare, a indicare l'assenza di scostamenti significativi dei pezzi prodotti dal modello nominale. Quando il processo è fuori controllo, invece, si possono riscontrare cambiamenti nella pendenza, nell'intercetta o nei residui del Q - Q plot (Figura 2.6). I coefficienti relativi a intercetta e pendenza sono stimati tramite il metodo dei minimi quadrati, mentre la varianza dei residui è stimata dall'errore quadratico medio. Queste quantità vengono poi monitorate attraverso tre carte tipo *Exponentially Weighted Moving Average (EWMA)* (si veda Qiu, 2014). I limiti di controllo della carta combinata vengono calcolati in modo da garantire l'ARL nomina-

le. Questo metodo basato sul *Q-Q plot*, tuttavia, è opportuno solo se risulta ragionevole assumere la normalità della distribuzione delle deviazioni, cosa non sempre vera e difficile da verificare. Quando viene segnalato un allarme, inoltre, è utile avere a disposizione uno strumento di analisi post-diagnostica che permetta di individuare il difetto. Dato che con il *Q-Q plot* si perde ogni informazione spaziale, Wells et al. (2013) propongono di ispezionare la scansione del pezzo fuori controllo e confrontarla con il modello nominale tramite un software di *3D scanning*.

Altri metodi, ad esempio quelli proposti da He et al. (2017) e Stankus e Castillo-Villar (2019), cercano, invece, di mantenere le informazioni spaziali relative ai punti rilevati attraverso la definizione di opportune regioni di interesse (ROI). Anche in questo caso, vengono prima calcolate le deviazioni tra l'oggetto prodotto e il modello nominale. Successivamente, si specificano le ROI e si ottiene la media delle deviazioni in ciascuna di esse. Tali medie vengono sorvegliate con una statistica MGLR (*Multivariate Generalized Likelihood Ratio*) (si veda Qiu, 2014). L'analisi post-diagnostica si può effettuare facendo un grafico delle deviazioni medie suddivise per ogni ROI. Un primo limite di questi metodi è dato dalla mancanza di un criterio per selezionare il numero e la dimensione delle ROI; inoltre, non viene tenuta in considerazione la dipendenza tra le deviazioni medie delle varie ROI, pertanto l'utilizzo di una statistica MGLR, che assume invece indipendenza tra le variabili (cioè le ROI), potrebbe non essere appropriato.

Recentemente Bui e Apley (2022) hanno proposto, invece, una metodologia che permette di identificare e visualizzare le fonti di variazione tra un prodotto e l'altro attraverso l'applicazione di tecniche di *manifold learning* alle deviazioni dal modello nominale. Tali algoritmi basati su misure di dissimilarità permettono di rappresentare le nuvole di punti in uno spazio di coordinate di dimensione inferiore. Uno dei più noti è, per esempio, il t-SNE (*t-distributed Stochastic Neighbor Embedding*), che viene spesso utilizzato per visualizzare dati complessi in uno spazio bidimensionale o tridimensionale, cercando di preservare le relazioni di similarità tra i punti originari (Maaten e Hinton, 2008). L'algoritmo t-SNE calcola la similarità tra punti nello spazio originario (o di partenza) e crea una mappa di probabilità che rappresenta tali relazioni. Successivamente, utilizza un processo di ottimizzazione per tro-



Figura 2.7: Fonti di variazione dei cilindri rispetto a una delle coordinate dello spazio ridotto da tre diverse prospettive (lato destro e sinistro e dall'alto)

vare una rappresentazione a bassa dimensione che minimizza la divergenza di Kullback-Leibler (Kullback e Leibler, 1951) tra la mappa di probabilità nello spazio di partenza e la mappa di probabilità nella rappresentazione a bassa dimensione. Questo processo viene ripetuto iterativamente fino a raggiungere una soluzione stabile. Questi metodi permettono, poi, di visualizzare le principali fonti di variabilità del processo basandosi sulle coordinate dello spazio ridotto (Figura 2.7). Non si tratta, perciò, di una metodologia standard di sorveglianza statistica, in quanto non viene proposta una carta di controllo, bensì di un utile strumento grafico.

Infine, un ulteriore contributo appartenente a questa tipologia di approcci basati sull'analisi di profili di deviazione è presentato da Scimone et al. (2022). Si tratta di una metodologia molto generale che, a differenza dei metodi proposti in precedenza da Wells et al. (2013), He et al. (2017) e Stankus e Castillo-Villar (2019), evita assunti distributivi, i quali potrebbero risultare inappropriati o troppo semplicistici (come l'assunzione di normalità o di indipendenza delle ROI). La proposta di Scimone et al. (2022), di particolare interesse in quanto costituisce una delle prime procedure di sorveglianza statistica applicabili anche a geometrie tridimensionali molto complesse, rappresentate in forma di nuvole di punti 'non strutturate', sarà approfondita nel prossimo Capitolo.

Capitolo 3

Procedure di sorveglianza per geometrie complesse

Come evidenziato nel Paragrafo 1.1, le principali sfide per l'SPC nell'ambito dell'AM sono date dalla necessità di definire nuovi metodi di sorveglianza in grado di gestire la complessità geometrica dei prodotti e di sviluppare delle procedure che siano applicabili anche nel contesto della produzione di tipo *short run*. In questo Capitolo viene descritto un caso di studio rappresentativo di tali criticità, che riguarda un numero ridotto di gusci trabecolari a forma di uovo prodotti da un processo di AM. I metodi descritti nel Capitolo 2 non risultano adeguati per la sorveglianza di tali prodotti a causa dei limiti modellistici e computazionali discussi in precedenza. Inoltre, alcuni di essi assumono di disporre di un elevato numero di pezzi per poter stimare in modo accurato i parametri del processo in-controllo.

In questo Capitolo verranno, dunque, presentate due metodologie di Fase I piuttosto generali in grado di gestire in modi diversi la complessità dei dati derivanti dal caso di studio. Il primo metodo (Scimone et al., 2022), che appartiene alla categoria degli approcci basati sui profili di deviazione, combina una mappa bidirezionale per modellare le distanze tra la geometria ricostruita e quella del modello nominale con un approccio di sorveglianza di profili. Il secondo metodo (Zhao e Del Castillo, 2021), che, invece, rappresenta un approccio del tutto innovativo rispetto a quelli finora discussi, sfrutta le proprietà geometriche intrinseche degli oggetti prodotti, sintetizzate

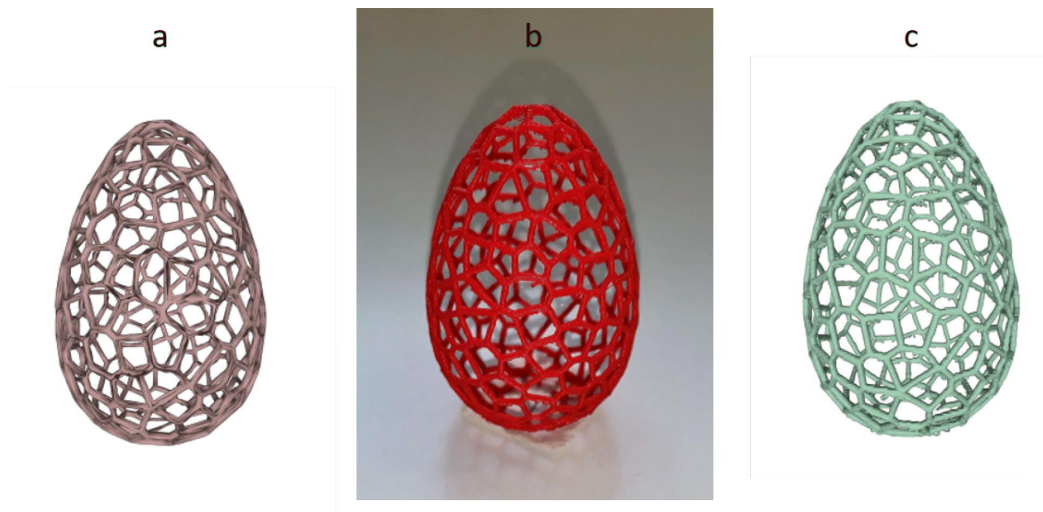


Figura 3.1: Ovetto: modello CAD (a), prodotto finito (b) e ricostruzione tramite X-Ray CT (c)

dallo spettro dell'operatore di Laplace-Beltrami (Kreyszig, 2013), che viene monitorato tramite una carta di controllo multivariata non parametrica.

3.1 Descrizione del caso di studio

Il caso di studio, presentato nell'articolo di Scimone et al. (2022), riguarda una struttura trabecolare a forma di uovo, che per semplicità nel seguito verrà chiamata 'ovetto' (Figura 3.1). Tale struttura deriva da una tassellatura di Voronoi (Moller, 2012). Essa è un particolare tipo di decomposizione di uno spazio metrico in regioni, dette 'poligoni di Voronoi', definite a partire dai punti di un dato insieme discreto, chiamati 'punti di generazione'. Tramite questo partizionamento ciascun poligono contiene un solo punto di generazione e ogni altro punto del poligono è più vicino a tale punto di generazione che a qualunque altro punto di generazione (Figura 3.2). La struttura trabecolare degli ovetti è definita da 'montanti' e 'nodi', dove i montanti sono componenti strutturali disposte lungo i bordi dei poligoni e i nodi rappresentano il punto in cui si incontrano due o più montanti.

Ogni ovetto ha un'altezza di 60 mm con un diametro massimo di 40 mm

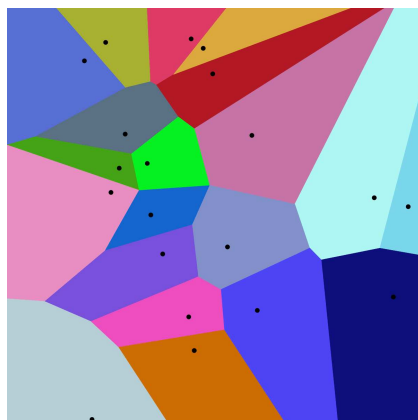


Figura 3.2: Tassellazione di Voronoi di un piano: i puntini neri sono i punti di generazione a partire dai quali sono state definiti i poligoni di Voronoi (aree colorate)

ed è stato prodotto tramite un particolare processo di AM, di ‘Modellazione a Deposizione Fusa’ (FDM), in cui un filamento di materiale termoplastico viene riscaldato, estruso attraverso una punta di ugello e depositato su una piattaforma per costruire oggetti strato dopo strato. Gli ovetti stampati sono stati poi scansionati e ricostruiti tramite X-ray CT, l’unico strumento utilizzabile in questo contesto per ispezionare la geometria della superficie di tali oggetti. Le misurazioni mediante X-ray CT, come si è detto nel Paragrafo 1.2, producono ricostruzioni basate su voxel, che possono essere anche rappresentate come una serie di immagini in scala di grigi, in cui una bassa intensità di grigio corrisponde alle zone del prodotto a bassa densità di materiale (come i buchi o gli spazi vuoti degli ovetti), mentre un elevato valore di grigio indica i punti di materiale ad alta densità (come i montanti che costituiscono la struttura trabecolare degli ovetti). La superficie ricostruita in questo modo può essere convertita in formato di maglia (*mesh*) tramite opportuni software di progettazione e disegno tecnico.

Nel contesto dell’AM, una *mesh* è una riproduzione poligonale di una superficie 3D o di un solido, che risulta particolarmente utile in fase di modellazione via CAD, poiché permette di convertire una geometria continua in una rappresentazione discreta composta da triangoli o altri poligoni (*triangulated* o *polygonal mesh*), così da semplificare il processo di visualizzazione,

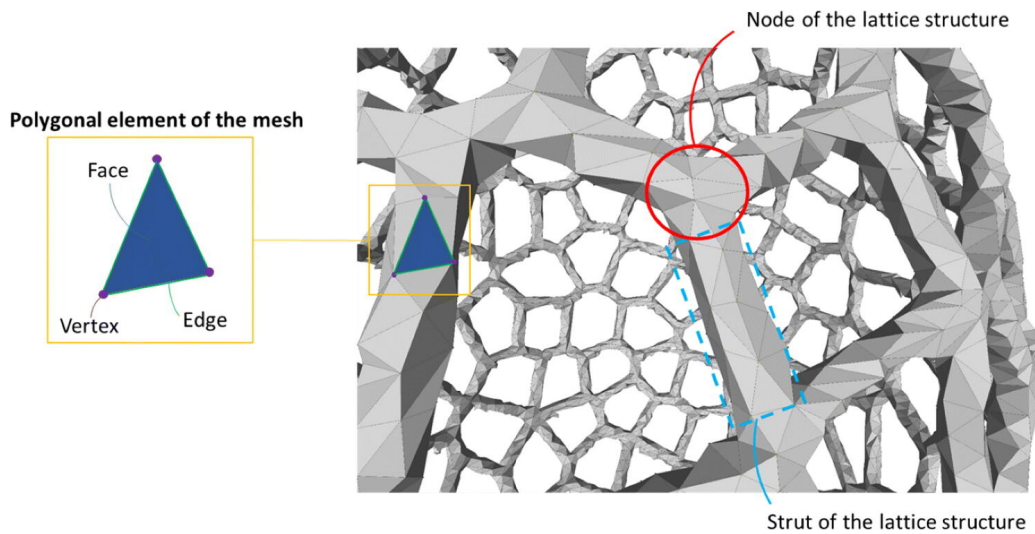


Figura 3.3: Terminologia specifica della struttura trabecolare (nodi e montanti) e della *mesh* (facce, archi e vertici)

modifica e analisi dell'oggetto. I termini 'faccia', 'arco' e 'vertice' sono utilizzati per indicare gli elementi geometrici di ciascun poligono di cui è composta la *mesh*. Nel presente caso di studio, le facce sono triangoli e gli archi e i vertici sono rispettivamente i lati e i vertici di tali triangoli. La Figura 3.3 mostra un dettaglio della *mesh* ottenuta da un ovetto scansionato tramite X-Ray CT e chiarisce la terminologia utilizzata nel resto della trattazione. In alternativa, spesso si utilizza una rappresentazione basata solo sui vertici della *mesh*, in quanto si assume che questi siano così densi che tralasciare gli archi comporti una perdita di informazione trascurabile o nulla. In questo caso si ottiene una 'nuvola di punti', come mostrato in Figura 3.4.

Nel processo di produzione degli ovetti è ovviamente possibile riscontrare dei difetti geometrici, data la complessità delle forme considerate. In particolare, vengono qui analizzate due tipologie di non conformità rappresentative delle anomalie più comuni nell'FDM e nei processi di AM in generale. Tipicamente si tratta di distorsioni geometriche locali, pori e crepe.

