

UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA DELL'INFORMAZIONE

Caratterizzazione del contenuto di point cloud basata su feature

RELATORE:

PROF.SSA FEDERICA BATTISTI

CORRELATORE:

GIUSEPPE VALENZISE

Université Paris-Saclay

CANDIDATO:

ANNALISA

GALLINA

2000271

Abstract

Il recente sviluppo di nuove tecnologie multimediali ha portato all'affermazione della point cloud come modello di rappresentazione di oggetti tridimensionali in vari settori e, conseguentemente, è diventato essenziale perfezionare tecniche di compressione che permettano di trasmettere questa notevole mole di dati in maniera efficiente e compatta. Parimenti, è utile conoscere la complessità di una point cloud per determinare il miglior compromesso tra le risorse disponibili e la qualità finale raggiunta a seguito della compressione. A tale scopo, questo studio si propone di definire metriche oggettive generali e computazionalmente semplici da calcolare che permettano di valutarla. In particolare, le misure proposte e analizzate in questo lavoro di tesi considerano prevalentemente la densità del contenuto e la rugosità della superficie con l'obiettivo di definire la complessità geometrica di una point cloud. I risultati ottenuti dimostrano che tali fattori influenzano in maniera significativa la facilità di compressione di ciascun oggetto, pur rimanendo strettamente legati al metodo di codifica utilizzato.

Questa tesi risulta essere il punto di partenza per un'analisi più approfondita da effettuare al fine di generalizzare i risultati ottenuti.

*Alla mia famiglia per avermi supportato e sopportato,
ai nonni Anna e Pietro per aver sempre creduto in me,
alla mia relatrice e a tutti gli insegnanti che mi hanno aiutato a capire cosa voglio fare da
grande*

Indice

Abstract	III
1 Introduzione	1
1.1 Motivazione	1
1.2 Point Cloud	2
1.3 Stato dell'arte	3
1.4 Scopo della tesi	4
2 Dataset e procedure applicate	7
2.1 Dataset	7
2.2 MPEG G-PCC	7
2.2.1 Voxelizzazione	9
2.2.2 Codifica Octree	10
2.2.3 Codifica Trisoup	12
2.3 Metriche di distorsione	13
2.3.1 Distanza Point to Point	15
2.3.2 Distanza Point to Plane	15
3 Metodo Proposto	19
3.1 Curve di distorsione	19
3.1.1 Tasso di compressione	20
3.1.2 Livello di distorsione	20
3.1.3 Curve ottenute	21
3.2 Misura della qualità di compressione	23
3.2.1 Bitrate a parità di livello di qualità	25
3.2.2 Bitrate Medio	25
3.2.3 Integrale della curva di distorsione interpolata	26

3.3	Misura della complessità	33
3.3.1	Distanza dai punti più vicini	33
3.3.2	Curvatura	37
3.3.3	Densità	42
4	Risultati sperimentali	45
4.1	Distanza dai punti più vicini	45
4.2	Curvatura	48
4.3	Densità	50
5	Conclusioni	55
5.1	Sviluppi futuri	56
	Acronimi	57
	Bibliografia	57

Elenco delle figure

1.1	Esempio di estrazione dell'informazione spaziale	3
2.1	8i Voxelized Surface Light Field (8iVSLF) Dataset	8
2.2	Esempio di una point cloud voxelizzata	9
2.3	Schema di codifica Octree	12
2.4	Approssimazione planare della superficie di una point cloud	13
2.5	Distanza dal punto più vicino	16
2.6	Proiezione dell'errore nella direzione normale alla superficie	17
3.1	Curve di distorsione del dataset 8iVSLF	22
3.2	Curve di distorsione interpolate: in alto i grafici interpolati relativi alle curve octree, mentre in basso i medesimi risultati relativi a trisoup	28
3.3	Interpolazione della curva di distorsione al di fuori dell'intervallo di definizione	29
3.4	Ricerca dei punti più vicini	36
3.5	Curvature principali di una superficie	39
4.1	Scattering dei punti relativi alla media (a sinistra) e alla deviazione standard (a destra) della distanza dal punto più vicino per codifica octree (sopra) e trisoup (sotto)	47
4.2	Scattering dei punti relativi alla deviazione standard della distanza media dai 50 punti più vicini per codifica octree (destra) e trisoup (sinistra)	48
4.3	Scattering dei punti relativi alla rugosità per codifica octree (sinistra) e trisoup (destra)	50
4.4	Scattering dei punti relativi alla media (a sinistra) e alla deviazione standard (a destra) della curvatura per codifica octree (sopra) e trisoup (sotto)	51
4.5	Scattering dei punti relativi alla densità globale dei contenuti: (1) considera la densità calcolata su tutti i livelli, (2) valuta unicamente i primi 5 livelli e (3) è relativa ai primi 3 livelli	52

4.6 Scattering dei punti relativi alla densità locale dei contenuti: (1) considera la densità calcolata su tutti i livelli, (2) valuta unicamente gli ultimi 4 livelli e (3) è relativa agli ultimi 2 livelli 52

Elenco delle tabelle

2.1	Parametro di quantizzazione octree	10
3.1	Octree	22
3.2	Trisoup	22
3.3	Correlazione delle misure di complessità per octree: (a) integrale centrato sulla media, (b) integrale sul supporto comune, (c) integrale sul supporto singolo, (d) bitrate medio	31
3.4	Correlazione delle misure di complessità per trisoup: (a) integrale centrato sulla media, (b) integrale sul supporto comune, (c) integrale sul supporto singolo, (d) bitrate medio	32
3.5	Percentuale di punti validi per il calcolo della curvatura con raggio 2.225 . . .	41
4.1	Correlazione tra le misure di complessità legate alla distanza tra i punti e la misura di difficoltà di compressione	45
4.2	Correlazione tra la distanza dal punto più vicino e la misura di difficoltà di compressione, esclusa thaidancer	47
4.3	Correlazione tra le misure di complessità legate alla rugosità del contenuto e la misura di difficoltà di compressione	48
4.4	Correlazione della curvatura calcolata con raggio 4.5, escluso thaidancer . . .	49
4.5	Correlazione tra le misure di complessità legate alla densità del contenuto e la misura di difficoltà di compressione	51
4.6	Correlazione tra le misure di complessità legate alla probabilità di occupazione fissata la dimensione della cella e la misura di difficoltà di compressione	53

Capitolo 1

Introduzione

1.1 Motivazione

Nel recente passato è aumentato considerevolmente l'utilizzo delle point cloud come formato di rappresentazione per contenuti tridimensionali in numerosi ambiti: per esempio, hanno affiancato le 3D meshes nella rappresentazione dello spazio circostante in scenari riguardanti la realtà virtuale o aumentata, allo stesso tempo sono estremamente usate anche nella ricostruzione dell'ambiente per veicoli a guida autonoma.

Questa rapida crescita è determinata dalla grande facilità di rappresentazione ed elaborazione e, parallelamente, dai notevoli passi avanti fatti nell'acquisizione di tali contenuti: ciò ha portato ad avere oggetti di qualità sempre più alta, permettendo all'utente di avere un'esperienza estremamente realistica e immersiva.

Dall'altra parte, un'inevitabile conseguenza è stata quella di avere oggetti rappresentati con un numero sempre crescente di dati, arrivando ad ottenere point cloud costituite da diversi miliardi di punti. Da questo, segue la necessità di procedere su due diversi fronti: da una parte è diventato essenziale perfezionare tecniche di compressione che permettano di rappresentare la grande mole di dati in maniera efficiente e compatta, dall'altra ha assunto importanza crescente l'esigenza di definire delle metriche di valutazione della complessità di una point cloud, che permettano di stimare a priori la difficoltà della compressione e il trade-off ottimale tra lo spazio occupato e la qualità del risultato.

Questa necessità deriva sicuramente dall'urgenza di diminuire le dimensioni di tali oggetti per poterli trasmettere e processare agevolmente senza intaccare l'esperienza dell'utente, ma anche dall'impossibilità di testare di volta in volta il modello di compressione che meglio si adatta alle caratteristiche della point cloud e all'ambito di utilizzo.

Inoltre, questo tipo di metrica potrebbe permettere di fare un'analisi della complessità che si svincoli totalmente dal codificatore utilizzato e vada a definire un risultato universale: nel caso delle point cloud, questo aspetto è ancora marginale, in quanto i sistemi di codifica a disposizione sono molto limitati e funzionano in maniera alquanto diversa. Tuttavia, riportandoci al caso bidimensionale, questo aspetto assume grande importanza.

Chiaramente, la metrica ottenuta non sarà mai in grado di determinare la complessità geometrica del contenuto meglio di quanto non sia possibile fare confrontando a posteriori l'oggetto originale con il risultato della compressione, ma può comunque essere estremamente utile fornendo una stima ad un costo computazionale pressochè nullo.

1.2 Point Cloud

Una point cloud, dall'inglese "nuvola di punti", è una collezione di punti rappresentati nello spazio tridimensionale e, di conseguenza, corrisponde ad un insieme di coordinate nello spazio euclideo XYZ. Ciascun punto, oltre alla sua rappresentazione spaziale, può essere dotato di alcuni attributi che lo caratterizzano. In particolare, ad ognuno di essi può essere associata la rappresentazione RGB del relativo colore, così come il vettore normale alla superficie in corrispondenza del punto stesso.

Tali punti sono completamente indipendenti tra loro e questo significa che non sono collegati gli uni agli altri in nessun modo, non andando a definire delle superfici che approssimino l'oggetto rappresentato, a differenza di quanto succede per altri formati di rappresentazione di contenuti tridimensionali.

A seconda della vicinanza media tra i punti, possiamo distinguere due diverse categorie di point cloud: point cloud dense e point cloud sparse. Una point cloud sparsa è caratterizzata, in generale, da una distanza media tra i punti superiore, che comporta un livello di dettaglio inferiore e una minore precisione nella rappresentazione della geometria. D'altra parte, però, abbattendo la quantità di punti, aumenta notevolmente la velocità di trasmissione del segnale e questo rende le point cloud sparse particolarmente utili in alcuni ambiti di utilizzo, in cui la velocità è da privilegiare rispetto alla precisione, come per esempio in applicazioni relative a veicoli a guida autonoma.