In particolare, nel caso degli ovetti gli errori geometrici sono legati a un eccesso o a una mancanza di materiale causato da vari fattori. Un eccesso di materiale potrebbe derivare dal cosiddetto 'effetto di stringatura', che

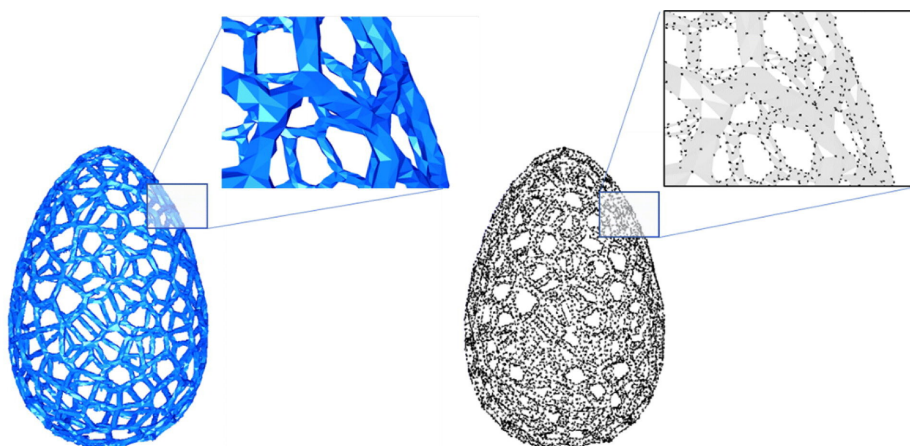


Figura 3.4: Ovetto ricostruito tramite X-Ray CT: rappresentazione in forma di *mesh* e di nuvola di punti

porta alla formazione di sottili ed antiestetici fili di plastica in diverse parti dell'oggetto. Questo si verifica quando, mentre l'ugello della stampante si sposta per passare da un punto all'altro, il filamento fuso che rimane dentro all'ugello fuoriesce e si solidifica durante lo spostamento, formando dei fili attaccati alle parti stampate. Solitamente l'effetto di stringatura è dovuto all'utilizzo non ottimale di alcuni parametri del processo (quali la velocità di stampa e la temperatura dell'ugello).

In Figura 3.5 si possono osservare le irregolarità locali riguardanti i montanti di un ovetto non conforme. Tale ovetto è stato prodotto con un pigmento diverso da quello usato per gli ovetti in-controllo. La colorazione del materiale termoplastico determina comportamenti anomali quando il filamento polimerico viene riscaldato ed estruso, con conseguenti effetti sulla qualità finale e sulla struttura del pezzo.

L'assenza di materiale, invece, potrebbe essere causata da una mancanza di adesione locale tra due strati, da un flusso di materiale interrotto o da un ugello non allineato. La Figura 3.6 mostra un esempio di un ovetto con un montante mancante, indicato da una linea rossa. In questo caso il difetto è stato inserito artificialmente: il montante è presente nella geometria nominale, ma è stato rimosso dall'ovetto stampato in 3D. Un'anomalia di questo tipo può chiaramente compromettere le prestazioni meccaniche dell'intero pezzo.

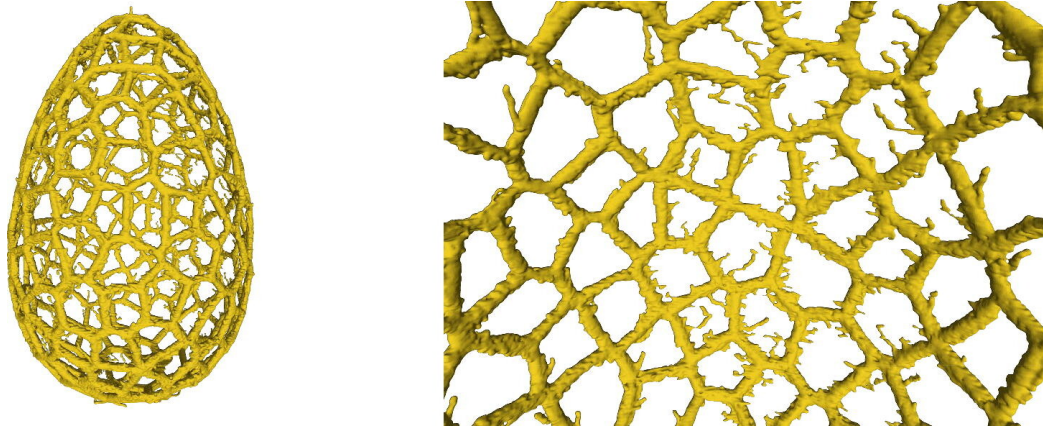


Figura 3.5: Ovetto caratterizzato da un eccesso di materiale

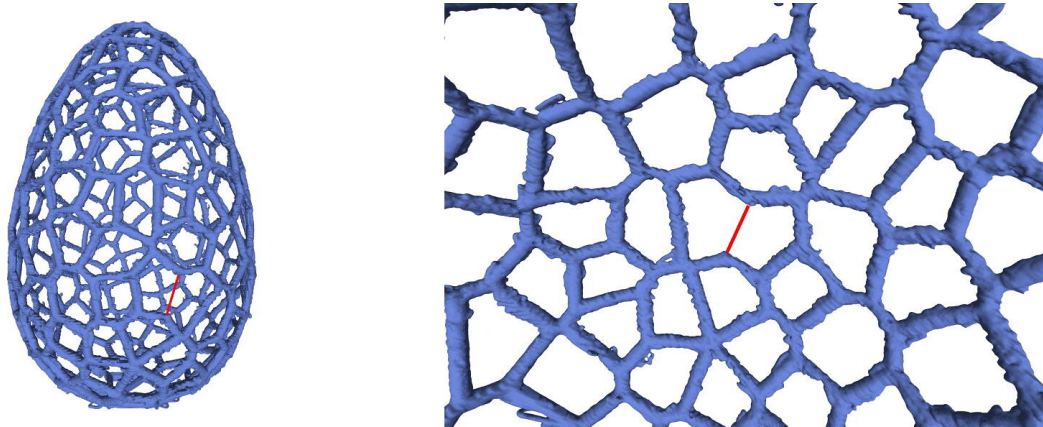


Figura 3.6: Ovetto caratterizzato dalla mancanza di materiale

Nel caso di studio si dispone di 13 ovetti: 8 presentano caratteristiche geometriche conformi al modello nominale e sono pertanto considerati in-controllo, mentre 5 sono fuori controllo. In particolare, uno di questi è caratterizzato dalla mancanza di un montante e gli altri quattro sono stati prodotti utilizzando pigmenti diversi, che hanno generato anomalie di varie intensità legate a un eccesso di materiale.

L'obiettivo della presente trattazione è quello di descrivere e implementare delle procedure di SPC adatte al caso di studio in questione. È evidente che gli approcci basati sulla sorveglianza di superfici, presentati nel Paragrafo 2.1, non risultano applicabili in questo contesto, poiché gli ovetti sono

geometrie completamente *free-form* e tridimensionali. Pertanto, non è possibile né descrivere tali forme tramite modelli parametrici o non parametrici né esprimere una delle dimensioni in funzione delle altre due (come nel caso delle superfici 2.5 dimensionali). Inoltre, l'elevata dimensionalità delle nuvole di punti rendono computazionalmente intrattabile la fase di registrazione delle superfici così come prevista da Zang e Qiu (2018a,b).

D'altra parte, gli approcci basati sulle deviazioni dal modello nominale, presentati nel Paragrafo 2.2, richiedono ipotesi distributive difficili da verificare, come la normalità della distribuzione delle deviazioni (Wells et al., 2013) o l'indipendenza delle ROI (He et al., 2017; Stankus e Castillo-Villar, 2019). Inoltre, questi metodi considerano insiemi di dati 'strutturati', in cui esiste una corrispondenza o punto per punto o tra ROI di due nuvole di punti. In questo modo, tali approcci ignorano il problema non trascurabile dell'allineamento di nuvole di punti 'non strutturate'. Tuttavia, nel presente caso di studio il numero di punti ottenuti dalla scansione tramite X-Ray CT varia tra un ovetto e l'altro e non sembra esserci un modo ragionevole per definire delle ROI. Infine, non vengono considerate le criticità derivanti dalla produzione di tipo *short run*: la stima dei parametri del *Q-Q plot* diviene più accurata all'aumentare del numero di osservazioni (Wells et al., 2013), così come la *performance* della carta MGLR migliora al crescere della numerosità dell'insieme di osservazioni pre-campionarie (He et al., 2017; Stankus e Castillo-Villar, 2019).

Le metodologie presentate nel seguito costituiscono, pertanto, due strategie per la modellazione e la sorveglianza di geometrie effettivamente complesse, che superano i limiti degli approcci descritti finora. Uno dei contributi della presente trattazione riguarda l'applicazione di tali metodi in situazioni di bassa numerosità campionaria. Vengono prese in considerazione due carte di Fase I, poiché in un contesto di produzione di tipo *short run*, come quello degli ovetti, non sembra ragionevole considerare un approccio di sorveglianza *online*, dato il numero di osservazioni disponibili necessariamente limitato.

Inoltre, in questo elaborato la *performance* delle metodologie proposte da Scimone et al. (2022) e da Zhao e Del Castillo (2021) è valutata in condizioni differenti rispetto a quelle considerate dai rispettivi autori. La carta di Scimone et al. (2022) è stata disegnata appositamente per lo studio di ovetti

analoghi a quelli descritti nel presente Paragrafo, tuttavia gli autori prendono in considerazione nelle loro applicazioni solo difetti isolati e molto marcati. Per contro, nella presente trattazione verranno considerati difetti persistenti e di varia intensità. La carta di Zhao e Del Castillo (2021), invece, è stata ideata come metodologia di sorveglianza *online*, mentre in questo elaborato viene proposta una sua applicazione in un contesto di Fase I.

3.2 Metodo basato sull'analisi delle componenti principali funzionale simpliciale (SFPCA)

Il metodo presentato da Scimone et al. (2022) appartiene alla classe di approcci basati sull'analisi dei profili di deviazione. Esso è costituito da tre passaggi principali, che verranno dettagliati nel seguito: in prima battuta si calcolano le deviazioni di ciascun prodotto, ricostruito tramite sistemi di sensori senza contatto, dal modello CAD; si stima, poi, la funzione di densità di probabilità (PDF) empirica di tali deviazioni; infine si applica alle PDF uno schema di sorveglianza di profili basato su una variante della PCA, l'analisi delle componenti principali funzionale simpliciale (SFPCA) (Hron et al., 2016).

Affinché sia possibile calcolare le deviazioni, questa metodologia assume che sia le ricostruzioni dei prodotti sia il modello nominale siano rappresentati in forma di nuvole di punti allineate. In Scimone et al. (2022) l'allineamento di nuvole di punti 'non strutturate' viene considerato un'operazione preliminare. Si tratta, in realtà, anche in questo caso di un aspetto molto critico. Si ricorda che il problema della registrazione di superfici è stato trattato precedentemente solo da Zang e Qiu (2018a,b), i quali propongono di utilizzare una trasformazione rigida. Anche Scimone et al. (2022) ricorrono a tale trasformazione che, però, in questo caso viene stimata tramite l'algoritmo *Iterative Closest Point* (ICP) molto utilizzato nell'ambito dell'elaborazione di immagini (Besl e McKay, 1992). Applicando l'algoritmo ICP, inizialmente si selezionano manualmente o in modo automatico dei punti di riferimento sulle due nuvole di punti da allineare e successivamente si iterano fino a convergenza due passi: ad ogni punto della prima nuvola viene associato il punto

più vicino della seconda e successivamente le coppie di punti così ottenute vengono utilizzate per stimare le matrici di rotazione e traslazione della trasformazione rigida, in modo da minimizzare una certa metrica di errore. Si tratta di un problema di ottimizzazione discreta non lineare, in quanto non solo bisogna stimare una trasformazione rigida, ma è anche necessario ricercare la corrispondenza puntuale ottimale tra le nuvole di punti sulla base dei punti di riferimento selezionati. Nella sua versione originaria l'algoritmo ICP ha un costo computazionale pari a $O(N_1 \cdot N_2)$, dove N_1 e N_2 sono il numero di punti delle due nuvole da allineare, che risulta proibitivo nel caso di nuvole di punti a elevata dimensionalità. Per questo motivo, le versioni più recenti, come quella utilizzata da Scimone et al. (2022) e descritta poco sopra, effettuano la registrazione su un sottoinsieme di punti (i punti di riferimento). La selezione di tali punti, tuttavia, è piuttosto arbitraria e soggettiva e influenza molto il risultato finale dell'algoritmo. Infine, poiché si tratta di un algoritmo iterativo di ottimizzazione non convessa, è possibile che venga raggiunta solo una soluzione subottimale e non un ottimo globale.

Si supponga che all'inizio della procedura di sorveglianza le nuvole di punti siano allineate e, in particolare, sia S_j , $j = 1, \dots, N$, la nuvola di punti $\{p_1^j, \dots, p_{M_j}^j\}$ costituita da M_j punti relativa al j -esimo prodotto, e P la nuvola di punti $\{p_1^p, \dots, p_{M_p}^p\}$ costituita da M_p punti relativa al modello CAD. Per la prima volta nella letteratura sulla sorveglianza di *manifold*, Scimone et al. (2022) utilizzano come misura di discrepanza dal modello nominale la distanza di Hausdorff (Hausdorff, 1914). La distanza di Hausdorff tra le nuvole di punti S_j e P si definisce come:

$$d_H(S_j, P) := \max \left\{ \sup_{s \in S_j} \inf_{p \in P} d(s, p), \sup_{p \in P} \inf_{s \in S_j} d(s, p) \right\}, \quad (3.1)$$

dove $d(\cdot, \cdot)$ è la distanza euclidea. La distanza di Hausdorff determina, dunque, per ogni nuvola di punti S_j , la seguente coppia di mappe di deviazione:

$$d_{S_j} : P \rightarrow C_j, d_{S_j}(p) := \left\{ \inf_{s \in S_j} d(s, p) \right\} \quad (3.2)$$

e

$$d_P^j : S_j \rightarrow K_j, d_P^j(s) := \left\{ \inf_{p \in P} d(s, p) \right\}, \quad (3.3)$$

dove K_j e C_j sono intervalli in \mathbb{R} . La (3.2) esprime la distanza di ogni punto $p \in P$ dalla nuvola di punti S_j , mentre la (3.3) quella di ogni punto $s \in S_j$ dalla nuvola di punti P . In generale, $\sup_{s \in S_j} d_P^j(s) \neq \sup_{p \in P} d_{S_j}(p)$, in quanto S_j e P potrebbero avere due domini diversi. Per esempio, nel caso in cui S_j presenti un numero di punti maggiore rispetto a P , ci si aspetta che le distanze calcolate a partire dalla mappa d_P^j siano elevate, poiché il suo dominio S_j contiene dei punti che non appartengono a P . Questa situazione si verifica se un pezzo prodotto mostra difetti dovuti a un eccesso di materiale (come gli ovetti soggetti all'effetto di stringatura). Viceversa, se S_j presenta un numero di punti inferiore rispetto a P , ci si aspetta che siano elevate le distanze calcolate a partire dalla mappa d_{S_j} , poiché il suo dominio P contiene dei punti che non appartengono a S_j . Questa situazione si verifica, invece, nel caso di difetti dovuti alla mancanza di materiale (come l'assenza di un montante negli ovetti). Pertanto, sembra opportuno sfruttare la natura bidirezionale della distanza di Hausdorff, modellando e sorvegliando entrambe le mappe di deviazione, al fine di cogliere e caratterizzare in modo più completo le differenze geometriche tra il pezzo prodotto e il modello nominale.