Al contrario, le point cloud dense sono oggetti in cui la vicinanza tra i punti è molto elevata, garantendo un livello di precisione superiore nella rappresentazione. Generalmente, la generazione di questo tipo di contenuti avviene osservando l'oggetto da differenti punti di vista e combinando assieme i risultati delle diverse acquisizioni.

Una successiva distinzione può essere fatta sulla base della tempo invarianza del contenuto: in questo caso distinguiamo point cloud statiche e dinamiche. Definendo una point cloud statica, si intende che il suo contenuto è tempo invariante, ossia la posizione dei punti non si modifica nel corso del tempo. Le point cloud dinamiche, invece, hanno un contenuto che varia nel tempo e sono utilizzate per applicazioni che richiedono di poter catturare il movimento dell'oggetto o la dinamica della scena. Una point cloud dinamica è costituita da un insieme di frame, ciascuno dei quali rappresenta una singola istantanea temporale dell'oggetto e contiene le coordinate tridimensionali dei punti in quel determinato momento. Nello sviluppo di questa tesi faremo riferimento sempre a point cloud statiche, anche se l'attributo verrà omesso.

Ulteriori caratteristiche della point cloud, come la densità di punti e il passo di campionamento scelto, variano molto a seconda dell'oggetto e del metodo di acquisizione utilizzato. In generale, il passo di campionamento usato non è costante e questo comporta un campionamento non uniforme del contenuto, introducendo un ulteriore elemento di difficoltà nella descrizione e nella compressione della point cloud.

1.3 Stato dell'arte

Per quanto è alla nostra conoscenza, il problema proposto non è mai stato affrontato e risolto nella letteratura, per quanto concerne l'analisi di contenuti tridimensionali. Questo implica che non esistono delle metriche di riferimento per la valutazione della complessità geometrica di una point cloud.

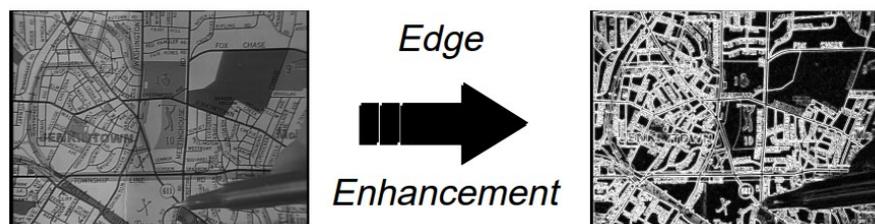


Figura 1.1: Esempio di estrazione dell'informazione spaziale

Nel caso bidimensionale, la complessità di un contenuto viene valutata attraverso la Spatial Information, che è una misura della presenza di alte frequenze nello spettro dell'immagine. Entrando nel dettaglio, secondo quanto è riportato in [6], tale misura può essere ottenuta applicando un filtro passa alto orizzontale e verticale a ciascun pixel dell'immagine, in modo da rendere più evidenti le differenze nelle due direzioni. In particolare, il filtro verticale evidenzia i bordi orizzontali mentre l'equivalente orizzontale serve per rendere più evidenti i bordi

verticali. I risultati dei due filtraggi possono, in seguito, essere composti in modo da ottenere una misura univoca per ciascun pixel. La complessità dell'intera immagine viene, poi, definita effettuando una statistica su tutti i valori ottenuti per i singoli pixel.

In particolare, lo studio citato riporta che la misura che meglio predice la complessità di un'immagine è la media dei valori precedentemente ottenuti, che ottiene dei risultati migliori rispetto alla deviazione standard e all'errore quadratico medio:

$$SI_{mean} = \sqrt{\frac{1}{P} \sum SI_r^2} \quad (1.1)$$

La validità della misura appena illustrata è stata provata in quanto risulta essere sperimentalmente quella maggiormente correlata con qualsiasi metrica di complessità di compressione scelta, che funge da banco di prova per la correttezza della stima a priori della complessità del contenuto.

Nel caso specifico, la misura di complessità utilizzata risulta essere il tasso di compressione necessario per codificare l'immagine ad un certo livello di qualità, definito come rapporto tra la taglia del file originale e quella del file compresso, fissato il parametro di quantizzazione. Tuttavia, i risultati ottenuti sono quasi del tutto equivalenti per diversi livelli di qualità valutati o per diverse definizioni di complessità di compressione.

1.4 Scopo della tesi

In analogia con quanto proposto per le immagini e vista l'assenza di un corrispettivo nel caso tridimensionale, l'obiettivo di questa tesi è quello di definire delle caratteristiche che permettano di classificare la complessità del contenuto in modo da determinare il trade-off ottimale tra il livello di compressione e la qualità raggiunta.

In particolare, l'idea è di andare a definire delle metriche di complessità ridotta e quanto più generali possibile, che permettano di classificare diverse point cloud a seconda della loro geometria e che siano quanto più possibile legate a misure di complessità della compressione, in modo da raggiungere l'obiettivo proposto.

Per farlo è doveroso procedere su due fronti: innanzitutto, è necessario andare a determinare una metrica che valuti la complessità dell'oggetto in relazione a quanto è difficile compri-merlo. Nel fare questo, l'approccio seguito è molto simile a quello utilizzato nel caso bidimensionale, andando a definire una metrica di valutazione del bitrate della point cloud codificata sulla base qualità raggiunta. Dall'altra parte l'obiettivo è caratterizzare un parametro oggettivo

vo che definisca la complessità a seconda delle caratteristiche geometriche dell'oggetto, ossia il corrispettivo della Spatial Information per quanto riguarda le immagini.

La tesi procederà come segue: dapprima viene fornita un'esposizione del metodo utilizzato, partendo dalla descrizione del dataset e dei software utilizzati per la compressione delle point cloud e per la valutazione della qualità dei risultati ottenuti. Successivamente, viene descritto il processo di costruzione delle curve di distorsione e la definizione della metrica di valutazione della complessità di compressione. A seguire, viene fatta una disamina delle misure di complessità geometrica proposte e della loro validità di predizione, sulla base della correlazione con la metrica oggettiva adottata. Infine, i risultati verranno discussi e commentati, avanzando proposte di lavori futuri per espandere ulteriormente l'analisi.

Capitolo 2

Dataset e procedure applicate

2.1 Dataset

Il dataset utilizzato per i test di valutazione della validità delle features proposte è 8i Voxelized Surface Light Field (8iVSLF) Dataset[3], che è il dataset di riferimento MPEG (Moving Picture Expert Group) per point clouds dense. Esso comprende 6 point cloud statiche voxelizzate che rappresentano delle figure umane (Figura 2.1). Con point cloud voxelizzata, si intende una point cloud raffigurata su una griglia regolare: ciascun elemento della griglia risulta pieno se al suo interno è presente almeno un punto dell'insieme iniziale e in tal caso ad esso sono associati anche degli attributi, primo fra tutti il colore.

In particolare, nel dataset considerato le point cloud sono raffigurate su uno spazio di $4096 \times 4096 \times 4096$ voxels: cinque di esse sono rappresentate sulla maggior parte dello spazio a disposizione, occupando l'intera altezza, mentre l'ultima occupa circa un ottavo dello spazio a disposizione. Questa peculiarità la rende estremamente densa rispetto alle altre, che, a prescindere da ciò, sono da considerarsi point clouds dense, con un numero di punti che si aggira, per ognuna di esse, intorno ai 3 milioni.

2.2 MPEG G-PCC

Per effettuare la compressione delle point cloud precedentemente descritte è stato utilizzato il modello di compressione standardizzato MPEG, elaborato tra il 2020 e il 2021. In particolare, gli standard proposti sono due: *Geometry based Point Cloud Compression* (GPCC), utilizzato per la compressione di point cloud statiche e *Video based Point Cloud Compression* (VPCC), pensato, invece, per le point cloud dinamiche [1].

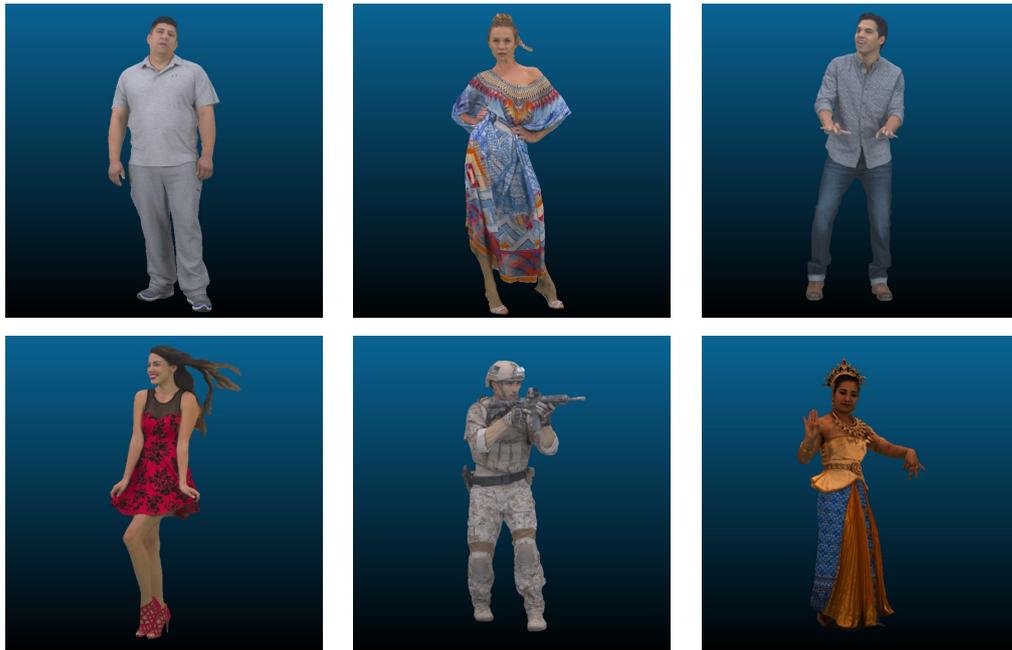


Figura 2.1: 8i Voxelized Surface Light Field (8iVSLF) Dataset

Il dataset considerato include unicamente point cloud statiche, per cui l'unico modello utilizzato e qui descritto è GPCC.

GPCC è un software per la compressione di point cloud che offre tre diversi schemi di codifica per la geometria:

- Octree: si basa su una suddivisione ricorsiva dello spazio tridimensionale in ottanti.
- Trisoup: si basa sulla codifica octree, ma migliora la qualità dell'oggetto compresso attraverso il fitting di superfici locali.
- PredGeom: codifica il punto corrente sulla base di uno schema di predizione applicato ai punti precedentemente acquisiti.