A partire dalle mappe di deviazione, d_P^j e d_{S_j} , per ogni nuvola di punti S_j si ottengono due vettori di deviazioni, D_P^j e D_{S_j} , che sono a elevata dimensionalità (si ricorda, infatti, che la nuvola di punti è densa) e presentano dimensioni differenti al variare di j . Perciò, è necessario sintetizzare le informazioni contenute nelle coppie di vettori di deviazioni in modo opportuno. Poiché l'utilizzo di indici di sintesi, come media, varianza e quantili delle deviazioni, porterebbe a un'eccessiva perdita di informazione, seguendo la proposta di Menafoglio et al. (2018), i vettori di deviazioni sono sintetizzati dalla loro funzione di densità di probabilità (PDF) empirica. Una stima della PDF delle deviazioni contenute nel vettore D_{S_j} si ottiene attraverso lo stimatore polinomiale di Bernstein (Leblanc, 2010; Vitale, 1975), che è uno stimatore liscio e consistente, ed è data da:

$$\hat{f}_{S_j}(t) = M_j \sum_{k=0}^{M_j-1} [(\bar{F}_{S_j}((k+1)/M_j) - \bar{F}_{S_j}(k/M_j))b_{k,M_j}(t)], \quad (3.4)$$

dove \bar{F}_{S_j} è la funzione di ripartizione empirica delle deviazioni in D_{S_j} e b_{k,M_j} è il k -esimo polinomio di Bernstein di ordine M_j . Si procede in modo analogo

anche per le deviazioni in D_P^j , così da ottenere \hat{f}_P^j . Al termine di questo passaggio si ottengono due insiemi di PDF: $\mathbf{f}_S := \{\hat{f}_{S_j}, j = 1, \dots, N\}$ e $\mathbf{f}_P := \{\hat{f}_P^j, j = 1, \dots, N\}$, che sintetizzano la natura bidirezionale delle mappe di deviazione d_{S_j} e d_P^j .

A questo punto, per sorvegliare la stabilità del processo e individuare eventuali pezzi fuori controllo, si applica la SFPCA agli insiemi di dati \mathbf{f}_S e \mathbf{f}_P separatamente. La SFPCA è un'estensione della PCA funzionale nel caso in cui le funzioni considerate appartengano allo spazio di Bayes B^2 , definito come lo spazio delle funzioni positive che integrano a una costante e il cui logaritmo è due volte integrabile, cioè:

$$B^2 = \left\{ f : f > 0, \int_0^t f(t)dt = c, \log(f) \in L^2 \right\}. \quad (3.5)$$

La SFPCA trova le direzioni ξ_1, \dots, ξ_{N-1} in B^2 lungo le quali i dati hanno variabilità massima. Più formalmente, considerando l'insieme \mathbf{f}_S , la prima componente ξ_1 massimizza:

$$\sum_{j=1}^N \langle \hat{f}_{S_j} - \bar{f}_S, \xi^{(S)} \rangle^2 \text{ soggetto al vincolo } \|\xi^{(S)}\| = 1, \quad (3.6)$$

dove $\langle \cdot, \cdot \rangle$ è il prodotto interno e \bar{f}_S è la media campionaria di \mathbf{f}_S , $j = 1, \dots, N$. Per $j > 1$, ξ_j massimizza:

$$\sum_{j=1}^N \langle \hat{f}_{S_j} - \bar{f}_S, \xi^{(S)} \rangle^2 \text{ soggetto ai vincoli } \|\xi^{(S)}\| = 1, \langle \xi^{(S)}, \xi_i^{(S)} \rangle = 0, i < j. \quad (3.7)$$

Dopo aver applicato la SFPCA a \mathbf{f}_S , si ottiene una sequenza di $N - 1$ coppie di autovalori e autofunzioni $(\lambda_i^{(S)}, \xi_i^{(S)})$, $i = 1, \dots, N - 1$, della covarianza campionaria, dove la sequenza di autovalori è non crescente e gli autovettori sono fra loro ortogonali. Inoltre, è possibile ottenere il punteggio di \hat{f}_{S_j} , $j = 1, \dots, N$, lungo l' i -esima componente principale, $i = 1, \dots, N - 1$, come $z_{ji}^{(S)} = \langle \hat{f}_{S_j} - \bar{f}_S, \xi_i^{(S)} \rangle$, che rappresenta la proiezione della PDF originale j lungo la direzione i . Infine, il numero $K^{(S)}$ di componenti principali è scelto sulla base della frazione di varianza spiegata. Si procede in modo analogo per \mathbf{f}_P , ottenendo le $K^{(P)}$ componenti principali $\xi_i^{(P)}$, i relativi autovalori $\lambda_i^{(P)}$ e punteggi $z_{ji}^{(P)}$.

Come usuale negli approcci di sorveglianza basati sulla PCA, anche in questo caso si definiscono due statistiche di controllo: la statistica T^2 di Hotelling, che identifica anomalie nelle PDF sulla base dei punteggi lungo le componenti principali, e la statistica Q , che identifica anomalie nella componente residua, fornendo un'informazione complementare a quella della statistica T^2 . Nel caso dell'insieme \mathbf{f}_S esse sono date da:

$$T_j^{2(S)} = \sum_{i=1}^{K^{(S)}} \frac{z_{ji}^{(S)2}}{\lambda_i^{(S)}} \quad (3.8)$$

e

$$Q_j^{(S)} = \|\hat{f}_{S_j} - f_{S_j}^*\|^2, \quad (3.9)$$

dove $f_{S_j}^* = \bar{f}_S + \sum_{i=1}^{K^{(S)}} z_{ji} \xi_i^{(S)}$ è la ricostruzione di \hat{f}_{S_j} ottenuta considerando le prime $K^{(S)}$ componenti principali. Dunque, per \mathbf{f}_S si ottiene una coppia di statistiche di controllo $(T^{2(S)}, Q^{(S)})$ e, procedendo allo stesso modo, anche per \mathbf{f}_P si ha la coppia di statistiche di controllo $(T^{2(P)}, Q^{(P)})$. Per determinare i limiti di controllo delle coppie di statistiche, si fissa una probabilità di falso allarme (FAP) complessiva pari a α^* . Sfruttando la correzione di Bonferroni, la FAP di ciascuna coppia di carte è data da $\alpha' = \alpha^*/2$ e, assumendo l'indipendenza tra le statistiche T^2 e Q , si ricava $\alpha = 1 - \sqrt{1 - \alpha'}$ per ciascuna delle quattro carte di controllo.

Il limite di controllo delle carte T^2 è dato da:

$$UCL_\alpha(T^2) = \frac{(N-1)^2}{N} q_{Beta(1-\alpha, K/2, (N-K-1)/2)}, \quad (3.10)$$

dove $K = K^{(S)}$, nel caso della carta T^2 relativa a \mathbf{f}_S , oppure $K = K^{(P)}$, nel caso della carta T^2 relativa a \mathbf{f}_P , e q_{Beta} è il quantile $(1-\alpha)$ della distribuzione Beta (Montgomery, 2019).

Il limite di controllo delle carte Q è dato da:

$$UCL_\alpha(Q) = \theta_1 \left\{ 1 - \theta_2 h_0 \left(\frac{1 - h_0}{\theta_1^2} \right) + \frac{\sqrt{z_\alpha(2\theta_2 h_0^2)}}{\theta_1} \right\}^{\frac{1}{h_0}}, \quad (3.11)$$

dove z_α è il quantile $(1-\alpha)$ della distribuzione normale, $h_0 = 1 - \frac{(2\theta_1\theta_3)}{3\theta_2^2}$, $\theta_r = \sum_{j=K+1}^{\infty} \lambda_j^r$, $r = 1, 2, 3$, e λ_j sono gli autovalori ottenuti dalla SFPCA di \mathbf{f}_S e \mathbf{f}_P (Zhu et al., 2007).

La carta combinata segnala un allarme quando almeno una delle quattro carte individua un pezzo fuori controllo. Data la natura bidirezionale degli insiemi \mathbf{f}_S e \mathbf{f}_P , come spiegato in precedenza, ci si aspetta che la coppia di statistiche $(T^{2(P)}, Q^{(P)})$ segnali un allarme nel caso in cui un prodotto presenti un eccesso di materiale rispetto al modello CAD, mentre nel caso di mancanza di materiale l'allarme dovrebbe essere segnalato dalla coppia di statistiche $(T^{2(S)}, Q^{(S)})$. Ciò rappresenta un semplice strumento interpretativo molto utile in fase di analisi post-diagnostica, in quanto le carte non solo individuano i pezzi fuori controllo, ma indicano anche quale tipologia di difetto si è verificata.

Dunque, la metodologia presentata da Scimone et al. (2022) ha il vantaggio di essere molto generale e adattabile a qualunque geometria arbitrariamente complessa. Infatti, questo approccio può essere applicato anche a insiemi di dati, rappresentati in forma di nuvole di punti, che non presentano una corrispondenza puntuale fra loro o sui quali non è possibile definire delle ROI (che hanno, cioè, domini diversi). Un ulteriore vantaggio è dato dal fatto che questo metodo non richiede alcuna assunzione distributiva sulle deviazioni calcolate, costituendo di fatto una generalizzazione dell'approccio descritto da Wells et al. (2013), quando l'ipotesi di normalità alla base dell'utilizzo del *Q-Q plot* è violata, come si dimostra in Menafoglio et al. (2018). La criticità maggiore, tuttavia, rimane la fase preliminare di registrazione. Sebbene l'algoritmo ICP rappresenti ormai la soluzione standard nelle procedure di allineamento di nuvole di punti 'non strutturate', i problemi relativi al costo computazionale e al raggiungimento di risultati subottimali non possono essere trascurati.

3.3 Metodo basato sullo spettro dell'operatore di Laplace-Beltrami (LB)

Il metodo proposto da Zhao e Del Castillo (2021) rappresenta un contributo del tutto innovativo nell'ambito della sorveglianza di *manifold*, che non rientra in nessuna delle due categorie di approcci descritte nel Capitolo 2. Non solo questo metodo non richiede che le ricostruzioni degli oggetti ab-

biano lo stesso numero di punti, ma evita anche la registrazione delle nuvole di punti o delle *mesh*. La metodologia è costituita da due passaggi principali, che verranno illustrati nel seguito: inizialmente si utilizzano strumenti di geometria differenziale per ricavare informazioni sulla struttura del *manifold* in esame, sintetizzate dagli autovalori di un'estensione del laplaciano, l'operatore di Laplace-Beltrami (Kreyszig, 2013); successivamente tali autovalori vengono sorvegliati attraverso una carta non parametrica multivariata. La metodologia descritta nel seguito è sviluppata per dati in formato di *mesh*, ma può essere estesa anche a dati rappresentati come nuvole di punti o come voxel.

L'idea principale consiste nello sfruttare l'assunzione nota in letteratura come 'ipotesi di *manifold*'. Come definito nel Paragrafo 1.1, un *manifold* è uno spazio curvo che, se osservato su un dominio ristretto, assomiglia a uno spazio euclideo. Si pensi, per esempio, alla Terra: essa è uno sferoide in \mathbb{R}^3 , tuttavia localmente sulla sua superficie valgono gli assiomi e i postulati della geometria euclidea. Dunque, spesso le superfici o i volumi sono *manifold* k -dimensionali (con $k = 2$ o 3 , rispettivamente) contenuti in uno spazio euclideo tridimensionale. Infatti, se si considera di nuovo la Terra, per individuare un punto sulla sua superficie sono necessarie solo due coordinate, sebbene la sua geometria sia tridimensionale. L'ipotesi di *manifold* assume che i dati contenuti in uno spazio curvo n -dimensionale, detto anche 'spazio ambientale', giacciono in realtà in uno spazio di dimensione k inferiore, chiamata anche 'dimensione intrinseca' del *manifold*. Qualunque proprietà di un *manifold* \mathcal{M} , che può essere calcolata senza ricorrere alle coordinate dello spazio ambientale, ma utilizzando solo le coordinate x^1, \dots, x^k intrinseche del *manifold*, è detta 'proprietà intrinseca' del *manifold*. Invece, qualunque proprietà geometrica di un *manifold*, che rimane invariata dopo l'applicazione di una data trasformazione, è detta 'invariante' rispetto a tale trasformazione. Le proprietà intrinseche di un *manifold* in uno spazio euclideo, sulle quali si basa la metodologia proposta da Zhao e Del Castillo (2021), sono invarianti rispetto a trasformazioni rigide, per questo motivo è possibile evitare la fase di registrazione delle nuvole di punti o delle *mesh*.

Le proprietà intrinseche di un oggetto sono spesso calcolate tramite operatori di geometria differenziale. In questa branca della matematica, a partire

dalle proprietà che caratterizzano l'intorno di un punto su una superficie, si deducono proprietà che governano l'intera struttura dell'oggetto in analisi. Si richiamano ora brevemente alcuni elementi di geometria differenziale necessari per introdurre l'operatore di Laplace-Beltrami utilizzato nella carta proposta da Zhao e Del Castillo (2021). Per ulteriori approfondimenti in merito ai concetti di geometria differenziale si rimanda a Kreyszig (2013).

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione due volte differenziabile. Il suo gradiente è definito come il vettore delle derivate parziali di f rispetto alle coordinate x_1, \dots, x_n :

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right). \quad (3.12)$$

La divergenza del vettore ∇f è definita come la somma delle derivate parziali dei suoi elementi rispetto alle coordinate x_1, \dots, x_n :

$$\operatorname{div} \nabla f = \frac{\partial}{\partial x_1} \left(\frac{\partial f}{\partial x_1} \right) + \dots + \frac{\partial}{\partial x_n} \left(\frac{\partial f}{\partial x_n} \right). \quad (3.13)$$

L'operatore di Laplace o laplaciano della funzione f è dato da:

$$\Delta f = -\operatorname{div} \nabla f = -\sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}. \quad (3.14)$$

Poiché l'operatore di Laplace è basato sulle derivate seconde parziali di f , esso costituisce di fatto un'indicazione della curvatura di f nel suo dominio, che è uno spazio euclideo n -dimensionale.