Nel caso specifico, ai fini dell'obiettivo da raggiungere, le point cloud considerate sono state compresse utilizzando unicamente i primi due modelli, che sono descritti più approfonditamente nel seguito.

Per quanto riguarda la codifica degli attributi, invece, le modalità di compressione consentite sono due:

- Predictive/Lifting Transform: suddivide i punti in livelli di dettaglio e utilizza i punti vicini del *Level of Detail* (LoD) precedente per predire gli attributi del punto corrente.
- Raht: segue la suddivisione octree, ma percorre l'albero ottenuto con approccio bottom-up, partendo dai singoli voxels, fino ad arrivare alla radice. La codifica degli attributi di ciascun nodo si basa su una predizione fatta sugli attributi dei nodi figli già visitati.

La trattazione dei metodi di codifica degli attributi non verrà affrontata con ulteriore dettaglio, in quanto l'analisi svolta successivamente si concentra esclusivamente sulla caratterizzazione della geometria delle point cloud.

2.2.1 Voxelizzazione

Indipendentemente dal modello di codifica della geometria utilizzato, la point cloud è sottoposta ad una fase di preprocessing, definita Voxelizzazione.

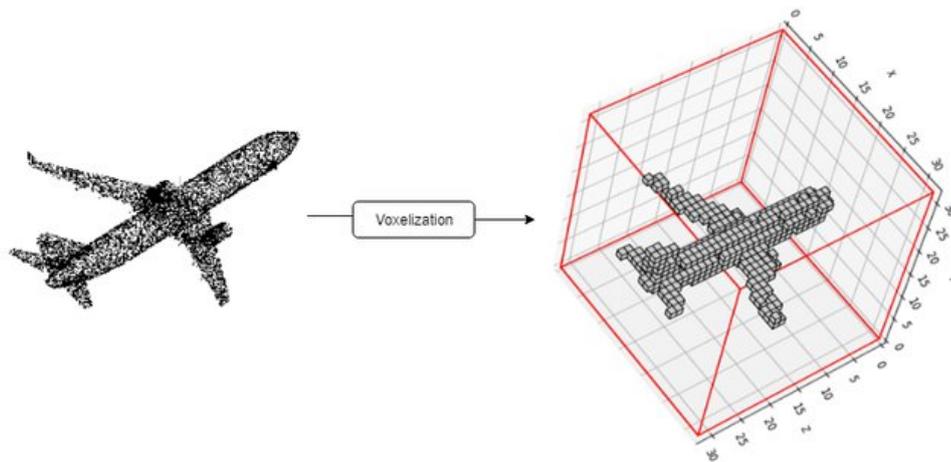


Figura 2.2: Esempio di una point cloud voxelizzata

Come è stato messo in evidenza nell'introduzione, ciascuna point cloud è costituita da un insieme di punti rappresentati nello spazio tridimensionale e quindi ciascun punto viene univocamente determinato sulla base di tre numeri reali che ne identificano le coordinate. Al fine di regolarizzare la struttura e facilitare la trasmissione, tali coordinate vengono mappate in una griglia regolare del tipo $[0, 2^d)^3$, dove d è il numero di bit utilizzati per rappresentare ciascuna coordinata. In particolare, la trasformazione applicata a ciascun punto è:

$$\hat{X}_n = \text{round}\left(\frac{X_n^{\text{orig}} - T}{s}\right) \quad (2.1)$$

dove X_n^{orig} rappresenta il vettore delle coordinate iniziali del punto, T è stabilito sulla base dei valori minimi assunti dalle coordinate originali, con l'obiettivo di andare a centrare l'oggetto, mentre s è un fattore di scala che è determinato dal passo di quantizzazione della point cloud e quindi è collegato al livello di compressione del contenuto. Infatti, per valori più grandi di s si ha che un numero sempre maggiore di punti cadono nell'intervallo tra due interi consecutivi e, quindi, vengono approssimati dalla funzione round con lo stesso intero. Gli attributi di tale voxel vengono, poi, calcolati sulla base di quelli dei punti presenti nell'intervallo.

L'insieme dei processi di trasformazione delle coordinate, rimozione dei punti duplicati e ridefinizione degli attributi viene definito voxelizzazione (Figura 2.2). Questo processo garantisce la presenza di un pattern regolare, molto simile al caso delle immagini, semplificando molto la processazione del contenuto.

Nel caso specifico, sono stati definiti sei diversi livelli di compressione per quanto riguarda la codifica octree e quattro livelli per quel che concerne la codifica trisoup. Le configurazioni utilizzate per le diverse point cloud del dataset considerato sono costanti in quanto tutte sono rappresentate utilizzando lo stesso numero di bit per ciascuna coordinata. Nel caso di octree, i diversi livelli di qualità differiscono tra loro per il valore del parametro di quantizzazione e i valori utilizzati sono riportati nella Tabella 2.1. Per quanto riguarda trisoup, invece, il parametro di quantizzazione è costante (pari a quello utilizzato in r05 per la codifica octree) e la variazione del livello di compressione è determinata da una variazione del massimo numero di suddivisioni octree da effettuare.

Tabella 2.1: Parametro di quantizzazione octree

Compression Level	Position Quantization Scale
r01	0.03125
r02	0.0625
r03	0.125
r04	0.25
r05	0.5
r06	0.75

2.2.2 Codifica Octree

Il modello di codifica octree si basa su una suddivisione ricorsiva in ottanti dello spazio in cui è definita la point cloud, fino al livello di profondità desiderato (Figura 2.3). In particolare, ciascun sottoinsieme viene ulteriormente suddiviso in otto parti solamente se al suo interno è presente almeno un punto, altrimenti diventa una foglia del relativo albero della ricorsione.

Con le opzioni di default del codificatore, la suddivisione procede finché ciascun nodo del livello corrente comprende al massimo un punto al suo interno, in alternativa, impostando opportunamente i parametri di cui dispone il codificatore stesso, è possibile definire la profondità massima dell'albero, ossia il livello di suddivisioni massime che possono essere effettuate: nel caso in cui più punti ricadessero nello stesso intervallo, verrebbero approssimati con un unico punto.

La geometria di ciascun nodo dell'albero comprendente almeno un punto viene codificata con un byte, in cui ciascun bit rappresenta l'occupazione del relativo nodo figlio: se nell'ottante preso in considerazione è presente almeno un punto, allora il bit viene fissato a 1, al contrario assume il valore nullo nel caso di un voxel vuoto. Tali sequenze di bit vengono, poi, ulteriormente codificate con una codifica aritmetica, la stessa utilizzata per codificare tutte le altre informazioni presenti nel bitstream.

Il vantaggio di questo tipo di rappresentazione deriva dal fatto che i voxel pieni sono, generalmente, un numero molto inferiore rispetto a quelli vuoti e questo permette di bloccare la ricorsione ad una profondità molto inferiore, diminuendo drasticamente il numero di byte trasmessi.

Al contrario, la codifica octree diventa poco efficiente nel caso di point cloud molto sparse: in questa circostanza, il numero di byte necessari per poter raggiungere una profondità dell'albero sufficiente a garantire un adeguato livello di qualità è molto superiore rispetto allo spazio necessario per codificare direttamente i punti. In questo caso, è preferibile la codifica diretta *Direct Coding Mode* (DCM), che consiste nell'andare a trasmettere direttamente le coordinate di ciascun punto nello spazio.

Infatti, se non c'è correlazione tra i punti è estremamente sconveniente cercare di codificare la geometria nel modo precedentemente descritto: la ricorsione procederebbe per un numero molto limitato di punti e questo comporterebbe un aumento della quantità di informazione trasmessa, senza un relativo aumento della qualità dell'oggetto considerato. Inoltre, l'assenza di correlazione tra i punti porterebbe anche ad una codifica entropica estremamente poco vantaggiosa.

Al contrario, la codifica diretta non è conveniente nel caso di punti non isolati, perché non sfrutta le caratteristiche dell'oggetto considerato. In genere, il numero di bit utilizzati per codificare ciascuna coordinata garantendo una discreta qualità, è abbastanza elevato: per esempio, nel caso del dataset considerato ciascuna coordinata di ogni punto è rappresentata con 12 bit, quindi nel complesso sono necessari più di 4 byte per codificare direttamente il singolo punto.

Per questo, viene adottata una modalità di codifica ibrida: lo spazio tridimensionale viene suddiviso in ottanti fino a quando il numero di punti all'interno del volume considerato è inferiore ad una certa soglia. Giunti a questo punto, i punti vengono codificati direttamente. L'informazione sulla modalità di codifica utilizzata deve essere necessariamente trasmessa affinché la decodifica della point cloud possa andare a buon fine.

Il processo di decodifica funziona in maniera del tutto simmetrica: prima di tutto viene

decodificata la dimensione dello spazio tridimensionale su cui è definita la point cloud e la stessa struttura dell'octree utilizzata per la codifica viene ricreata. Ogni volta che viene raggiunta una foglia dell'albero viene decodificato il numero di punti compresi in quella zona e un numero di punti pari a quello precedentemente decodificato viene posizionato all'interno dell'ottante considerato.