Un'estensione del laplaciano, che si utilizza nel caso in cui il dominio di f non sia \mathbb{R}^n ma un *manifold* \mathcal{M} k -dimensionale, racchiuso in uno spazio euclideo n -dimensionale, è data dall'operatore di Laplace-Beltrami. Più formalmente, per una funzione $f : \mathcal{M} \rightarrow \mathbb{R}$ l'operatore di Laplace-Beltrami è definito come:

$$\Delta_{\mathcal{M}} f = -\operatorname{div}_{\mathcal{M}} \nabla_{\mathcal{M}} f, \quad (3.15)$$

dove $\nabla_{\mathcal{M}} f$ e $\operatorname{div}_{\mathcal{M}}$ sono rispettivamente il gradiente e la divergenza calcolati sul *manifold* \mathcal{M} .

Nel presente caso di studio, la dimensione k di \mathcal{M} è pari a 2, poiché è d'interesse sorvegliare la superficie degli ovetti, mentre la dimensione n dello spazio ambientale è pari a 3, in quanto gli ovetti sono strutture tri-dimensionali. Dunque, la superficie del *manifold* può essere descritta da

$\mathbf{p}(x^1, x^2) = (x(x^1, x^2), y(x^1, x^2), z(x^1, x^2))'$. Siano \mathbf{p}_{x^1} e \mathbf{p}_{x^2} i vettori delle derivate parziali degli elementi di $\mathbf{p}(x^1, x^2)$ rispetto alle coordinate x^1 e x^2 . Sia \mathbf{g} la matrice:

$$\begin{pmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{pmatrix},$$

dove $g_{11} = \langle \mathbf{p}_{x^1}, \mathbf{p}_{x^1} \rangle$, $g_{12} = \langle \mathbf{p}_{x^1}, \mathbf{p}_{x^2} \rangle$ e $g_{22} = \langle \mathbf{p}_{x^2}, \mathbf{p}_{x^2} \rangle$. L'operatore di Laplace-Beltrami di una funzione $f(x^1, x^2)$, nel seguito abbreviato con LB, è dato da:

$$\Delta_{\mathcal{M}} f = -\frac{1}{\sqrt{\det(\mathbf{g})}} \sum_{j=1}^2 \frac{\partial}{\partial x^j} \left(\sqrt{\det(\mathbf{g})} \sum_{i=1}^2 g^{ij} \frac{\partial f}{\partial x^i} \right), \quad (3.16)$$

dove g^{ij} sono gli elementi di \mathbf{g}^{-1} . L'operatore di LB è una misura intrinseca della curvatura locale di una funzione, che tiene in considerazione non solo la curvatura dello spazio ambientale ma anche quella del *manifold*.

Tuttavia, spesso nella pratica, come nel caso di studio, non si dispone di una funzione f che descriva il *manifold*, bensì di una nuvola di punti o di una *mesh*, che costituiscono una versione discretizzata del *manifold*. Pertanto, la funzione continua f definita su \mathcal{M} si riduce a un vettore, i cui elementi sono i valori della funzione valutata nei punti rilevati, e l'operatore LB viene approssimato da una sua versione discreta. L'approssimazione utilizzata da Zhao e Del Castillo (2021) è quella proposta da Li et al. (2015) per dati in forma di *mesh* ed è definita nel modo seguente.

Sia \mathcal{M} un *manifold* su cui vengono rilevati m punti e $f(p_i)$, $i = 1, \dots, m$ la funzione f valutata nei punti rilevati. L'operatore discreto di Laplace-Beltrami è una matrice L^t , di dimensione $m \times m$, i cui elementi sono dati dai valori dell'operatore calcolato in $f(p_i)$, $i = 1, \dots, m$:

$$L^t f(p_i) = \frac{1}{4\pi t^2} \sum_{p_j: d_{geo}(p_i, p_j) \leq r} \frac{A(p_j)}{3} e^{-\frac{d_{geo}(p_i, p_j)^2}{4t}} [f(p_i) - f(p_j)], \quad j = 1, \dots, m, \quad (3.17)$$

dove $d_{geo}(\cdot, \cdot)$ è la distanza geodesica, r è un valore prefissato, $A(p_j)$ è l'area dell'intorno del punto p_j e t è un parametro legato all'ampiezza dell'area di interesse intorno al punto p_i (al crescere di t , aumenta l'area considerata). La (3.17) è un'approssimazione 'locale' dell'operatore discreto di Laplace-Beltrami, in quanto per calcolare l'operatore in $f(p_i)$ vengono considerati

solo i punti p_j , $j = 1, \dots, m$, che si trovano entro un certo raggio r di distanza da p_i . In questo modo è possibile ridurre lo spazio di memoria e il costo computazionale necessari per il calcolo dell'operatore LB, in quanto la matrice risultante è sparsa. È interessante notare che la distanza geodesica (si veda il Paragrafo 2.1.1) è una misura di distanza intrinseca e invariante rispetto a trasformazioni rigide, a differenza di quella euclidea, che invece è un esempio di distanza invariante rispetto a trasformazioni rigide ma non intrinseca.

Le informazioni contenute nell'operatore LB sono ben sintetizzate dal suo spettro, cioè dalla sequenza dei suoi autovalori $\{\lambda_i\}_{i=0}^{\infty}$, che riassumono le proprietà geometriche e topologiche del *manifold* (Reuter et al., 2009). Tali autovalori sono crescenti e non negativi (Chavel, 1984) e si ottengono come soluzione del problema agli autovalori:

$$\Delta_{\mathcal{M}}f = \lambda f, \quad (3.18)$$

dove $\Delta_{\mathcal{M}}f$ è stato definito nella (3.16). Anche il vero spettro dell'operatore LB, così come la funzione f che descrive il *manifold*, non è noto nella maggior parte dei casi. Si ricorre, pertanto, a una sua approssimazione, che è data semplicemente dalla decomposizione in autovalori e autovettori della matrice L^t dell'operatore discreto di Laplace-Beltrami (3.17). Si dimostra che il calcolo dello spettro approssimato è numericamente stabile (Patanè, 2017). Inoltre, gli studi di simulazione condotti da Zhao e Del Castillo (2021) mostrano che solo la parte inferiore dello spettro è utile ai fini dell'implementazione di una procedura di sorveglianza statistica, pertanto non è necessario ottenere tutti gli m autovalori. Il costo computazionale relativo al calcolo dei primi K autovalori è sostenibile ed è pari a $O(K \cdot m)$ nei più recenti algoritmi (per esempio, nell'algoritmo di Arnoldi (Arnoldi, 1951) implementato nella funzione `eigs()` di Matlab).

Una volta ottenuta la parte inferiore dello spettro dell'operatore discreto di Laplace-Beltrami per ogni pezzo prodotto, gli autovalori devono essere sorvegliati attraverso una procedura non parametrica, in quanto la loro distribuzione è non-normale (Zhao e Del Castillo, 2021). A tale scopo gli autori utilizzano una versione modificata della carta *self-starting* proposta da Chen et al. (2016). Le carte *self-starting* (Qiu, 2014) spesso necessitano di un insieme di osservazioni pre-campionarie piuttosto elevato (circa 50, nelle carte

di Zhao e Del Castillo (2021) e Chen et al. (2016)) e hanno difficoltà nell'individuare osservazioni fuori controllo, se la situazione di non conformità si verifica negli istanti iniziali della sorveglianza. Pertanto, Zhao e Del Castillo (2021) suggeriscono di affiancare una metodologia di Fase I come procedura complementare alla carta *self-starting* da loro sviluppata. A questo proposito, gli autori scelgono di utilizzare il metodo proposto da Capizzi e Masarotto (2017), poiché esso rappresenta un punto di riferimento tra le carte multivariate non parametriche di Fase I. È importante sottolineare che la procedura di Capizzi e Masarotto (2017) è stata sviluppata nel contesto tradizionale della sorveglianza retrospettiva di dati multivariati, pertanto la sua applicazione nell'ambito della sorveglianza di dati complessi, come i *manifold*, rappresenta in questo elaborato un potenziale contributo innovativo. Inoltre, nella presente trattazione, tale metodologia viene impiegata con una finalità diversa rispetto a quanto fatto da Zhao e Del Castillo (2021). Infatti, qui, la carta di Capizzi e Masarotto (2017) viene utilizzata come procedura di Fase I a sé stante, mentre Zhao e Del Castillo (2021) scelgono tale carta come strumento di supporto alla carta *self-starting* nei primi stadi della sorveglianza.

Nel seguito si presenta brevemente la metodologia di Capizzi e Masarotto (2017), adattata al caso di studio in esame. Sia x_{ji} , l' i -esimo autovalore, $i = 1, \dots, K$, relativo al j -esimo pezzo prodotto, $j = 1, \dots, N$. Si assume $K < N$. Sui dati standardizzati si calcolano i ranghi con segno multivariati (Oja, 2010), indicati dai vettori K -dimensionali \mathbf{u}_j , $j = 1, \dots, N$. Si tratta di una trasformazione dei dati in grado di riflettere sia la direzione sia l'ampiezza della distanza di un'osservazione dal centro dell'insieme di dati da cui questa proviene. Successivamente, si adatta ai ranghi con segno multivariati un modello di regressione lineare:

$$\mathbf{u}_j = \boldsymbol{\beta}_{comune} + \sum_{\tau=2}^{N-1} \boldsymbol{\beta}_{persistente,\tau} I(j \geq \tau) + (\text{residuo})_j, \quad (3.19)$$

dove i $\boldsymbol{\beta}$ sono parametri K -dimensionali ignoti e $I(\cdot)$ è la funzione indicatrice. Il vettore $\boldsymbol{\beta}_{comune}$ rappresenta il livello medio dei ranghi con segno, mentre $\boldsymbol{\beta}_{persistente,\tau}$ introduce un cambiamento di livello a partire dall'istante di tempo τ e che coinvolge tutte le osservazioni successive. Attraverso un algoritmo di ricerca passo a passo in avanti si selezionano, tra gli $N - 1$ parametri, gli

$H < (N - 1)$ parametri associati agli istanti di tempo in cui più plausibilmente si sono verificati dei cambiamenti, sulla base di quanto suggerito dai dati. La carta aggrega poi le H statistiche di controllo elementari, definite per individuare la presenza di $1, \dots, H$ cambiamenti persistenti, e viene calcolato un p -value complessivo per testare la stabilità globale del processo tramite un approccio basato sulle permutazioni (Pesarin, 2001). Quando viene segnalato un allarme, la procedura individua gli autovalori che hanno portato alla condizione di non conformità attraverso l'algoritmo LASSO adattivo (Zou, 2006).

Infine, Zhao e Del Castillo (2021) propongono di visualizzare i difetti sul prodotto stesso, così da poter risalire alla tipologia di anomalia che ha interessato il prodotto. Per fare ciò gli autori utilizzano l'algoritmo ICP per allineare la *mesh* del pezzo fuori controllo a quella del modello nominale, al fine di identificare le zone in cui sono presenti delle discrepanze. L'uso dell'algoritmo ICP, in questo contesto, è computazionalmente sostenibile, poiché viene applicato in fase post-diagnostica e solo ai prodotti non conformi. Dunque, la metodologia proposta da Zhao e Del Castillo (2021) ha il merito di aver introdotto nell'ambito dell'SPC di *manifold* un nuovo tipo di approccio, che si basa su proprietà intrinseche e locali degli oggetti, opportunamente colte dall'operatore LB. L'approssimazione discreta di esso, infatti, è in grado di sfruttare le informazioni di adiacenza contenute nella rappresentazione in formato di *mesh* dei prodotti di AM.

Capitolo 4

Implementazione delle procedure di sorveglianza per geometrie complesse

In questo Capitolo le metodologie proposte da Scimone et al. (2022) e da Zhao e Del Castillo (2021) vengono applicate al caso di studio riguardante gli ovetti. L'insieme di dati è costituito da 13 osservazioni, di cui 5 fuori controllo. Come descritto nel Paragrafo 3.1, l'ovetto caratterizzato da una mancanza di materiale (il nono) è stato ottenuto artificialmente, rimuovendo un montante da un ovetto in-controllo, mentre i quattro caratterizzati da un eccesso di materiale (dal decimo al tredicesimo) sono stati prodotti utilizzando dei filamenti polimerici di colore blu, arancione, bianco e rosso, che hanno causato difetti di intensità variabile. Il presente caso di studio costituisce un esempio paradigmatico di geometrie complesse derivanti da una produzione di tipo *short run*. L'obiettivo di questa applicazione è quello di valutare l'adeguatezza di due differenti approcci di sorveglianza in tale contesto.

Il metodo di Scimone et al. (2022) è stato sviluppato proprio nell'ambito della sorveglianza di strutture completamente *free-form* e applicato a ovetti analoghi a quelli descritti in precedenza. Tuttavia, nel sopracitato articolo, gli ovetti fuori controllo presi in considerazione sono solo due: quello privo di un montante e quello caratterizzato da un eccesso di materiale più accentuato (ovetto blu). Si tratta di difetti isolati e molto evidenti anche a occhio nudo.

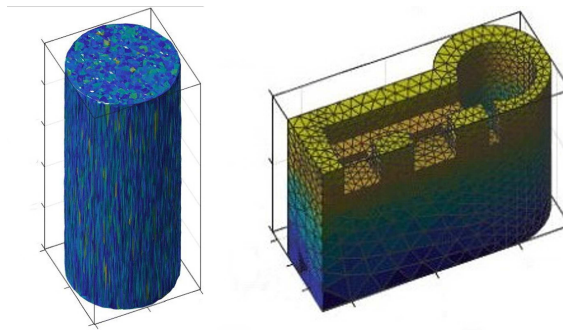


Figura 4.1: *Manifold* studiati da Zhao e Del Castillo (2021)

In questo Capitolo verrà valutata la *performance* della carta di Scimone et al. (2022) in presenza di più difetti di varie intensità.

Il metodo di Zhao e Del Castillo (2021) è stato implementato, in particolare, in una situazione diversa da quella qui considerata. Gli autori applicano la loro procedura a *manifold* meno complessi rispetto agli ovetti, come quelli rappresentati in Figura 4.1, assumono di avere a disposizione un cospicuo numero di unità statistiche e si concentrano prevalentemente sulla sorveglianza *online* dei prodotti. Al contrario, in questo Capitolo la metodologia da loro proposta verrà applicata a forme geometriche *free-form* (gli ovetti) in un contesto di bassi volumi di produzione e di sorveglianza *offline*.

In questo modo è possibile eseguire un confronto, non ancora presente in letteratura, tra l'approccio di Zhao e Del Castillo (2021) e la metodologia proposta da Scimone et al. (2022) nell'ambito della sorveglianza di geometrie complesse.

4.1 Sorveglianza dei profili di deviazione tramite SFPCA

Nel seguito vengono riportati i risultati relativi all'applicazione della metodologia presentata da Scimone et al. (2022) al caso di studio in esame. Le analisi sono state eseguite utilizzando il software R e il codice riportato in Appendice A.1.

I dati sono contenuti in file di tipo *Standard Triangulation Language* (STL), che è un formato di file utilizzato nei software per il disegno digitale, come CAD, per rappresentare una *mesh*. Le analisi descritte in questo Paragrafo si basano su una rappresentazione dei dati in forma di nuvola di punti, pertanto vengono mantenuti solo i vertici delle *mesh*. Il numero di punti di ciascuna nuvola è pari a circa 15 000 e può variare leggermente tra un ovetto e l'altro.

Poiché i dati a nostra disposizione sono già stati precedentemente registrati dagli autori dell'articolo, si procede direttamente al calcolo dei vettori di deviazioni, D_P^j e D_{S_j} , $j = 1, \dots, 13$, tra il modello nominale e gli ovetti prodotti. Vengono calcolate sia la mappa d_P^j , che misura la distanza dei punti del j -esimo ovetto dal modello nominale P , sia la mappa d_{S_j} , che misura la distanza dei punti del modello nominale P dal j -esimo ovetto, $j = 1, \dots, 13$. Come è ragionevole aspettarsi, i vettori di deviazioni presentano molti elementi vicini o pari a zero. Pertanto, le deviazioni vengono trasformate in scala logaritmica, in modo da ridurre l'asimmetria della loro distribuzione.

Le PDF dei vettori di deviazioni, \hat{f}_P^j e \hat{f}_{S_j} , $j = 1, \dots, 13$, si ottengono a partire dalle funzioni di ripartizione empiriche di D_P^j e D_{S_j} . Le PDF vengono stimate attraverso lo stimatore di Bernstein (si veda la (3.4)) e valutate su un insieme discreto di punti appartenenti al loro supporto.

Il calcolo della SFPCA, così come definito nelle (3.6)–(3.7), è lungo e complesso, poiché le PDF appartengono allo spazio di Bayes B^2 . Tuttavia, Hron et al. (2016) dimostrano che calcolare la SFPCA è equivalente a calcolare la FPCA su un'opportuna trasformazione delle PDF, chiamata 'trasformazione log-rapporto centrato' (clr).

Sia $f(t)$ una densità con supporto $[a, b]$. La trasformata clr di $f(t)$ è definita come:

$$[\text{clr}(f)](t) = \log(f(t)) - \frac{1}{b-a} \int_a^b \log(f(\tau)) d\tau, \quad t \in [a, b]. \quad (4.1)$$

La (4.1) permette di trasformare le densità, che sono elementi dello spazio B^2 , in elementi dello spazio delle funzioni reali due volte integrabili, L^2 , in cui è possibile effettuare la FPCA tradizionale. La trasformazione clr (4.1) viene calcolata a partire dalle PDF stimate al passo precedente e valutata su un insieme discreto di punti del supporto delle PDF.

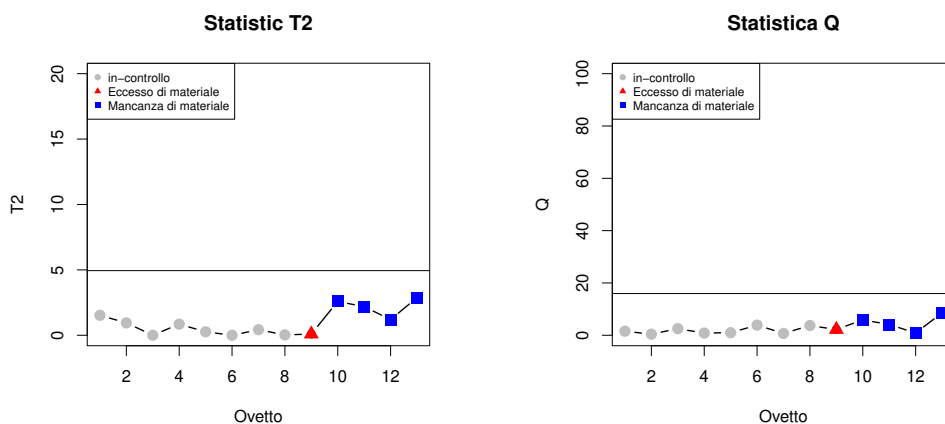


Figura 4.2: Carte di controllo relative alle funzioni di densità \hat{f}_P^j

Le componenti principali delle PDF, così trasformate, si ottengono dalla decomposizione spettrale della loro matrice di covarianza (Hron et al., 2016). Il numero di componenti principali è selezionato in modo tale da spiegare il 98% della variabilità totale, come indicato da Menafoglio et al. (2018). In questo caso, è sufficiente una sola componente principale (la prima). Le statistiche di controllo T^2 e Q e i rispettivi limiti di controllo $UCL_\alpha(T^2)$ e $UCL_\alpha(Q)$ sono calcolati come nelle (3.8)–(3.11), fissando una FAP complessiva $\alpha^* = 0.05$.

In Figura 4.2 sono riportate le carte di controllo relative alla coppia di statistiche $(T^{2(P)}, Q^{(P)})$ e i rispettivi limiti di controllo, rappresentati dalle linee nere orizzontali. Si ricorda che tali statistiche di controllo sono calcolate a partire dalla SFPCA delle funzioni di densità \hat{f}_P^j , $j = 1, \dots, 13$, che sintetizzano le informazioni relative alle distanze dei punti del j -esimo ovetto dal modello CAD. Dunque, queste carte di controllo dovrebbero essere in grado di individuare osservazioni che presentano non conformità dovute a un eccesso di materiale. Tuttavia, in questo caso, nessuna delle due carte segnala un allarme.

Tale risultato è dovuto al fatto che, in generale, le carte di tipo Shewhart non sono in grado di cogliere cambiamenti persistenti e di piccole intensità, come quelli considerati in questo caso di studio. Infatti, ciascuna statistica T_j^2 è costruita sulla base delle informazioni relative alla j -esima osservazione,

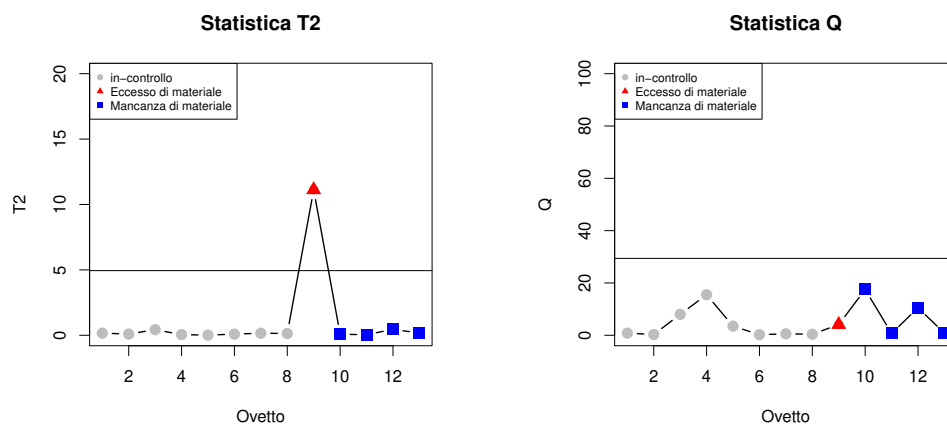


Figura 4.3: Carte di controllo relative alle funzioni di densità \hat{f}_{S_j}

pertanto è sensibile solo a cambiamenti di tipo isolato. Inoltre, poiché la statistica T^2 è basata su una quantità (in questo caso, i punteggi lungo le componenti principali) elevata al quadrato, essa è adatta per individuare difetti di grandi dimensioni, ma non cambiamenti di intensità ridotte.

In Figura 4.3 si riportano le carte di controllo relative alla coppia di statistiche $(T^{2(S)}, Q^{(S)})$, che sono calcolate a partire dalla SFPCA delle funzioni di densità \hat{f}_{S_j} , $j = 1, \dots, 13$. Poiché queste PDF sintetizzano le informazioni relative alle distanze dei punti del modello nominale dal j -esimo ovetto, le rispettive carte di controllo dovrebbero essere in grado di individuare osservazioni anomale caratterizzate da una mancanza di materiale. In questo caso, la carta segnala correttamente un allarme in corrispondenza dell'ovetto privo di un montante. Si tratta, infatti, di un difetto isolato e di grandi dimensioni, che la statistica T^2 è in grado di cogliere.

4.2 Sorveglianza non parametrica dello spettro dell'operatore LB discreto

Nel seguito è descritta l'applicazione della metodologia proposta da Zhao e Del Castillo (2021) al caso di studio in esame. In Appendice A.2 sono riportati i codici Matlab e R necessari all'implementazione della procedura.

In questo caso, ogni oggetto è rappresentato da una *mesh* con facce triangolari costituita da $m \simeq 15\,000$ vertici. Per ogni oggetto si calcola l'operatore discreto di Laplace-Beltrami proposto da Li et al. (2015) e definito nella (3.17). La matrice L^t di dimensione $m \times m$, contenente l'operatore LB discreto valutato nei punti p_i , $i = 1, \dots, m$, è ottenuta come:

$$L^t = D - W, \quad (4.2)$$

dove $W_{ij} = \frac{1}{12\pi t^2} A(p_j) e^{d_{geo}(p_i, p_j)^2/4t}$ e D è una matrice diagonale con elementi $D_{ii} = \sum_j W_{ij}$ (Zhao e Del Castillo, 2021).

La matrice L^t può essere rappresentata come il prodotto di due matrici tali che:

$$L^t = B^{-1}C, \quad (4.3)$$

dove B è una matrice diagonale e C è una matrice simmetrica semidefinita positiva (Patanè, 2017). Gli elementi sulla diagonale della matrice B sono dati dalle aree $A(p_i)$, $i = 1, \dots, m$, degli intornoi dei punti (Zhao, 2022).

La (4.3) implica che L^t è invertibile e garantisce che gli autovalori di L^t siano reali. Inoltre, si dimostra che $B^{-1}C$ ha gli stessi autovalori della matrice simmetrica $B^{-1/2}CB^{-1/2}$ (Zhao, 2022). Poiché dalla (4.3) si ricava che $C = BL^t$, si ha che trovare gli autovalori di $B^{-1}C$ è equivalente a calcolare gli autovalori della matrice simmetrica:

$$L_{sym} = B^{1/2}L^tB^{-1/2}. \quad (4.4)$$

Pertanto, è possibile effettuare la decomposizione spettrale della (4.4) anziché della (4.3). Questo permette di sfruttare i vantaggi computazionali derivanti dalla simmetria di L_{sym} per ottenere gli autovalori della matrice L^t (Zhao, 2022).

Dunque, la parte inferiore dello spettro dell'operatore LB discreto viene calcolata a partire dalla matrice L_{sym} . Attualmente non esiste un criterio per determinare il numero di autovalori da calcolare e da utilizzare nella sorveglianza. A causa di errori di approssimazione numerica, è possibile che i primi autovalori presentino valori negativi molto vicini a zero, pertanto non sono da considerare nelle analisi successive (Zhao e Del Castillo, 2021). In questo caso di studio vengono calcolati i primi 30 autovalori, tuttavia i primi 15 vengono rimossi, poiché assumono valori nell'ordine di 10^{-22} .

Una volta ottenuta la parte inferiore dello spettro dell'operatore LB per ogni ovetto, si applica la procedura di sorveglianza proposta da Capizzi e Masarotto (2017). Si ricorda che il numero di autovalori considerati si assume minore del numero di osservazioni disponibili (Paragrafo 3.3). Data l'esigua numerosità campionaria degli ovetti, il numero di autovalori qui presi in analisi è compreso tra 2 e 5.

La *performance* della carta viene valutata anche al variare della lunghezza minima di un cambiamento di tipo persistente (se presente). In questo contesto, poiché il campione è piccolo, sembra ragionevole ipotizzare la presenza di cambiamenti anche molto brevi, pertanto si assume una lunghezza minima adeguata, pari a 1, 2 o 3.

In Tabella 4.1 si riportano i risultati ottenuti dall'applicazione della procedura di Capizzi e Masarotto (2017) al caso di studio in esame, fissando una FAP $\alpha = 0.05$. I risultati suggeriscono che 2, 3 o 4 autovalori sono sufficienti per individuare correttamente tutti gli ovetti fuori controllo caratterizzati da un eccesso di materiale (ovetti dal decimo al tredicesimo). In tutti questi scenari, la carta indica che il cambiamento riguarda il secondo autovalore.

Si osservi che, una volta segnalato un allarme, le osservazioni identificate come fuori controllo vengono rimosse e la procedura viene ripetuta. Il *p-value* complessivo relativo alla stabilità globale del processo indica, correttamente, che il processo è ora in-controllo.

Se si utilizzano 5 autovalori, la carta segnala erroneamente un allarme in corrispondenza del quarto ovetto, che in realtà è in-controllo. Questo risultato è in accordo con quanto indicato da Zhao e Del Castillo (2021), i quali, sulla base di studi di simulazione, mostrano che considerare un numero elevato di autovalori è controproducente in termini di *performance* della carta. Ciò è dovuto al fatto che solo la parte inferiore dello spettro dell'operatore LB discreto contiene informazioni rilevanti relative alle proprietà geometriche del prodotto, mentre gli autovalori più alti contengono anche del 'rumore' (Reuter et al., 2009).

L'ovetto privo di un montante (il nono) viene segnalato solo se si considerano 3 o 4 autovalori e una lunghezza minima del cambiamento pari a 3. Tale osservazione risulta più difficile da individuare, dato che, a differenza delle altre, presenta un difetto molto localizzato e aggiunto artificialmente

Numero di autovalori	Lunghezza minima del cambiamento	Istante del cambiamento	Livello di significatività
2	1	10	0.040
	2	10	0.016
	3	–	0.075
3	1	10	0.021
	2	10	0.015
	3	9	0.045
4	1	10	0.010
	2	10	0.008
	3	9	0.027
5	1	4 e 10	0.012
	2	4 e 10	0.006
	3	4 e 9	0.025

Tabella 4.1: Risultati della sorveglianza dei 13 ovetti in base al numero di autovalori considerati e alla lunghezza minima del cambiamento persistente (se presente). Sono indicati l'istante (ovetto) in cui ha inizio il cambiamento, che riguarda tutte le osservazioni successive fino al termine del processo o fino a che non viene individuato un altro cambiamento, e il *p-value* relativo al test di stabilità globale del processo. In grassetto sono evidenziate le *performance* migliori e '–' indica che non è stato segnalato alcun allarme dalla procedura.