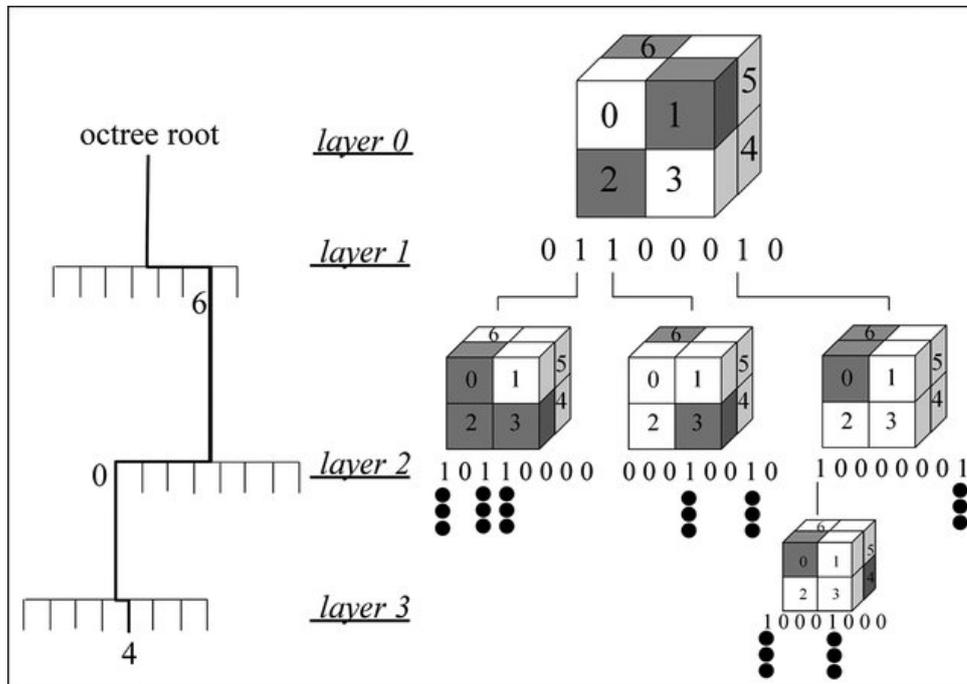


Figura 2.3: Schema di codifica Octree

2.2.3 Codifica Trisoup

Il modello di codifica Trisoup si basa sulla codifica Octree, fornendo delle funzionalità ulteriori che migliorano la qualità percepita del risultato finale a parità di bitrate. In particolare, trisoup applica la suddivisione octree fino ad un livello prestabilito, che può essere impostato modificando opportunamente i parametri dell'encoder. Successivamente, procede cercando di approssimare la superficie della point cloud considerata con una serie di superfici triangolari.

L'approssimazione parte da blocchi che contengono almeno 8 voxel, ottenuti tramite le precedenti suddivisioni e va a determinare la superficie che meglio rappresenta l'insieme dei punti che appartengono a quel cubo. L'intersezione tra questa superficie e i bordi dell'ottante preso in considerazione da origine al più a 12 vertici, uno per ogni lato del solido, che vengono usati per approssimare la superficie inizialmente definita tramite delle superfici triangolari (Figura 2.4).

La codifica avviene in due passi: nel primo passo viene creato un vettore che identifica tra tutti i bordi quelli che contengono effettivamente un vertice, mentre nel secondo passo è codificata la posizione del singolo vertice nel segmento. Il bitstream corrispondente alla geometria, infine, è un insieme della codifica dell'octree, dell'indicatore dei segmenti e della posizione quantizzata dei vertici.

In fase di decodifica, tali superfici vengono ricostruite a partire da sottoinsiemi di tre punti opportunamente scelti (tre punti possono definire univocamente una superficie). Successivamente, da tali superfici vengono estratti dei punti che approssimano quelli di partenza.

Questo metodo di codifica ottiene buoni risultati nel momento in cui i punti sono situati approssimativamente su un piano, mentre se questi sono molto sparsi, a parità di bitrate si otterrà un risultato molto peggiore perché i punti ricostruiti si allontanano da quelli della point cloud originale.

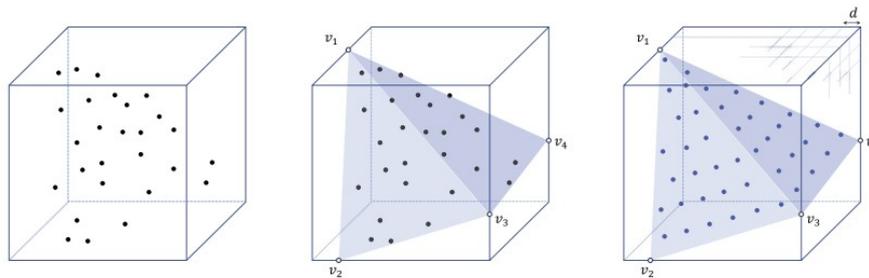


Figura 2.4: Approssimazione planare della superficie di una point cloud

2.3 Metriche di distorsione

Come sottolineato nell'introduzione, la compressione di una point cloud è fondamentale per l'enorme quantità di dati che servono per rappresentarla e parallelamente per l'impossibilità di trasmettere questa mole di informazioni in maniera rapida ed efficiente. Inevitabilmente, la compressione porta ad una perdita di dati, determinando una distorsione e una conseguente diminuzione della qualità percepita: questo ha reso fondamentale lo sviluppo di metriche che permettano di valutare il livello di distorsione introdotto dalla codifica e parimenti la diminuzione della qualità percepita.

A questo scopo, delle metriche soggettive di valutazione della qualità sarebbero lo strumento più adeguato per rilevare la variazione nella percezione del contenuto, tuttavia tali metriche richiedono un grande mole di lavoro, derivante dalla necessità di condurre gli esperimenti soggettivi e fare un'analisi statistica dei risultati ottenuti, per garantire la validità della

misura derivata e, inoltre, una qualsiasi metrica di valutazione soggettiva presenta una serie di criticità che ne ostacolano l'utilizzo.

Da quanto appena esposto, si evince la necessità di sviluppare delle metriche oggettive che permettano di valutare il livello di distorsione di un contenuto compresso.

Nel caso specifico, è stato utilizzato il software MPEG che permette di calcolare alcune delle metriche oggettive suddette: in particolare, vengono fornite due diverse metriche di valutazione della distorsione, che verranno illustrate in seguito, e, per ognuna di esse, viene restituito sia il *Mean Square Error* (MSE) che il PSNR.

La definizione di PSNR applicata è la seguente:

$$PSNR = 10 \log_{10} \left(\frac{3 \times MAX(I)^2}{MSE} \right) \quad (2.2)$$

$MAX(I)$ rappresenta il massimo valore che può essere assunto da ciascuna coordinata, mentre il fattore 3, non presente nella definizione comunemente nota del PSNR, deriva dal fatto che il contenuto preso in considerazione è definito su uno spazio tridimensionale.

Il software può essere configurato impostando opportunamente i parametri messi a disposizione: tra questi è di fondamentale importanza l'opzione che permette di andare a definire la risoluzione, basata sul numero di bit utilizzati per rappresentare ciascuna coordinata. Ulteriori specifiche possono essere fornite, per esempio, per stabilire come comportarsi nel caso di punti equidistanti o punti duplicati.

È importante sottolineare che la scelta delle metriche non è univoca ma, al contrario, esistono diverse strategie per valutare la qualità del contenuto compresso. Questa varietà è in parte determinata dal fatto che la struttura non regolare della point cloud rende più complesso valutarne le differenze e, inoltre, non esiste una corrispondenza completa tra i punti della point cloud originale e di quella compressa. Nel caso bidimensionale, questo obiettivo può essere perseguito in maniera molto più semplice confrontando il valore di ciascun pixel, mentre l'assenza di una corrispondenza univoca tra i punti comporta una maggiore difficoltà.

Al fine di dimostrare la non univocità della scelta, vengono qui riportate due differenti strategie per valutare la corrispondenza tra i punti della point cloud originale e di quella compressa: nel software utilizzato la corrispondenza, come illustrato in seguito, viene effettuata considerando per ciascun punto il relativo punto più vicino nell'altra point cloud. In alternativa, è possibile effettuare un fitting di una superficie quadratica sulla base dei punti più vicini a quello considerato e definire la proiezione di tale punto sulla superficie originata, come viene effettuato in [4].

2.3.1 Distanza Point to Point

La distanza punto a punto è una metrica oggettiva di valutazione del livello di distorsione di una point cloud che si basa sul calcolo della distanza dal punto più vicino. In particolare, per ogni punto appartenente al contenuto non compresso, viene utilizzato un algoritmo di ricerca per trovare il punto più vicino nella corrispondente point cloud considerata. La distanza da tale riferimento viene presa come misura d'errore per il singolo punto, mentre la valutazione complessiva viene effettuata mediando sul risultato ottenuto per ciascun punto:

$$e_{(A,B)}^{p2p} = \frac{1}{N_A} \sum_{a \in A} \| E(i, j) \|^2 \quad (2.3)$$

dove $E(i, j)$ è il vettore errore, definito dalla differenza tra il punto nella point cloud di riferimento e il suo vicino più prossimo nell'oggetto compresso, mentre N_A rappresenta il numero di punti della point cloud originale.

Il principale difetto di questo approccio è il fatto che non tiene in considerazione della disposizione dei punti e quindi della direzione dell'errore: infatti, i punti di una point cloud tendono ad essere disposti in delle superfici e questo comporta che uno spostamento dei punti all'interno della superficie provochi una diminuzione della qualità molto inferiore rispetto ad uno spostamento della stessa quantità nella direzione perpendicolare.

Per sopperire a questa mancanza, è stata sviluppata una nuova metrica (point to surface) che va a definire una superficie che approssima localmente la point cloud e calcola la distanza di ciascun punto rispetto a tale superficie interpolata.

Il problema di questa seconda alternativa è il fatto che la metrica suddetta è estremamente dipendente dal metodo utilizzato per definire la superficie ed è per questo che non viene presa in considerazione.

2.3.2 Distanza Point to Plane

La distanza point to plane è un indicatore della qualità di una point cloud compressa che viene definito nell'ottica di evitare le criticità riscontrate con le due metriche precedentemente illustrate (distanza point to point e distanza point to surface).

Il calcolo di tale metrica si basa sulla ripetizione dello stesso approccio, successivamente descritto, considerando alternativamente una delle due point cloud da confrontare come riferimento e il risultato finale si ottiene mantenendo unicamente il peggiore tra i due valori ottenuti.

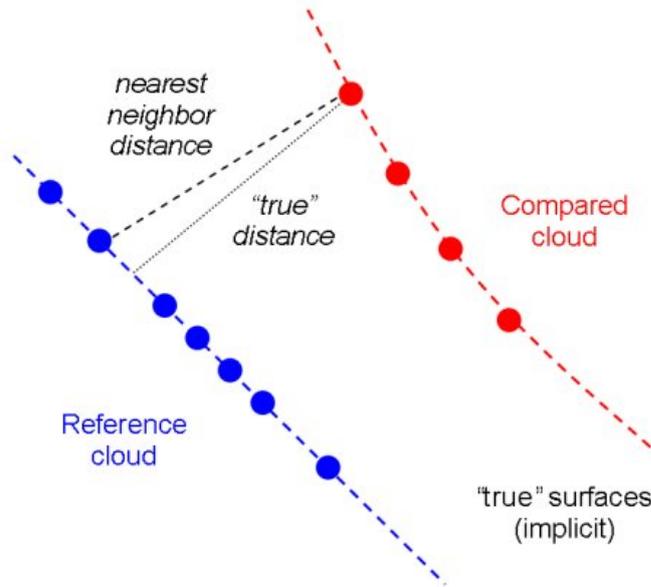


Figura 2.5: Distanza dal punto più vicino

Il procedimento è articolato come segue: innanzitutto, vengono calcolati i vettori normali alla superficie per la point cloud di riferimento, qualora quest'ultima non fosse già fornita di tale attributo. Nel caso del dataset considerato, le point cloud originali ne sono fornite, mentre quelle compresse no. Questo è determinato dal fatto che il metodo di compressione utilizzato codifica unicamente il colore come attributo.

Successivamente, per ogni punto viene definito il suo corrispondente nella point cloud di riferimento. Anche in questo caso, la corrispondenza viene effettuata utilizzando come metodo la distanza dal punto più vicino. Tuttavia, diversamente dalla metrica precedente, non è tale distanza a determinare la bontà del risultato compresso.

Determinato il piano che comprende i due punti in esame e la normale della point cloud di riferimento, definiamo il vettore errore come la proiezione del vettore congiungente i due punti nella direzione parallela al vettore normale. L'errore per il punto suddetto risulta essere la norma di tale vettore (Figura 2.6).

Infine, la misura complessiva della qualità della point cloud compressa viene effettuata mediando rispetto a tutti i punti del contenuto originale:

$$e_{(A,B)}^{p2p} = \frac{1}{N_A} \sum_{a \in A} \| \hat{E}(i, j) \|^2 \quad (2.4)$$

Tale metrica tiene effettivamente in considerazione ciò che, al contrario, la distanza punto-punto ignorava, ossia il fatto che una dispersione dei punti nella direzione normale alla superficie comporta una perdita di qualità del contenuto molto superiore rispetto ad uno spostamento

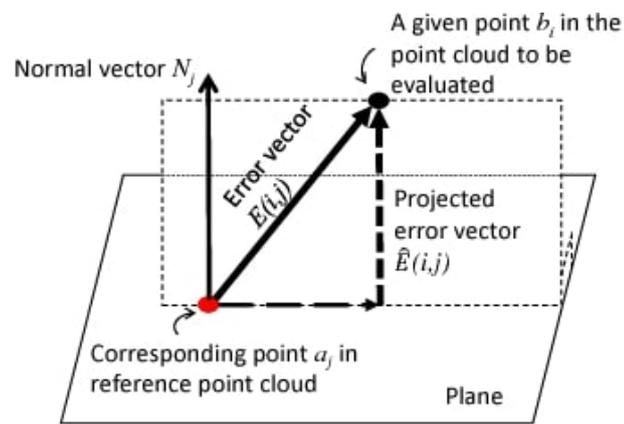


Figura 2.6: Proiezione dell'errore nella direzione normale alla superficie

della stessa intensità parallelamente alla superficie stessa.

Questo aspetto è particolarmente importante nel caso in cui il metodo di codifica utilizzato preveda l'approssimazione di superfici, come trisoup (Sezione 2.2.3). Questo comporta che, in generale, i risultati ottenuti comprimendo con tale metodo siano migliori se la metrica di valutazione utilizzata è la distanza point to plane, piuttosto che point to point.

Capitolo 3

Metodo Proposto

Come è stato evidenziato nell'introduzione, l'obiettivo proposto di questa tesi è quello di sviluppare delle metriche che permettano di descrivere opportunamente la complessità di una point cloud in funzione delle caratteristiche della sua geometria. L'intento è quello di avere una metrica semplice che permetta di riuscire a priori a determinare il miglior compromesso tra il livello di compressione dei dati e la qualità del contenuto ottenuto.

Al fine di perseguire tale scopo, certamente è fondamentale determinare delle caratteristiche che permettano di descrivere adeguatamente la struttura dell'oggetto preso in considerazione, ma è altrettanto indispensabile provare la validità di tali metriche: questo può essere fatto mettendole a confronto con una misura oggettiva di valutazione della qualità di un contenuto compresso.

3.1 Curve di distorsione

Il primo passo che è stato effettuato è l'elaborazione di opportune curve di distorsione che mettano in relazione il livello di compressione con la qualità della point cloud compressa. Tali curve sono state realizzate sulla base dei risultati ottenuti in fase di codifica, quindi il numero di punti che le definiscono dipende dalla quantità di contenuti compressi originati.

Come già riportato in fase di descrizione del software di codifica, sono stati utilizzati sei diversi valori del parametro di quantizzazione nel caso di codifica octree (riportati nella Tabella 2.1), mentre per trisoup i livelli di qualità definiti sono unicamente quattro e sono determinati variando il parametro di codifica `trisoupNodeSizeLog2`, che assume valori interi consecutivi compresi tra 5 (qualità peggiore) e 2 (qualità migliore).

3.1.1 Tasso di compressione

Come misura del tasso di compressione, è stata effettuata la scelta di considerare il rapporto tra il numero di bit necessari per codificare la geometria della point cloud in esame e il numero di punti del contenuto non compresso. Tale decisione non è univoca ma risulta comunque essere quella più comune nella letteratura.

Alcune alternative altrettanto valide per la misura del tasso di compressione sono il numero di bit per unità di volume o la quantità di bit necessari per codificare ciascun voxel. Tuttavia, nel caso del dataset preso in considerazione, non vi sarebbe stata alcuna differenza rispetto alla misura prescelta, essendo tutte le point cloud già voxelizzate.

Il numero di byte necessari per codificare una point cloud viene fornito direttamente dal codec al termine della codifica del contenuto: la decisione presa è quella di considerare unicamente i byte necessari per la codifica della geometria, anch'essi forniti in output dal software utilizzato. Nel caso in cui il numero di slices definite sia superiore a uno, i byte considerati risultano essere la somma dei byte necessari per codificare la geometria di ciascuna slice.

Tale scelta è stata effettuata con l'obiettivo di svincolarsi completamente dal modello di codifica degli attributi, al variare del quale si riscontra un cambiamento del numero di byte necessari per rappresentare il contenuto. Questa differenza è determinata unicamente dal numero di byte utilizzati per codificare il colore ed essendo quest'ultimo completamente irrilevante per lo studio effettuato, risulta più opportuno non tenere in considerazione i byte suddetti.

Quindi, fissato il parametro di quantizzazione q , definiamo $B(q)$ il numero di byte necessari per la codifica della geometria e n il numero di punti che costituiscono la point cloud originale e descriviamo il bitrate del contenuto compresso come:

$$R = \frac{8 \cdot B(q)}{n}. \quad (3.1)$$

Il codificatore fornisce in output anche una misura del bitstream ottenuto per la rappresentazione della geometria, tuttavia tale risultato viene calcolato diversamente da quanto precedentemente illustrato e quindi è stato ritenuto più opportuno non utilizzarlo.

3.1.2 Livello di distorsione

Per quanto concerne il livello di distorsione della point cloud, non è opportuno utilizzare direttamente il parametro di quantizzazione come misura del tasso di distorsione del contenuto

per diverse ragioni: la principale è il fatto che, variando il contenuto originale, cambia anche la qualità del contenuto compresso a parità di parametro di quantizzazione. Infatti, la struttura dell'oggetto può condizionare in maniera diversa la codifica, non consentendo di considerare quest'ultimo come una misura valida della qualità del contenuto compresso.

La scelta della misura di caratterizzazione della qualità del contenuto compresso che è stata effettuata è, quindi, quella di utilizzare il PSNR, che viene fornito in output dal software per la stima del livello di distorsione ed è calcolato come riportato nell'Equazione 2.2.

L'MSE viene definito a seconda della metrica di distorsione scelta, mentre $MAX \{I\}$ fa riferimento al massimo valore che può essere raggiunto da ciascuna coordinata di un punto appartenente all'oggetto, quindi nel caso del dataset considerato è 4095.

La scelta di utilizzare il PSNR piuttosto che servirsi direttamente dello scarto quadratico medio, deriva principalmente da due considerazioni: innanzitutto, perché il PSNR è misura che tiene conto sia dell'errore tra la point cloud originale e quella ricostruita, sia della dinamica dei dati utilizzata per rappresentare la stessa e quindi tiene conto anche della qualità percettiva del risultato ottenuto, diversamente da quanto fatto dall'MSE, che, di conseguenza, generalmente risulta meno accurato nella valutazione della distorsione del contenuto. In secondo luogo, il vantaggio dell'utilizzo del PSNR deriva dalla scala logaritmica utilizzata per definirlo. La scala dei decibel consente di rappresentare grandi variazioni di valore in modo più compatto. Le variazioni nella qualità o nell'errore possono essere espressi in un intervallo più ristretto di numeri, semplificando l'interpretazione visiva. Queste motivazioni fanno in modo che, in generale, nella letteratura sia più comune la scelta del PSNR come metrica di valutazione della distorsione, anche se in alcuni casi l'accuratezza del MSE risulta sufficiente.

Il software utilizzato fornisce una duplice misura del PSNR, ottenuta considerando di volta in volta una delle due point cloud confrontate come riferimento. Per evitare inconvenienti che possono accadere nel determinare la corrispondenza tra le due, si considera sempre il peggiore dei due risultati ottenuti.

3.1.3 Curve ottenute

Le curve di distorsione ottenute utilizzando le metriche precedentemente descritte, sono riportate in seguito nella Figura 3.1.