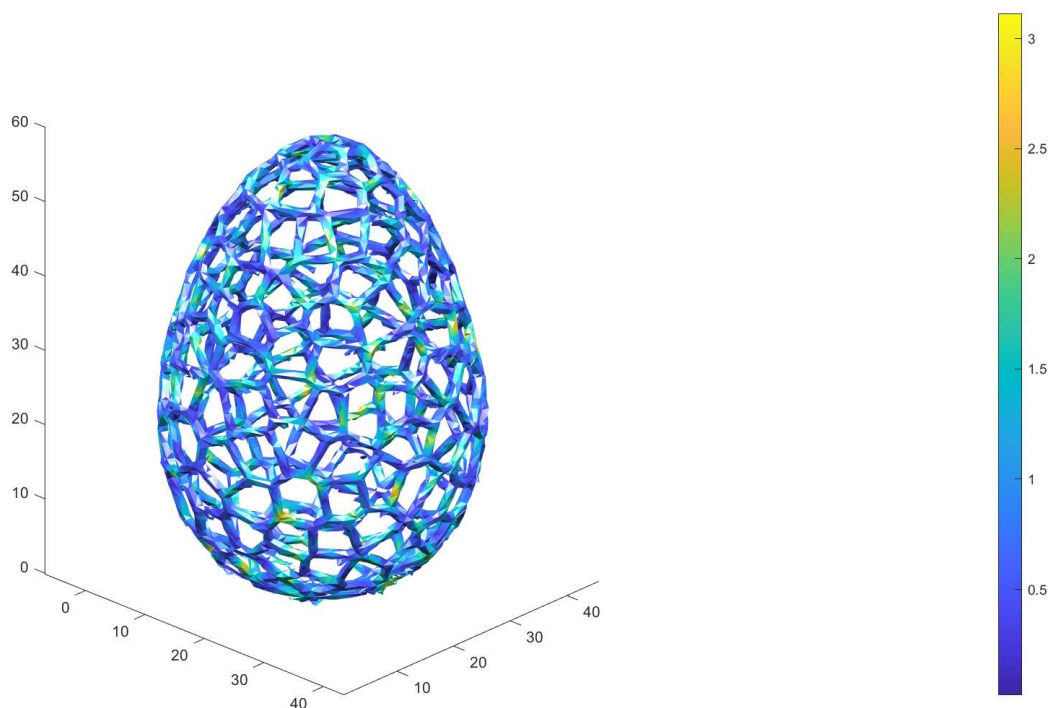


Figura 4.4: Analisi post-diagnostica del decimo ovetto. Le zone gialle indicano la presenza di un eccesso di materiale nel pezzo prodotto rispetto al modello nominale, mentre le zone blu indicano che non sono presenti discrepanze.

a partire da un ovetto in-controllo. Infatti, come evidenziano anche Zhao e Del Castillo (2021), la presenza di un unico difetto isolato non altera in modo sostanziale la struttura geometrica complessiva del prodotto, pertanto lo spettro dell'operatore LB subisce cambiamenti di intensità minori, che per la carta sono più critici da segnalare.

In Figura 4.4 è mostrato il grafico risultante dall'analisi post-diagnostica del decimo ovetto, eseguita tramite l'algoritmo ICP. I colori riportati sull'ovetto rispecchiano l'intensità delle discrepanze tra il pezzo prodotto e il modello CAD. Grazie a tale grafico è possibile individuare in modo molto intuitivo la presenza di difetti diffusi.

Conclusioni

La continua evoluzione dei processi produttivi, soprattutto nel contesto dell'AM, ha reso possibile la realizzazione di prodotti caratterizzati da forme sempre più libere e complesse. Tuttavia, i pezzi ottenuti via AM sono spesso soggetti a difetti che difficilmente possono essere individuati dai metodi di sorveglianza tradizionali. Perciò, una delle principali sfide dell'SPC in questo ambito è data dalla necessità di definire nuove procedure per la modellazione e il controllo della qualità di geometrie complesse.

Nel presente elaborato sono state descritte due metodologie di sorveglianza innovative, che si distinguono dagli altri approcci esistenti in letteratura in quanto sono applicabili a oggetti effettivamente tridimensionali, che possono assumere forme anche molto irregolari. Inoltre, questi metodi sono computazionalmente in grado di gestire l'elevata dimensionalità dei dati relativi ai prodotti di AM, che solitamente provengono da sistemi di sensori senza contatto, i quali permettono di ricostruire gli oggetti sotto forma di nuvole di punti o di *mesh*.

Nella metodologia proposta da Scimone et al. (2022) ogni prodotto è rappresentato da una coppia di vettori di deviazioni, che descrivono la natura bidirezionale della distanza di Hausdorff tra la nuvola di punti del prodotto ricostruito e quella del modello nominale. Le informazioni contenute in tali vettori, che sono costituiti da migliaia di punti, sono sintetizzate dalle loro funzioni di densità stimate. Su di esse viene applicata una tecnica di riduzione della dimensionalità, la SFPCA, che permette di individuare le principali fonti di variabilità presenti nei dati. I punteggi lungo le componenti principali, riscaldati per i loro autovalori, sono sorvegliati con la statistica T^2 di Hotelling, mentre la componente residua è sorvegliata dalla statistica Q .

Nella metodologia proposta da Zhao e Del Castillo (2021), invece, ogni prodotto è rappresentato in base alle sue proprietà geometriche intrinseche, che dipendono solo dalle caratteristiche locali del *manifold*. Tali proprietà sono sintetizzate dalla parte inferiore dello spettro dell'operatore discreto di Laplace-Beltrami (LB), che costituisce un'estensione del concetto di laplaciano. L'operatore LB sfrutta le informazioni di adiacenza, contenute negli archi e nelle facce della *mesh* del prodotto, per fornire una misura della curvatura locale della sua superficie sulla base dei punti rilevati. Gli autovalori dell'operatore LB sono sorvegliati tramite una procedura non parametrica multivariata, che combina diverse statistiche di controllo elementari, definite per identificare la presenza di uno o più cambiamenti nel processo.

Nel presente elaborato, le metodologie di Scimone et al. (2022) e Zhao e Del Castillo (2021) sono implementate e applicate a un caso di studio rappresentativo delle sfide poste dall'AM. Tale applicazione riguarda la sorveglianza di geometrie complesse in un contesto di produzione di tipo *short run*. Per la prima volta in letteratura, per quanto noto allo stato attuale, viene confrontata la *performance* dei due metodi. L'approccio basato sulla SFPCA risulta adeguato solo nel caso di difetti isolati e di grande intensità. Al contrario, la procedura non parametrica basata sull'operatore LB è in grado di individuare cambiamenti persistenti, anche di piccole intensità, purché non siano eccessivamente localizzati. Dunque, entrambi i metodi permettono di gestire in modo opportuno la complessità geometrica dei prodotti e l'elevata dimensionalità dei dati e sono applicabili anche in contesti di bassi volumi di produzione. Per quanto sia difficile stabilire in modo univoco quale delle due metodologie sia da preferire in termini di *performance*, viste le diverse applicazioni, è tuttavia possibile osservare che entrambe le procedure presentano sia punti di forza sia aspetti critici.

I vantaggi dell'approccio proposto da Scimone et al. (2022) sono dati dall'interpretabilità della natura bidirezionale della distanza di Hausdorff, legata all'eccesso o alla mancanza di materiale, e dalla familiarità del metodo basato sull'analisi dei profili di deviazione, che è ben noto nell'ambito dell'SPC di forme geometriche. Il principale limite è dato dalla fase preliminare di registrazione delle nuvole di punti tramite l'algoritmo *Iterative Closest Point* (ICP), che può risultare onerosa e computazionalmente instabile. Viceversa,

il metodo proposto da Zhao e Del Castillo (2021) costituisce un approccio del tutto innovativo, che evita la procedura di allineamento delle *mesh*, poiché si basa su proprietà geometriche intrinseche e locali dei prodotti. Qui l'algoritmo ICP viene utilizzato solo in fase di analisi post-diagnostica per individuare in modo semplice e intuitivo i difetti sui prodotti. Tuttavia, un aspetto critico di questo metodo riguarda la selezione del numero di autovalori da considerare nella sorveglianza, in quanto non esiste un criterio univoco per tale scelta.

Appendice A

Codice utilizzato

A.1 Sorveglianza dei profili di deviazione tramite SFPCA

In [A.1.1](#) sono definite delle funzioni ausiliarie, necessarie per la stima delle PDF dei vettori di deviazioni, in [A.1.2](#) sono implementati i passaggi della procedura di sorveglianza, descritta nel Paragrafo [3.2](#), e in [A.1.3](#) è presente il codice per riprodurre i grafici delle carte di controllo.

A.1.1 Funzioni ausiliarie

```
# Stimatore polinomiale di Bernstein della PDF valutata in x,  
# data la funzione di ripartizione empirica Fn e il supporto [a,b]  
fd.01 = function(Fn, x, a=0, b=1) {  
  n=length(knots(Fn))  
  m=n  
  y=scInv.01(x,a,b)  
  ret=rep(0,length(y))  
  for(k in 0:(m-1)) {  
    bp=choose(m-1,k)*(y^k)*(1-y)^(m-1-k)  
    ret=apply(rbind(ret,(Fn(sc.01((k+1)/m,a,b))-Fn(sc.01(k/m,a,b)))*bp),2,sum)  
  }  
  ret=m*ret*dscInv.01(x,a,b)
```

```
    return(ret)
}

# Trasformata clr di una densita', dati i punti z del supporto e la
# lunghezza z_step del supporto
clr = function(density, z, z_step) {
  return(log(density)-trapzc(z_step,log(density))/(max(z)-min(z)))
}

# Valore dell'integrale di y nell'intervallo di lunghezza t_step
trapzc = function(t_step,y) {
  return(t_step*(0.5*y[1]+sum(y[2:(length(y)-1)]) +0.5*y[length(y)]))
}

# Supporto in [0,1] -> supporto in [a,b]
sc.01=function(y,a,b) {
  if(a==-Inf & b==Inf)
    return(tan(pi*y-pi/2))
  if(a==-Inf & b!=Inf)
    return(((1+b)*y-1)/y)
  if(a!=Inf & b!=Inf)
    return((b-a)*y+a)
  print('Cannot compute the transformation, support incorrect')
  return(-1)
}

# Supporto in [a,b] -> supporto in [0,1]
scInv.01=function(x,a,b) {
  if(a==-Inf & b==Inf)
    return(.5+atan(x)/pi)
  if(a==-Inf & b!=Inf)
    return(1/(1+b-x))
  if(a!=Inf & b!=Inf)
    return((x-a)/(b-a))
  print('Cannot compute the transformation, support incorrect')
  return(-1)
}
```

```
}

# Derivata di una trasformazione inversa dy/dx
dscInv.01=function(x,a,b) {
  if(a==-Inf & b==Inf)
    return(1/(pi*(1+x^2)))
  if(a==-Inf & b!=Inf)
    return(1/(1+b-x)^2)
  if(a!=Inf & b!=Inf)
    return(1/(b-a))
  print('Cannot compute the transformation, support incorrect')
  return(-1)
}

# Operazione di mediazione
mean_batch = function(v,batch) {
  ret = rep(0,ceiling(length(v)/batch))
  for (i in 1:length(ret))
    ret[i]= mean(v[((i-1)*batch+1):min((i*batch),length(v))]);
  return(ret);
}
```

A.1.2 Procedura di sorveglianza basata sulla SFPCA

La libreria `Rvcg` permette di leggere e caricare in R i file `.stl`, in cui sono contenuti i dati relativi agli ovetto, attraverso la funzione `vcgImport()`. Tale funzione crea oggetti della classe `mesh3d`, che sono descritti da i) una matrice contenente le coordinate dei vertici e ii) una matrice contenente i vertici che formano le facce triangolari della *mesh*.

La funzione `vcgMetro()` calcola le distanze tra i vertici di una coppia di oggetti `mesh3d`, che costituiscono delle nuvole di punti. Le deviazioni del modello nominale dal j -esimo ovetto sono contenute nel vettore `distances1`, mentre le deviazioni del j -esimo ovetto dal modello nominale sono contenute in `distances2`, $j = 1, \dots, 13$.

```
library(Rvcg)
source("funzioni_auxiliarie.R")

### Importo tutte le uova necessarie al lavoro
Ndata = 14
eggs=list()
for(i in 1:Ndata) {
  eggs[[i]] = vcgImport(paste0("dati/uovo",i,".stl"), silent = T)
}

### Calcolo tutte le distanze rispetto al modello nominale
confronto = eggs[[14]] #fisso come confronto l'uovo nominale
distanze_f = list()
distanze_b = list()
for (i in 1:(Ndata-1)) {
  distanze_f[[i]] = vcgMetro(confronto, eggs[[i]], silent = F)$distances1
  distanze_b[[i]] = vcgMetro(confronto, eggs[[i]], silent = F)$distances2
}
```

Inizialmente si analizzano le distanze d_{S_j} dal modello nominale al j -esimo ovetto, $j = 1, \dots, 13$, e si eseguono delle operazioni preliminari per la stima delle PDF \hat{f}_{S_j} dei vettori di deviazioni D_{S_j} . In particolare, i) si definiscono gli estremi del supporto delle PDF, ii) si ordinano le deviazioni in senso crescente e si applica loro la trasformata logaritmica e iii) si seleziona un numero ridotto di elementi dei vettori di deviazioni attraverso un'operazione di mediazione, implementata dalla funzione ausiliaria `mean_batch()`.

Questi elementi vengono utilizzati nel calcolo delle funzioni di ripartizione empiriche dei vettori D_{S_j} , $j = 1, \dots, 13$, che sono determinate tramite la funzione `ecdf()` della libreria `stats`. A partire da esse si stimano le PDF dei vettori di deviazioni attraverso la funzione ausiliaria `fd.01()`. Essa richiede come *input* i) la funzione di ripartizione empirica di un vettore di deviazioni, ii) un vettore di punti in cui valutare la PDF e iii) gli estremi del supporto della PDF e restituisce come *output* un vettore contenente la PDF del vettore di deviazioni, stimata attraverso lo stimatore polinomiale di Bernstein (3.4) e valutata sui punti dati in *input*.

```
# Estremo inferiore (la deviazione minima e' 1e-09, ma valori cosi'
# piccoli creano problemi numerici nei passaggi successivi)
cutleftacq = 0.0025
# Estremo superiore (la deviazione massima e' circa 2.54)
cutright = 3
# Logaritmo degli estremi del supporto
a_acq = log(cutleftacq)
b = log(cutright)

# Selezione i punti appartenenti al supporto, li ordino e
# applico il logaritmo per ridurre l'asimmetria della distribuzione
data = list();
for (i in 1:ntrain) {
  data[[i]] = distanze_f[[i]]
  data[[i]] = data[[i]][which(data[[i]]>cutleftacq & data[[i]]<cutright)]
}
data = lapply(data, sort)
data = lapply(data, log)

# Numero di punti da utilizzare nella stima delle ECDF
fin_num = 500
# Numero di punti in ogni lotto: numero di deviazioni totali/500
batches = floor(as.numeric(lapply(data, length))/fin_num)
# Operazione di mediazione: a ogni lotto di punti si assegna il suo
# valore medio
for (i in 1:ntrain)
  data[[i]] = mean_batch(data[[i]], batches[i])

# Stima delle funzioni di ripartizione empiriche
ecdf = lapply(data, ecdf)
# Punti del supporto su cui calcolare le PDF
t = seq(a_acq, b, length=302)
t = t[-c(1, 302)]
### Stima delle PDF
epdf.sm1=lapply(ecdf, fd.01, t, a_acq, b)
epdf.sm1a=simplify2array(epdf.sm1)
```

La SFPCA delle PDF stimate viene calcolata sfruttando la sua relazione con la FPCA applicata a trasformazioni clr di funzioni di densità (Hron et al., 2016). La funzione ausiliaria `clr()` calcola la trasformazione clr (4.1) di una densità. Essa richiede come *input* i) una PDF stimata attraverso la funzione `fd.01()`, ii) un vettore di punti in cui calcolare la trasformazione e iii) il *range* del supporto della PDF e restituisce come *output* un vettore contenente la trasformazione clr della PDF, valutata nei punti dati in *input*.

Le componenti principali si calcolano a partire dalla decomposizione spettrale della matrice di covarianza delle PDF, così trasformate, attraverso la funzione `eigen()` della libreria `base`. Infine, si ottengono i punteggi lungo le componenti principali e le statistiche T^2 e Q con i relativi limiti di controllo, come indicato nel Paragrafo 3.2.

```
# Lunghezza del supporto
t_step = t[2]-t[1]
# Numero di punti del supporto
n=length(t)
### Calcolo della trasformata clr nei punti del supporto
lepdf1 = lapply(epdf.sm1, clr, t, t_step)
lepdf1a=simplify2array(lepdf1)

# Calcolo della matrice di covarianza delle PDF trasformate tramite clr
lmod=lepdf1a
lmod=lmod*sqrt(t_step)
lmod[1,]=lmod[1,]/sqrt(2)
lmod[n,]=lmod[n,]/sqrt(2)
S=(ntrain-1)/ntrain*cov(t(lmod))

### Calcolo della PCA
# Decomposizione in autovettori e autovalori
pc=eigen(S)
pc$vec[1,]=pc$vec[1,]*sqrt(2)
pc$vec[n,]=pc$vec[n,]*sqrt(2)
pc$vec=pc$vec/sqrt(t_step)
K=Nmax.harm=10 # numero massimo di componenti principali
```

```
# Centro le osservazioni
m1=apply(lepdf1a, 1, mean)
lepdf1a.c=NULL
for(i in 1:ntrain)
  lepdf1a.c=cbind(lepdf1a.c, lepdf1a[,i]-m1)

# Calcolo i punteggi
sc.a=matrix(NA, ncol=Nmax.harm, nrow=ntrain)
for(i in 1:ntrain) {
  for(j in 1:Nmax.harm)
    sc.a[i,j]=trapzc(t_step, lepdf1a.c[,i]*pc$vec[,j])
}

# Selezione il numero di componenti principali sulla base della
# soglia di varianza spiegata
thr = 0.98;
K=min(which((cumsum(pc$val)[1:Nmax.harm]/sum(pc$val))>thr)))

# Salvo i risultati della PCA
pca=list()
pca$values=pc$values
pca$harmonics=pc$vec
pca$scores=sc.a
T2 = Q = matrix(0, 1, ntrain)
SC0 = rbind(sc.a)

### Calcolo delle statistiche di controllo T2 e Q
for (i in 1:ntrain) {
  for (k in 1:K)
    T2[i] = T2[i]+(SC0[i,k]^2)/pc$values[k]
  for (k in (K+1):Nmax.harm)
    Q[i] = Q[i] + (SC0[i,k]^2)
}
T2 = as.vector(T2)
Q = as.vector(Q)
```

```
### Calcolo dei limiti di controllo
# Probabilita' di falso allarme complessiva
alfatot = 0.05
# alfa per le coppie di carte
alfaside = alfatot/2
# alfa per una singola carta
alfasingle = 1-sqrt(1-alfaside)

# Limite per la carta T2
T2lim = (ntrain-1)^2/ntrain*qbeta(1-alfasingle,K/2,(ntrain-K-1)/2)

# Limite per la carta Q
th1 = sum(pca$values[(K+1):Nmax.harm]);
th2 = sum(pca$values[(K+1):Nmax.harm]^2)
th3 = sum(pca$values[(K+1):Nmax.harm]^3)
h0 = 1-2*th1*th3/(3*th2^2);
Qlim = th1*(1+qnorm(1-alfasingle)*sqrt(2*th2*h0^2)/th1+th2*h0*
(h0-1)/th1^2)^(1/h0);

### Creo un unico oggetto per salvare le quantita' calcolate
forward = list(data, epdf.sm1a, lepdf1a, m1, SCO, T2, Q, K,
  pca$harmonics, pca$values, T2lim, Qlim)
names(forward) = c("data", "smooth_dens", "ldens", "lmean", "scores",
  "T2", "Q", "k", "eigenf", "eigenv", "T2lim", "Qlim")
```

Si ripetono le stesse operazioni sulle distanze d_p^j dal j -esimo ovetto al modello nominale, $j = 1, \dots, 13$.

```
data = list();
for (i in 1:ntrain) {
  data[[i]] = distanze_b[check][[i]]
  data[[i]] = data[[i]][which(data[[i]]>cutleftacq & data[[i]]<cutright)]
}
data = lapply(data, sort)
data = lapply(data, log)
batches = floor(as.numeric(lapply(data, length))/fin_num)
for (i in 1:ntrain)
```

```

    data[[i]] = mean_batch(data[[i]], batches[i])
t = seq(a_acq, b,length=302)
t = t[-c(1, 302)]
ecdf = lapply(data, ecdf)
epdf.sm1=lapply(ecdf, fd.01, t, a_acq, b)
epdf.sm1a=simplify2array(epdf.sm1)
t_step = t[2]-t[1]
n=length(t)
lepdf1 = lapply(epdf.sm1, clr, t, t_step)
lepdf1a=simplify2array(lepdf1)
lmod=lepdf1a
lmod=lmod*sqrt(t_step)
lmod[1,]=lmod[1,]/sqrt(2)
lmod[n,]=lmod[n,]/sqrt(2)
S=(ntrain-1)/ntrain*cov(t(lmod))
pc=eigen(S)
pc$vec[1,]=pc$vec[1,]*sqrt(2)
pc$vec[n,]=pc$vec[n,]*sqrt(2)
pc$vec=pc$vec/sqrt(t_step)
K=Nmax.harm=10
m1=apply(lepdf1a, 1, mean)
lepdf1a.c=NULL
for(i in 1:ntrain)
  lepdf1a.c=cbind(lepdf1a.c, lepdf1a[,i]-m1)
sc.a=matrix(NA, ncol=Nmax.harm, nrow=ntrain)
for(i in 1:ntrain) {
  for(j in 1:Nmax.harm)
    sc.a[i,j]=trapzc(t_step, lepdf1a.c[,i]*pc$vec[,j])
}
thr = 0.98;
K=min(which((cumsum(pc$val)[1:Nmax.harm]/sum(pc$val)>thr)))
pca=list()
pca$values=pc$values
pca$harmonics=pc$vec
pca$scores=sc.a
T2 = Q = matrix(0, 1, ntrain)

```

```
SC0 = rbind(sc.a)
for (i in 1:ntrain) {
  for (k in 1:K)
    T2[i] = T2[i]+(SC0[i,k]^2)/pc$values[k]
  for (k in (K+1):Nmax.harm)
    Q[i] = Q[i] + (SC0[i,k]^2)
}
T2 = as.vector(T2)
Q = as.vector(Q)
alfatot = 0.05
alfaside = alfatot/2
alfasingle = 1-sqrt(1-alfaside)
T2lim = (ntrain-1)^2/ntrain*qbeta(1-alfasingle,K/2,(ntrain-K-1)/2)
th1 = sum(pca$values[(K+1):Nmax.harm]);
th2 = sum(pca$values[(K+1):Nmax.harm]^2)
th3 = sum(pca$values[(K+1):Nmax.harm]^3)
h0 = 1-2*th1*th3/(3*th2^2);
Qlim = th1*(1+qnorm(1-alfasingle)*sqrt(2*th2*h0^2)/th1+th2*h0*
(h0-1)/th1^2)^(1/h0);
backward = list(data, epdf.smla, lepdf1a, m1, SC0, T2, Q, K,
  pca$harmonics, pca$values, T2lim, Qlim)
names(backward) = c("data", "smooth_dens", "ldens", "lmean", "scores",
  "T2", "Q", "k", "eigenf", "eigenv", "T2lim", "Qlim")

### Salvo i risultati e le varie quantita' in un oggetto utile per
# fare i grafici
PcaData = list(a_acq,b,thr,t,n,forward,backward)
names(PcaData) = c("a_acq","b","thr","t","nt","forward","backward")
save(PcaData,file = "PcaResults.rdata")
```

A.1.3 Grafici

```
load("PcaResults.rdata")

### parametri grafici per la legenda
ncontrol = 8
```

```
ndefecttype = 2
defectnumerosity = c(1, 4)
ntrain = ncontrol + sum(defectnumerosity)
ncolors = 1 + ndefecttype
legend = c("in-controllo", "Eccesso di materiale", "Mancanza di materiale")
colors = c("gray", "red", "blue")
colvec = rep(colors[1], ncontrol)
lwds = rep(1, ncontrol)
lty = rep("solid", ncontrol)
for ( j in 1:ndefecttype) {
  colvec = c(colvec, rep(colors[j+1], defectnumerosity[j]))
  lwds = c(lwds, rep(1, defectnumerosity[j]))
}
lwds[9:13] = 2
lty = c(lty, "longdash", rep("dashed", 4))
pch = c(rep(16, ncontrol), 17, rep(15, 4))

# dati relativi alle d_Sj su cui fare i grafici
attach(PcaData$forward)

### carta T2
pdf(file = "T2_forward.pdf")
par(cex = 1.5)
plot(1:ntrain, T2, cex.lab = 1, cex.axis = 1, type='b', col = "black",
      xlim = c(1, ntrain), xlab = "Ovetto", main = "Statistica T2",
      ylim = c(0, 20), lwd = 2)
points(1:ntrain, T2, pch=pch, col = colvec, cex = 1.5)
abline(h = T2lim, col = "black")
legend("topleft", legend = legend, col = colors, cex=0.7, pch =
      c(19, 17, 15))
dev.off()

### carta Q
pdf(file = "Q_forward.pdf")
par(cex = 1.5)
plot(1:ntrain, Q, cex.lab = 1, cex.axis = 1, type='b', col = "black",
```

```
        xlim = c(1,ntrain), xlab = "Ovetto", main = "Statistica Q",
        ylim = c(0,100),lwd = 2)
points(1:ntrain, Q, col = colvec, cex = 1.5,pch = pch)
abline(h = Qlim, col = "black")
legend("topleft", legend = legend, col = colors, cex=0.7, pch =
      c(19,17,15))
dev.off()
detach(PcaData$forward)

# dati relativi alle d^j_P su cui fare i grafici
attach(PcaData$backward)

### carta T2
pdf(file = "T2_backward.pdf")
par(cex = 1.5)
plot(1:ntrain,T2,cex.lab = 1, cex.axis = 1, type='b',col = "black", xlim
     = c(1,ntrain), xlab = "Ovetto", main = "Statistic T2", ylim =
     c(0,20),lwd = 2)
points(1:ntrain,T2,col = colvec, cex = 1.5,pch = pch)
abline(h = T2lim, col = "black")
legend("topleft",legend = legend,col = colors, cex=0.7, pch = c(19,17,15))
dev.off()

### carta Q
pdf(file = "Q_backward.pdf")
par(cex = 1.5)
plot(1:ntrain,Q,cex.lab = 1, cex.axis = 1, type='b',col = "black", xlim =
     c(1,ntrain), xlab = "Ovetto", main = "Statistica Q", ylim =
     c(0,100),lwd = 2)
points(1:ntrain,Q,col = colvec, cex = 1.5,pch = pch)
abline(h = Qlim, col = "black")
legend("topleft",legend = legend,col = colors, cex=0.7, pch = c(19,17,15))
dev.off()
detach(PcaData$backward)
```


A.2 Sorveglianza non parametrica dello spettro dell'operatore LB discreto

In Appendice A.2.1 è riportato il codice Matlab per il calcolo dell'operatore discreto di Laplace-Beltrami, in A.2.2 è implementata in R la procedura di sorveglianza di Capizzi e Masarotto (2017) e in A.2.3 è effettuata in Matlab l'analisi post-diagnostica tramite l'algoritmo ICP.

A.2.1 Calcolo dell'operatore LB discreto

I dati relativi agli oggetti sono stati convertiti tramite il programma MeshLab da file `.stl` a file oggetto (`.obj`). Si tratta di un formato semplice e molto utilizzato per l'interscambio di dati, che rappresenta solamente la geometria tridimensionale dell'oggetto, ossia la posizione di ogni vertice e le facce che compongono il modello.

In questo modo gli oggetti possono essere letti e caricati in Matlab attraverso la funzione `readMesh()` del pacchetto `matlabmesh`. Ogni oggetto è rappresentato in forma di *mesh* ed è descritto da i) una matrice $m \times 3$ relativa ai vertici, contenente le coordinate tridimensionali dei vertici, ii) una matrice $N_{tri} \times 3$ relativa alle N_{tri} facce triangolari, in cui ogni riga contiene i vertici che definiscono una faccia, e iii) una matrice $m \times m$ sparsa relativa agli archi, in cui l'elemento $m_{ij} = 1$ se i vertici p_i e p_j sono collegati da un arco, $m_{ij} = 0$ altrimenti.

Successivamente, la funzione `LocalizedLaplacian()` calcola, per ogni oggetto, l'operatore discreto di Laplace-Beltrami proposto da Li et al. (2015) e definito nella (3.17). Essa richiede come *input* una *mesh*, definita come sopra, e un parametro `eta`, che indica il livello di approssimazione desiderato (di *default* è pari a 0.001). La funzione calcola le quantità definite nella (4.2). In particolare, la distanza geodesica tra le coppie di punti è ottenuta tramite la funzione `perform_fast_marching()` del pacchetto `fast marching toolbox`. La funzione `LocalizedLaplacian()` restituisce come *output* la matrice `L_t` di dimensione $m \times m$, contenente l'operatore LB discreto valutato negli m vertici della *mesh*, e un vettore `b` di dimensione $m \times 1$, contenente gli elementi sulla diagonale della matrice B , definita nella (4.3).