Per completezza, sono state riportate le curve ottenute considerando come metrica di distorsione sia la distanza point to point (Sezione 2.3.1) che la distanza point to plane (Sezione 2.3.2). Come si può vedere, l'andamento delle curve è molto simile nei due casi e questo com-

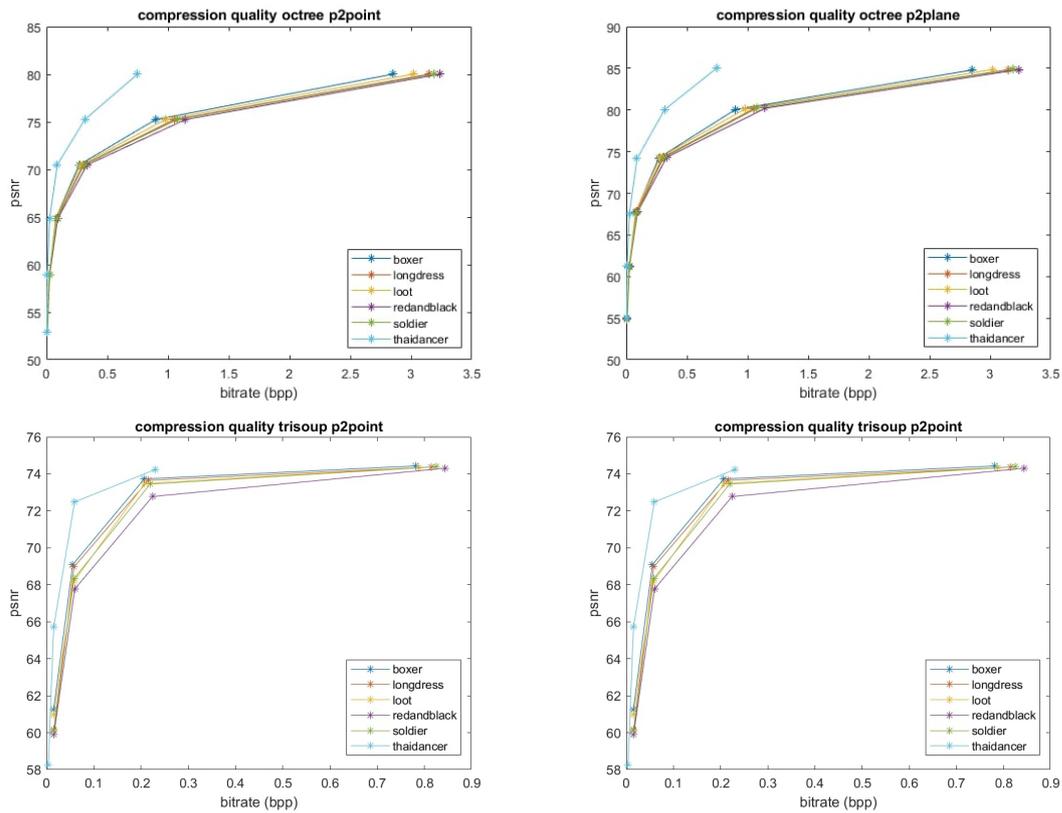


Figura 3.1: Curve di distorsione del dataset 8iVSLF

porta che sia abbastanza ininfluenza la scelta della metrica di distorsione da considerare. A dimostrazione di quanto appena detto, vengono qui riportati i risultati della correlazione tra la distorsione rilevata con le due diverse metriche, sia nel caso di codifica octree che nel caso di codifica trisoup.

Tabella 3.1: Octree

	p2point	p2plane
p2point	1.0	0.9996
p2plane	-	1.0

Tabella 3.2: Trisoup

	p2point	p2plane
p2point	1.0	0.9960
p2plane	-	1.0

Correlazione tra le misure di distorsione

In entrambi i casi, la qualità rilevata è superiore nel caso in cui la metrica utilizzata sia la distanza point to plane: questo era prevedibile perché tale valore rileva la dispersione dei punti solo nella direzione normale alla superficie, trascurando eventuali variazioni parallelamente ad essa. Come discusso anche nella sezione di descrizione del software di valutazione della distorsione, questo approccio è maggiormente coerente con la qualità percettiva rilevata: infatti, uno scattering nella direzione della superficie non comporta visivamente una grande perdita di definizione del contenuto. Sulla medesima idea si basa anche il metodo di codifica trisoup, i cui risultati, quindi, risultano particolarmente adatti ad essere descritti da questa metrica.

Ciò nonostante, l'estrema correlazione degli esiti ottenuti permette di considerare indistintamente le due metriche ai fini dello studio in corso.

Osservando le curve, si può notare un andamento molto simile per tutte le point cloud appartenenti al dataset, con l'unica eccezione rappresentata da thaidancer. L'estrema vicinanza può essere giustificata dal fatto che gli oggetti sono molto simili, sia per quanto riguarda il tipo di contenuto che per la densità. Il diverso andamento di thaidancer potrebbe essere causato dal fatto che tale point cloud è definita su uno spazio molto inferiore rispetto alle altre (circa 1/8 dello spazio messo a disposizione).

Come verrà sottolineato in seguito, tale somiglianza potrebbe in parte ostacolare il progetto di validazione delle misure di complessità proposte. Infatti, risulta molto più complesso dimostrare l'efficacia di tali metriche su point cloud con caratteristiche affini, in quanto i risultati ottenuti differiranno di poco gli uni dagli altri, rendendo più difficile il confronto.

3.2 Misura della qualità di compressione

Dai grafici riportati in Figura 3.1, è possibile dedurre la difficoltà nella compressione di ciascuna point cloud, tuttavia non è sufficiente ai fini dello studio proposto. Infatti, da essi possiamo stimare quale tra i contenuti del dataset risulta più critico, ma non abbiamo a disposizione una misura univoca della complessità di compressione che possa fungere da banco di prova per la correttezza della metrica di valutazione della geometria di ciascuna point cloud che verrà proposta. Senza tale metrica non sarebbe possibile validare i risultati ottenuti in quanto potremmo unicamente confrontarli con i grafici sopracitati, stabilendo se l'ordine rilevato corrisponde con quello determinato nella pratica: questo non è sufficiente perché comporterebbe, da una parte, l'impossibilità di definire qual è la differenza di costo di compressione tra point cloud del suddetto dataset e dall'altra l'assenza di una misura che descriva intrinsecamente il singolo oggetto, senza la necessità di un confronto.

Stabilita la necessità di tale misura della complessità di compressione, il punto di partenza nella sua definizione è stato lo studio analogo effettuato nel caso bidimensionale [6]: per quanto riguarda le immagini, la complessità di compressione viene stimata basandosi sul bitrate, ossia il numero di bit necessari per rappresentare l'immagine. Infatti, come riportato nell'introduzione, la validità dell'informazione spaziale come misura della complessità del contenuto bidimensionale è stata provata valutando la sua correlazione con il tasso di compressione, definito come rapporto tra la dimensione del contenuto compresso e di quello originale, che viene valutato fissando un valore predefinito del parametro di quantizzazione.

Tale approccio presenta principalmente tre criticità:

1. Non è una misura generale della difficoltà di compressione ma, al contrario, valuta tale complessità solamente per un valore specifico del parametro di quantizzazione. In altre parole, ci permette di valutare i risultati ottenuti soltanto per il valore considerato ma non garantisce nulla nel caso in cui il passo di quantizzazione sia differente.
2. Fissato il parametro di quantizzazione non è stabilita univocamente anche la qualità del contenuto risultante: lo stesso valore del parametro di quantizzazione può provocare livelli di distorsione molto differenti a seconda della struttura diversa delle immagini considerate. Maggiore è l'intensità delle alte frequenze nello spettro del contenuto considerato, maggiore sarà la perdita di qualità causata dalla compressione. Una considerazione analoga può essere fatta anche nel caso delle point cloud, per cui non sarebbe del tutto corretto valutare la complessità del contenuto sulla base del bitrate raggiunto, se la qualità dei diversi oggetti confrontati non è la stessa.
3. Nel caso delle point cloud, risulterebbe estremamente complicato effettuare questo tipo di misura in quanto livelli di compressione differenti si ottengono al variare di più parametri del codificatore, a seconda del metodo di compressione prescelto. Significativo è il caso della compressione effettuata utilizzando il metodo trisoup, precedentemente descritto: i diversi livelli di qualità sono ottenuti variando il numero di suddivisioni oc-tree prima di effettuare l'approssimazione delle superfici, senza modificare il parametro di quantizzazione.

Viste le criticità della misura precedentemente illustrata, in questa tesi non è stato ritenuto opportuno utilizzare il bitrate a parità di parametro di quantizzazione come misura della complessità di compressione di una point cloud.

Nel caso bidimensionale viene proposta come alternativa una metrica che tenga in considerazione il rapporto tra le dimensioni del contenuto originale e di quello codificato senza perdite. In questo modo ci si svincola dalla valore del parametro di quantizzazione scelto e si ottiene una misura altrettanto valida e molto correlata alla precedente. Tuttavia, nel caso tridimensionale questo risulta impraticabile perché non tutti i metodi di codifica prevedono la possibilità di non introdurre distorsione.

3.2.1 Bitrate a parità di livello di qualità

La prima alternativa valida alla misura precedentemente descritta, consiste nell'andare a valutare il bitrate a parità di qualità del contenuto compresso. Questa metrica risulterebbe più efficace della prima considerata in quanto risolverebbe il secondo dei problemi descritti in precedenza, ma manterrebbe comunque numerose criticità: innanzitutto, si tratta comunque di una misura non generale della complessità di compressione, in quanto ci permette di fare delle valutazioni unicamente per il livello di qualità considerato, senza la garanzia di poter generalizzare. In secondo luogo, tale misura sarebbe operativamente molto complessa da calcolare perché la qualità del contenuto compresso può essere valutata soltanto a posteriori, quindi sarebbe necessario determinare il valore del parametro di quantizzazione per la singola point cloud che permette di avvicinarsi maggiormente alla qualità richiesta. Questo può avere un costo computazionale estremamente elevato e, in alcuni casi, risulta addirittura impossibile: ciò significa che non sempre è possibile stabilire una configurazione del codificatore che permetta di raggiungere uno specifico livello di qualità per ogni contenuto. Per esempio, nel caso del metodo di codifica trisoup presente nel codificatore utilizzato, i livelli di qualità considerati sono ottenuti per valori consecutivi del parametro modificato e questo implica che non sarebbe possibile definire delle configurazioni intermedie che permettano di ottenere i livelli di qualità desiderati.

Anche in questo caso, quindi, la misura descritta non è del tutto adatta per valutare la complessità di compressione.

Le stesse difficoltà si riscontrano anche nel caso in cui si andasse a valutare la qualità del contenuto compresso a parità di bitrate ottenuto, in quanto anche quest'ultimo può essere stabilito solamente dopo aver eseguito la codifica.

3.2.2 Bitrate Medio

Con l'obiettivo di ottenere una misura più generale della complessità di compressione di un contenuto, è stata presa in considerazione la possibilità di andare a valutare il bitrate medio, definito mediando sui livelli di qualità utilizzati. Quindi, nel caso specifico, sono stati considerati sei diversi valori quando la codifica utilizzata è octree mentre solo quattro se il metodo scelto è trisoup. In questo modo, otteniamo una misura molto più generale della complessità di compressione in quanto non consideriamo unicamente uno dei risultati ottenuti.

Tuttavia, seguendo questo approccio permane il problema determinato dal fatto che, a parità di parametro di quantizzazione, i livelli di qualità valutati non sono costanti tra le point

cloud. Questo aspetto, fortunatamente, risulta abbastanza marginale nel caso in cui il metodo utilizzato sia octree, in quanto, per il dataset in esame, c'è una forte corrispondenza tra il parametro di quantizzazione e la distorsione. Invece, per quanto riguarda la codifica trisoup, la differenza risulta essere più marcata, inficiando in parte la validità della misura. Ciononostante, i risultati ottenuti rispecchiano quanto mostrato dalle curve di distorsione precedentemente illustrate.

Al fine di estendere la validità di questa misura in modo che possa essere un efficace strumento di confronto anche tra point cloud molto diverse tra loro, è utile cercare di svincolarsi dai livelli di qualità raggiunti. Come già sottolineato in precedenza, è pressoché impossibile modificare la configurazione del codificatore per raggiungere la qualità desiderata per ogni contenuto, quindi è necessario procedere diversamente per cercare di pervenire al risultato desiderato.

3.2.3 Integrale della curva di distorsione interpolata

La soluzione proposta per svincolarsi dai risultati specifici ottenuti comprimendo le point cloud del dataset è quella di interpolare le curve di distorsione ottenute. Partendo dal presupposto che le curve di distorsione debbano necessariamente essere non decrescenti, è ragionevole considerare l'andamento del livello di distorsione abbastanza regolare nell'intervallo di definizione in modo tale da assicurare la validità della soluzione proposta.

A questo scopo, sono stati considerati due diversi metodi di interpolazione delle curve di distorsione:

- Interpolazione spline
- Interpolazione pchip

Spline è un metodo d'interpolazione a tratti, che non definisce un unico polinomio di grado molto elevato per cercare di approssimare l'insieme dei punti di input, ma divide tali punti in sottoinsiemi che vengono approssimati da polinomi di grado limitato, generalmente di terzo grado.

È importante notare che l'interpolazione con spline non calcola una curva separata per ciascun gruppo di dati, ma utilizza tutti i punti forniti per calcolare una serie di funzioni cubiche che si adattano ai dati nel complesso. Questo significa che il numero di punti di controllo e quindi di spline cubiche calcolate dipende dal numero di punti forniti e dalla loro disposizione. In ogni caso, non è possibile effettuare l'interpolazione di una curva con questo metodo nel caso in cui i punti a disposizione siano in numero inferiore a quattro.

L'interpolazione con spline è, in generale, preferibile rispetto ad un'interpolazione polinomiale, anche se fatta considerando polinomi di grado sufficientemente elevato, in quanto tende a produrre un errore più contenuto rispetto alla curva originale e, inoltre, tende a fornire risultati più regolari.

Tuttavia, nel caso in cui i punti a disposizione non fossero sufficienti non è possibile interpolare in questo modo e conviene servirsi di una tecnica polinomiale, oppure abbassare il grado del polinomio utilizzato per approssimare ciascun sottoinsieme di punti (interpolazione spline lineare).

L'interpolazione pchip è, allo stesso modo, una tecnica di interpolazione a tratti che approssima ciascun sottoinsieme di punti con una funzione di terzo grado, detta spline cubica di Hermite. A differenza delle spline cubiche tradizionali, che si basano solo sui valori delle coordinate dei punti considerati, le spline cubiche di Hermite prendono in considerazione anche le derivate di tali punti.

Una spline cubica di Hermite viene definita da coppie di punti di controllo, ciascuna contenente le coordinate x e y dei punti, insieme alle rispettive derivate, che sono utili per rappresentare pendenze o variazioni locali nella funzione.

Tale metodo garantisce che la funzione interpolata attraversi i punti di controllo e che le derivate negli stessi punti corrispondano alle derivate specificate. Ciò consente un controllo più preciso sull'interpolazione della funzione, garantendo dei risultati tendenzialmente più regolari e limitando, in particolare, le oscillazioni della curva risultante. Tuttavia, le derivate nei punti considerati non sono sempre disponibili e il loro calcolo può aumentare la complessità dell'interpolazione.

La principale sfida nell'interpolazione delle curve di distorsione ottenute è determinata dalla limitatezza dei punti a disposizione. Per questo motivo non è stata presa in considerazione un'approssimazione polinomiale che richiede un numero di punti superiore rispetto a quelli a disposizione. Questa difficoltà risulta evidente, in particolare, per le curve che descrivono i risultati ottenuti con trisoup, in quanto i punti a disposizione sono unicamente quattro. Questo comporta delle grandi oscillazioni nelle curve ottenute utilizzando il primo metodo descritto, tali da non renderle minimamente una buona approssimazione dell'andamento della distorsione.

A titolo esemplificativo, sono riportate in Figura 3.1 le curve risultanti per una singola point cloud (boxer) a seguito dell'interpolazione effettuata utilizzando le due diverse tecniche illustrate. Com'era prevedibile, nel caso di octree le differenze sono pressoché nulle mentre per trisoup l'interpolazione spline introduce un errore estremamente elevato, tanto da portarci a

non considerare più tali risultati nel seguito.

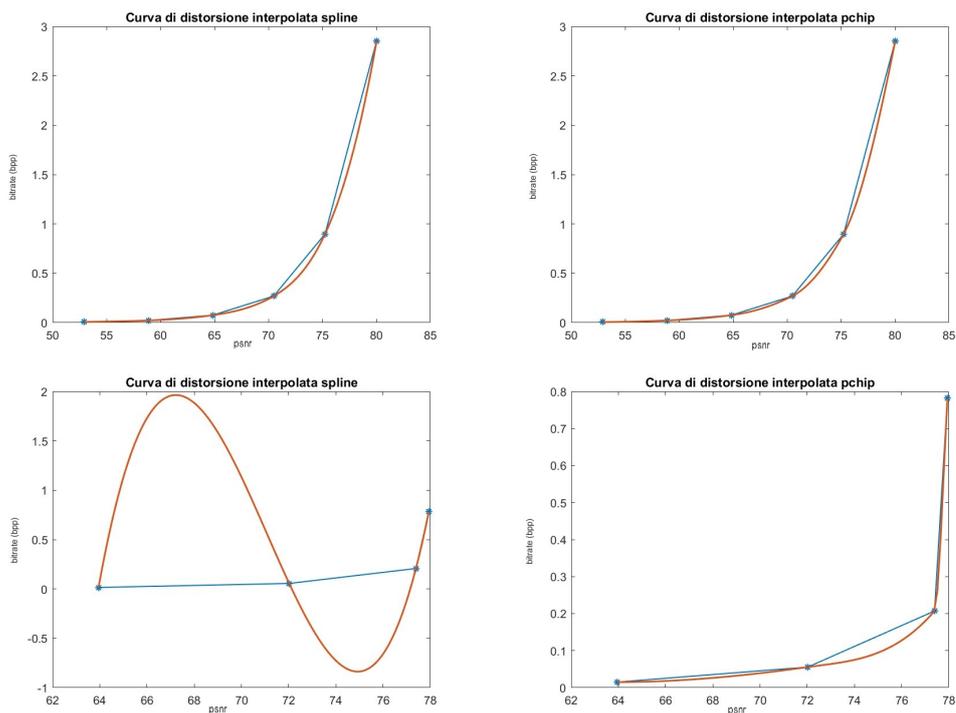


Figura 3.2: Curve di distorsione interpolate: in alto i grafici interpolati relativi alle curve octree, mentre in basso i medesimi risultati relativi a trisoup

A questo punto, è necessario un metodo di valutazione della curva ottenuta con l'obiettivo di ottenere una misura univoca che rappresenti la complessità di ciascuna point cloud. La scelta più ragionevole per raggiungere tale obiettivo è considerare l'integrale della curva interpolata. In questo modo, il risultato che si ottiene è una valutazione complessiva della complessità, svincolata dai valori specifici di distorsione raggiunti, pur rispettandone l'andamento.

Come detto in precedenza, valori di PSNR elevati corrispondono a point cloud compresse con un livello di distorsione basso. Questo implica che un contenuto può essere considerato poco complesso se riesce a introdurre un errore di compressione relativamente piccolo, utilizzando un numero di bit molto limitato: di conseguenza, ci si aspetta che integrando rispetto alla misura di qualità di contenuto, valori bassi dell'integrale corrispondano a point cloud con complessità limitata.

Ciò che rimane da definire è l'intervallo d'integrazione della curva. Per il dataset preso in considerazione si tratta di un problema abbastanza marginale, soprattutto se la codifica scelta è octree, in quanto il supporto su cui sono definite le curve di distorsione è quasi analogo per tutte le point cloud, vista l'estrema somiglianza dei contenuti considerati. In particolare, si rileva una differenza massima di supporto di 0,04 dB nel caso di Octree e 3,14 dB nel caso di trisoup, che risulta più difficilmente trascurabile.

Una soluzione possibile è quella di determinare un intervallo d'integrazione comune tra le curve. Nel seguito sono riportati diversi approcci per stabilire l'intervallo su cui integrare le curve e i rispettivi risultati ottenuti per il dataset utilizzato. Ancora una volta, è necessario ricordare che tali point cloud sono molto simili tra loro e di conseguenza i problemi esposti si rilevano solo marginalmente. Tuttavia, andando a considerare un dataset più ampio e soprattutto più vario, alcune delle scelte potrebbero risultare impraticabili.

I principali aspetti da tenere in considerazione nell'andare a definire un intervallo d'integrazione sono due: da una parte, è utile determinare un intervallo che sia sufficientemente ampio per avere una misura di complessità abbastanza generale, che tenga in considerazione diversi livelli di qualità. In secondo luogo, è fondamentale non andare ad integrare le curve in una regione lontana dal supporto di definizione iniziale.