Infine, si ricava la matrice `L_symm`, come indicato nella (4.4), e si calcola la parte inferiore del suo spettro tramite la funzione `eigs()`. Gli autovalori di ciascun ovetto vengono salvati in un file `.csv`, in cui le righe rappresentano i 13 ovetto e le colonne i primi 30 autovalori ottenuti.

```
% Numero di autovalori da calcolare
numEigs = 30;
% Numero di ovetto
numParts = 13;
Data = zeros(numParts,numEigs);
eigsopts = struct('disp',0);
% Calcolo lo spettro dell'operatore discreto LB per ogni ovetto
for i=1:numParts
    % carico un ovetto
    mesh = readMesh(['uovo', num2str(i), '.obj'], 'B');
    % calcolo l'operatore LB discreto
    [L_t, b] = LocalizedLaplacian(mesh);
    % calcolo gli autovalori
    L_symm = sqrt(b).*L_t./sqrt(b');
    Data(i,:) = real(eigs(L_symm, numEigs, -1e-8, eigsopts));
    % li salvo in un file .csv
    dlmwrite('spettro_uova.csv',Data(i,:), '-append')
end

% Definisco la funzione per calcolare l'operatore LB discreto
function [L_t, b] = LocalizedLaplacian(mesh,eta)

% Matrice m*3 contenente le coordinate dei vertici della mesh
v = mesh.v;
% Triangoli della mesh
f = mesh.f;
% Matrice m*m di 0 e 1 contenente gli archi della mesh
e = mesh.e;
```

```
% Numero di vertici
Nver = size(v,1);
% Numero di triangoli
Ntri = size(f,1);

% Calcolo la lunghezza media degli archi per determinare il
% valore del parametro t della formula (3.17) del presente
% elaborato
sumEdge = 0; count=0;
for i=1:(Nver-1)
    for j=(i+1):Nver
        if e(i,j)~=0
            sumEdge = sumEdge+norm(v(i,:)-v(j,:));
            count=count+1;
        end
    end
end
t = sumEdge/count;

% Calcolo le aree di tutti i triangoli della mesh
Atri = zeros(Ntri,1);
for i=1:Ntri
    % indici dei vertici dell'i-esimo triangolo
    ind = f(i,:);
    % vertici dell'i-esimo triangolo, ogni riga e' un vertice
    verts = v(ind,:);
    % funzione ausiliaria che calcola l'area
    Atri(i) = TriArea(verts(:,1), verts(:,2), verts(:,3));
end

% Valore di default di eta (minore e' eta, migliore
% e' l'approssimazione)
if ~ exist('eta', 'var')
    eta = 1.0e-2;
```

```
end

% Calcolo il raggio r utilizzando la formula (22) dell'articolo
% di Li et al. (2015)
AreaS = sum(Atri);
xlinf = max(abs(v));
T = 2*AreaS*norm(xlinf);
r = sqrt(-4*t*log(4*pi*t^2*eta/T));
option.constraint_map = r*ones(Nver,1);

% Calcolo le quantita' della formula (4.2) del presente
% elaborato, necessarie per ottenere l'operatore LB discreto
Aver = zeros(1,Nver);
Emat = sparse(Nver,Nver);
parfor i = 1:Nver
    % Distanze geodesiche tra il punto e agli altri punti
    [D,~,~] = perform_fast_marching_mesh(v', f', i, option);

    % Esponenziale delle distanze
    Emat(i,:) = exp(-D.^2/4/t)';

    % Area dell'intorno del punto i
    triId = [];
    for j=1:3
        triId = [triId; find(f(:,j)==i)];
    end
    if length(triId)~=length(unique(triId))
        fprintf('one point listed in a triangle twice!');
    end
    triId = unique(triId);
    Aver(i)=sum(Atri(triId));
end

% Calcolo le matrici W e D tali che  $L_t = D-W$  (formula (4.2))
```

```

% del presente elaborato)
W = Emat.*Aver/(12*pi*t^2);
D = diag(sum(W,2));
L_t = D-W;
b = Aver';

end

% Funzione ausiliaria che calcola l'area di un triangolo definito
% dai punti x1, x2, x3
function A = TriArea(x1, x2, x3)
l1 = x2-x1;
l2 = x3-x1;
A = sqrt((l1'*l1)*(l2'*l2)-(l1'*l2)^2)/2;
end

```

A.2.2 Procedura di sorveglianza non parametrica

Il file `.csv` viene caricato in R. La procedura di sorveglianza proposta da Capizzi e Masarotto (2017) è implementata nella funzione `mphase1()` della libreria `dfphase1`. La *performance* della carta viene valutata al variare del numero di autovalori considerati e dei valori dei parametri `lmin`, indica la lunghezza minima di un cambiamento di tipo persistente (se presente), e `gamma` della funzione `mphase1()`.

Il parametro di penalizzazione `gamma` è legato al criterio EBIC (*Extended Bayesian Information Criteria*) (Chen e Chen, 2008), che viene impiegato nell'algoritmo LASSO adattivo, ed è utilizzato per bilanciare la probabilità di individuare un vero cambiamento e la probabilità di segnalare un falso allarme. I valori presi in esame sono $\gamma = 0$, che corrisponde al criterio BIC standard, e $\gamma = 0.5$ o $\gamma = 1$ che offrono una migliore protezione rispetto ai falsi allarmi. I risultati ottenuti sono i medesimi al variare di `gamma`.

```

library(dfphase1)
dati=read.csv("spettro_uova.csv", header = F)

```

```
# Selezione gli autovalori di interesse
# Sono stati considerati gli autovalori:
# 16:17, 16:18, 16:19, 16:20
dati=dati[,16:17]
# Riorganizzo i dati in formato di array
dati2 <- array(t(dati), dim = c(ncol(dati), 1, nrow(dati)))
mphase1(dati2, lmin = 1, gamma=0.5)
# Sono stati considerati:
# lmin = 1, 2, 3
# gamma = 0, 0.5, 1
```

A.2.3 Analisi post-diagnostica

L'analisi post-diagnostica è effettuata tramite la funzione `icp()` del pacchetto `ICP`. Poiché gli ovetti sono già registrati, alle nuvole di punti viene aggiunto del rumore e vengono applicate trasformazioni rigide e permutazioni, in modo da simulare degli ovetti non allineati.

```
% Modello nominale
mesh1 = readMesh('uovo14.obj');
model = mesh1.v;
% Ovetto fuori controllo
mesh2 = readMesh('uovo10.obj');
part = mesh2.v;

% Rumore
part = part + 0.05*randn(size(part));
% Permutazione
per = randperm(size(part,1));
part = part(per,:);
% Rotazione
v1 = 0.6*(2*rand-1); v2 = 0.6*(2*rand-1); v3 = 0.6*(2*rand-1);
R1 = [1 0 0;0 cos(v1) -sin(v1);0 sin(v1) cos(v1)];
R2 = [cos(v2) 0 sin(v2);0 1 0;-sin(v2) 0 cos(v2)];
R3 = [cos(v3) -sin(v3) 0;sin(v3) cos(v3) 0;0 0 1];
```

```
part = part*R3'*R2'*R1';
% Traslazione
part = part + randn(1,3);

% Applico l'algoritmo ICP
% data: ovetto registrato
% vi(k): indice del punto del modello piu' vicino al punto k
% dell'ovetto (vi e' un vettore)
[~,~,data,~,vi]=icp(model, part);

% Calcolo delle deviazioni dal modello nominale
deviation=sqrt(sum((data'-model(vi,:)).^2,2));

% Grafico delle deviazioni
 [~,perInv] = sort(per);
plotMesh(mesh2,'fl',deviation(perInv));
```


Bibliografia

- Arnoldi, W. E. (1951). «The principle of minimized iterations in the solution of the matrix eigenvalue problem». In: *Quarterly of Applied Mathematics* 9.1, pp. 17–29.
- Besl, P. J. e McKay, N. D. (1992). «Method for registration of 3-D shapes». In: *Sensor fusion IV: control paradigms and data structures*. Vol. 1611. Spie, pp. 586–606.
- Box, G. E. et al. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Bui, A. T. e Apley, D. W. (2022). «Analyzing nonparametric part-to-part variation in surface point cloud data». In: *Technometrics* 64.4, pp. 457–474.
- Capizzi, G. e Masarotto, G. (2017). «Phase I distribution-free analysis of multivariate data». In: *Technometrics* 59.4, pp. 484–495.
- Chavel, I. (1984). *Eigenvalues in Riemannian geometry*. Academic press.
- Chen, J. e Chen, Z. (2008). «Extended Bayesian information criteria for model selection with large model spaces». In: *Biometrika* 95.3, pp. 759–771.
- Chen, N., Zi, X. e Zou, C. (2016). «A distribution-free multivariate control chart». In: *Technometrics* 58.4, pp. 448–459.
- Colosimo, B. M. e Pacella, M. (2007). «On the use of principal component analysis to identify systematic patterns in roundness profiles». In: *Quality and Reliability Engineering International* 23.6, pp. 707–725.
- Colosimo, B. M., Semeraro, Q. e Pacella, M. (2008). «Statistical process control for geometric specifications: on the monitoring of roundness profiles». In: *Journal of Quality Technology* 40.1, pp. 1–18.

- Colosimo, B. M. et al. (2014). «From profile to surface monitoring: SPC for cylindrical surfaces via Gaussian processes». In: *Journal of Quality Technology* 46.2, pp. 95–113.
- Colosimo, B. M. et al. (2018a). «Editors' note for special issue: "Quality engineering in advanced manufacturing"». In: *Journal of Quality Technology* 50.3, pp. 231–232.
- Colosimo, B. M. et al. (2018b). «Opportunities and challenges of quality engineering for additive manufacturing». In: *Journal of Quality Technology* 50.3, pp. 233–252.
- Del Castillo, E., Colosimo, B. M. e Tajbakhsh, S. D. (2015). «Geodesic Gaussian processes for the parametric reconstruction of a free-form surface». In: *Technometrics* 57.1, pp. 87–99.
- Hausdorff, F. (1914). *Grundzüge der Mengenlehre*, Leipzig.
- He, K. et al. (2017). «Enhancing the monitoring of 3D scanned manufactured parts through projections and spatiotemporal control charts». In: *Journal of Intelligent Manufacturing* 28, pp. 899–911.
- Hron, K. et al. (2016). «Simplicial principal component analysis for density functions in Bayes spaces». In: *Computational Statistics & Data Analysis* 94, pp. 330–350.
- Kalender, W. A. (2011). *Computed tomography: fundamentals, system technology, image quality, applications*. John Wiley & Sons.
- Kellner, T. (2017). *An Epiphany Of Disruption: GE Additive Chief Explains How 3D Printing Will Upend Manufacturing*. URL: <https://www.ge.com/news/taxonomy/term/2774?page=3>.
- Kreyszig, E. (2013). *Differential geometry*. Courier Corporation.
- Kullback, S. e Leibler, R. A. (1951). «On information and sufficiency». In: *The Annals of Mathematical Statistics* 22.1, pp. 79–86.
- Leblanc, A. (2010). «A bias-reduced approach to density estimation using Bernstein polynomials». In: *Journal of Nonparametric Statistics* 22.4, pp. 459–475.
- Li, X., Xu, G. e Zhang, Y. J. (2015). «Localized discrete Laplace–Beltrami operator over triangular mesh». In: *Computer Aided Geometric Design* 39, pp. 67–82.

-
- Liu, L. et al. (2008). «A local/global approach to mesh parameterization». In: *Computer graphics forum*. Vol. 27. 5. Wiley Online Library, pp. 1495–1504.
- Maaten, L. van der e Hinton, G. (2008). «Visualizing high-dimensional data using t-SNE». In: *Journal of Machine Learning Research* 9, pp. 2579–2605.
- Maleki, M. R., Amiri, A. e Castagliola, P. (2018). «An overview on recent profile monitoring papers (2008–2018) based on conceptual classification scheme». In: *Computers & Industrial Engineering* 126, pp. 705–728.
- Matérn, B. (2013). *Spatial variation*. Vol. 36. Springer Science & Business Media.
- Menafoglio, A. et al. (2018). «Profile monitoring of probability density functions via simplicial functional PCA with application to image data». In: *Technometrics* 60.4, pp. 497–510.
- Moller, J. (2012). *Lectures on random Voronoi tessellations*. Vol. 87. Springer Science & Business Media.
- Montgomery, D. C. (2019). *Introduction to statistical quality control*. John Wiley & Sons.
- Oja, H. (2010). *Multivariate nonparametric methods with R: an approach based on spatial signs and ranks*. Springer Science & Business Media.
- Patanè, G. (2017). «An introduction to Laplacian spectral distances and kernels: Theory, computation, and applications». In: *ACM SIGGRAPH 2017 Courses*, pp. 1–54.
- Pesarin, F. (2001). *Multivariate permutation tests: with applications in biostatistics*. New York: Wiley.
- Pratola, M. T. (2022). «Editorial: special issue on Industry 4.0». In: *Technometrics* 64.4, pp. 435–436.
- Qiu, P. (2014). *Introduction to statistical process control*. CRC press.
- Reuter, M. et al. (2009). «Laplace–Beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis». In: *Computer-Aided Design* 41.10, pp. 739–755.
- Saghaei, A., Noorossana, R. e Amiri, A. (2013). *Statistical analysis of profile monitoring*. Wiley Online Library.

- Scimone, R. et al. (2022). «Statistical modeling and monitoring of geometrical deviations in complex shapes with application to additive manufacturing». In: *Technometrics* 64.4, pp. 437–456.
- Śladek, J. A. (2016). «Coordinate metrology». In: *Springer Tracts in Mechanical Engineering*, pp. 1–13.
- Stankus, S. E. e Castillo-Villar, K. K. (2019). «An improved multivariate generalised likelihood ratio control chart for the monitoring of point clouds from 3D laser scanners». In: *International Journal of Production Research* 57.8, pp. 2344–2355.
- Tenenbaum, J. B., Silva, V. d. e Langford, J. C. (2000). «A global geometric framework for nonlinear dimensionality reduction». In: *Science* 290.5500, pp. 2319–2323.
- Vitale, R. A. (1975). «A Bernstein polynomial approach to density function estimation». In: *Statistical inference and related topics*. Elsevier, pp. 87–99.
- Wang, A., Wang, K. e Tsung, F. (2014). «Statistical surface monitoring by spatial-structure modeling». In: *Journal of Quality Technology* 46.4, pp. 359–376.
- Wells, L. J. et al. (2013). «Statistical process monitoring approach for high-density point clouds». In: *Journal of Intelligent Manufacturing* 24, pp. 1267–1279.
- Woodall, W. H. (2007). «Current research on profile monitoring». In: *Production* 17, pp. 420–425.
- Zang, Y. e Qiu, P. (2018a). «Phase I monitoring of spatial surface data from 3D printing». In: *Technometrics* 60.2, pp. 169–180.
- Zang, Y. e Qiu, P. (2018b). «Phase II monitoring of free-form surfaces: An application to 3D printing». In: *Journal of Quality Technology* 50.4, pp. 379–390.
- Zhao, X. (2022). «Intrinsic geometrical methods for statistical process control of complex data objects». Tesi di dottorato. Pennsylvania State University.
- Zhao, X. e Del Castillo, E. (2021). «An intrinsic geometrical approach for statistical process control of surface and manifold data». In: *Technometrics* 63.3, pp. 295–312.

- Zhu, L. et al. (2007). «On-line HPLC combined with multivariate statistical process control for the monitoring of reactions». In: *Analytica Chimica Acta* 584.2, pp. 370–378.
- Zou, H. (2006). «The adaptive lasso and its oracle properties». In: *Journal of the American Statistical Association* 101.476, pp. 1418–1429.