Questa necessità deriva dal fatto che la curva è stata ottenuta interpolando rispetto ai punti inizialmente a disposizione, quindi nella regione compresa tra tali punti approssima l'andamento della distorsione in maniera sufficientemente precisa, mentre al di fuori di tale intervallo niente assicura che l'approssimazione sia altrettanto buona. Al contrario, è possibile ottenere un andamento completamente opposto rispetto a quello effettivo e questo comporta che l'inclusione di una regione non appartenente al supporto iniziale della curva potrebbe inficiare la misura. A titolo esemplificativo, sono riportati i risultati ottenuti integrando una delle point cloud del dataset in un intervallo superiore a quello di definizione: come si può osservare in Figura 3.3, la curva non risulta predire bene l'andamento della distorsione, tanto da risultare decrescente al di fuori dell'intervallo di definizione.

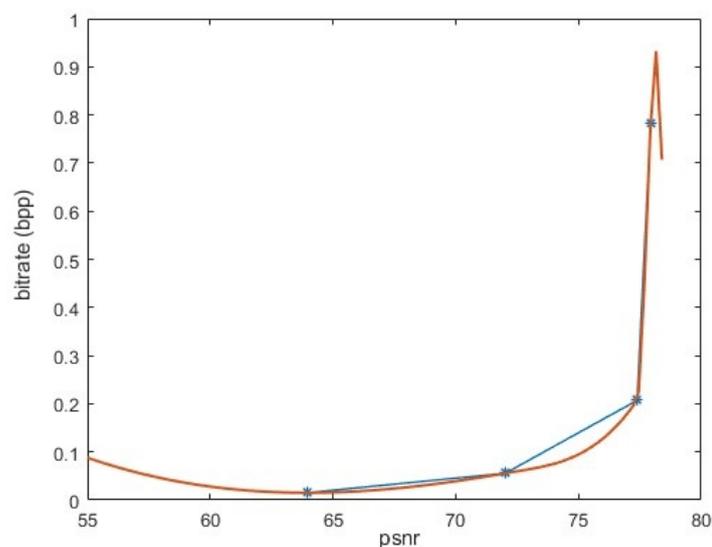


Figura 3.3: Interpolazione della curva di distorsione al di fuori dell'intervallo di definizione

Purtroppo questi aspetti risultano collidere tra loro nel momento in cui si tenga in considerazione un dataset sufficientemente vario. Infatti, la qualità del risultato finale che si ottiene può essere molto diversa a seconda del contenuto considerato e questo comporta che per avere un intervallo d'integrazione sufficientemente ampio sia necessario integrare al di fuori del supporto iniziale delle curve.

La prima soluzione che viene proposta è quella di andare ad integrare tutte le curve sull'intervallo comune a tutti i supporti iniziali, definito come l'intervallo che ha come estremo inferiore il massimo dei valori di qualità ottenuti per il livello peggiore (livello 1) e come estremo superiore il minimo dei valori di PSNR corrispondenti al livello migliore (livello 4 per trisoup e 6 per octree), in modo che tutte le curve siano integrate in un intervallo appartenente al loro supporto iniziale.

Tale soluzione funziona estremamente bene per il dataset preso in considerazione poiché la quasi totale sovrapposizione dei supporti comporta che l'intervallo comune ha un ampiezza sufficientemente elevata, garantendo degli ottimi risultati. Tuttavia, può essere abbastanza problematico in situazioni che comprendano una variabilità maggiore: infatti, è possibile che i supporti delle diverse curve siano disgiunti portando ad avere l'estremo d'integrazione inferiore più grande dell'estremo superiore. Questo comporta una misura dell'integrale negativa, ottenendo un risultato opposto rispetto a quello sperato, per cui le curve con integrale superiore sono le meno complesse.

La seconda soluzione proposta è quella di non definire un intervallo univoco per ciascuna point cloud da analizzare, ma andare a definire univocamente l'ampiezza di tale intervallo. Il vantaggio di questo tipo di approccio è che l'integrale viene effettuato sempre all'interno di una regione sufficientemente ampia appartenente al supporto iniziale delle curve, evitando eventuali problemi derivanti dall'interpolazione. Di contro, invece, l'integrale non viene effettuato in un intervallo comune a tutte le curve e questo comporta che le differenze di bitrate che si riscontrano possono essere determinate sia da differenze nella complessità del contenuto ma anche da differenze nella qualità raggiunta.

Un'ulteriore alternativa possibile consiste nell'andare a determinare un intervallo d'integrazione univoco, ossia un intervallo d'integrazione assoluto ragionevole ma svincolato da qualsiasi considerazione sulle point cloud considerate. L'aspetto maggiormente positivo di questo tipo di approccio è l'assoluta generalità della misura, che non è minimamente legata al dataset preso in considerazione e quindi risulta altrettanto valida per confrontare contenuti differenti. Dall'altra parte, invece, ci sono due criticità principali: la prima consiste nella strategia da utilizzare per determinare questo intervallo, mentre la seconda è collegata al fatto

che l'assoluta generalità della misura non fornisce garanzie riguardo al fatto che l'intervallo d'integrazione appartenga al supporto iniziale delle curve. Questo comporta che c'è la possibilità di andare ad integrare in delle regioni non comprese tra i punti iniziali, correndo i rischi precedentemente descritti.

L'ultima soluzione proposta consiste nell'andare a definire un intervallo di integrazione comune, centrato sul valore medio ottenuto considerando tutte le misure di PSNR a disposizione. L'ampiezza di tale intervallo può essere definita in diversi modi: se viene scelta in modo che corrisponda alla minima delle ampiezze dei supporti delle curve, si raggiunge l'obiettivo di integrare ciascuna di esse in un intervallo che non si discosti troppo dalla regione di definizione, evitando tutti i rischi precedentemente descritti.

In questo modo, si ottiene una misura della complessità di compressione generale, dato che viene considerato un intervallo d'integrazione sufficientemente ampio, e univoca, in quanto l'intervallo definito è lo stesso per tutte le curve.

L'unico aspetto ancora da considerare è la non univocità della misura, che rimane strettamente legata al dataset valutato. Tuttavia, non è detto che questo rappresenti un problema poiché tale misura viene definita unicamente per confrontare e validare i risultati specifici del dataset in questione. D'altra parte, considerando un dataset differente, è possibile ripetere la stessa misura, ottenendo una metrica di valutazione analoga: infatti, l'intervallo d'integrazione cambia tra due insiemi diversi ma è costante all'interno dello stesso dataset.

Le grandezze precedentemente descritte risultano essere tutte delle valide alternative per stimare la complessità di compressione di una point cloud. Come sottolineato precedentemente, ciascuna di esse risulta essere più adatta delle altre per certi aspetti ma ha delle criticità che potrebbero in qualche misura inficiare il risultato ottenuto.

Nonostante ciò, i risultati sperimentali che si ottengono applicando tali metriche al dataset (Figura 2.1) rendono evidente che, nel caso specifico, c'è una correlazione quasi totale tra le diverse misure di complessità di compressione, rendendo la valutazione della complessità abbastanza indipendente dal modo in cui viene calcolata.

	a	b	c	d
a	1.0000	1.0000	1.0000	0.9993
b	-	1.0000	1.0000	0.9994
c	-	-	1.0000	0.9994
d	-	-	-	1.0000

Tabella 3.3: Correlazione delle misure di complessità per octree: (a) integrale centrato sulla media, (b) integrale sul supporto comune, (c) integrale sul supporto singolo, (d) bitrate medio

In particolare, nel caso di octree (Tabella 3.3), la precisione utilizzata nel calcolo della misura di complessità di compressione non permette di rilevare differenze tra le tre differenti metodologie per il calcolo dell'integrale della curva interpolata: questo è determinato dal fatto che tutte le curve hanno un supporto pressoché costante e ciò comporta che le diverse scelte riguardo all'intervallo d'integrazione tendano a coincidere. Anche la misura del bitrate medio risulta descrivere molto bene la complessità: questo era prevedibile in quanto la qualità è quasi completamente fissata a parità di parametro di quantizzazione, rendendo, di conseguenza, la misura valida.

Al contrario, nel caso in cui il metodo di codifica sia trisoup (Tabella 3.4), i risultati ottenuti sono più significativi: infatti, risulta evidente che il bitrate medio è meno adatto a descrivere la complessità di compressione e questo è determinato dal fatto che non è possibile fissare la qualità del contenuto compresso ma unicamente la configurazione del codificatore e, a parità di parametri utilizzati, non si ottengono gli stessi risultati.

Invece, per quanto riguarda le diverse misure proposte per la valutazione della curva interpolata, la correlazione risulta molto più elevata, anche se non quanto riscontrato nel caso di octree. Tale difformità è determinata dal fatto che le differenze tra i supporti delle curve sono più marcate e questo comporta che le diverse scelte sull'intervallo d'integrazione abbiano un impatto sui risultati ottenuti. Ciò nonostante, la grande somiglianza dei contenuti considerati rende questa differenza estremamente lieve.

	a	b	c	d
a	1.0000	0.9860	0.9877	0.7846
b	-	1.0000	0.9996	0.8605
c	-	-	1.0000	0.8471
d	-	-	-	1.0000

Tabella 3.4: Correlazione delle misure di complessità per trisoup: (a) integrale centrato sulla media, (b) integrale sul supporto comune, (c) integrale sul supporto singolo, (d) bitrate medio

Come sottolineato in precedenza, la limitata varietà del dataset preso in considerazione non permette di rilevare del tutto gli aspetti più critici delle misure proposte. Questo, da una parte, potrebbe essere un problema in quanto non permette di definire univocamente la migliore misura di complessità di compressione.

Tuttavia, tale misura è necessaria nello studio in corso unicamente come banco di prova per la validità e l'efficacia delle misure di predizione della complessità geometrica di ciascun contenuto. Ne segue che è fondamentale che tale metrica descriva adeguatamente la com-

plexità di compressione per il dataset preso in considerazione, ma non è necessario che sia la misura che funziona meglio in assoluto.

3.3 Misura della complessità

Come già sottolineato nell'introduzione, l'intento di questa tesi è quello di andare a definire delle semplici misure di valutazione della complessità di una point cloud sulla base della sua struttura geometrica. Al fine di perseguire questo obiettivo, prima di tutto, è necessario andare ad effettuare delle ipotesi riguardo alle possibili misure di predizione e in seguito queste dovranno essere validate, andandole a confrontare con la grandezza precedentemente definita.

Tali ipotesi sono state effettuate tenendo in considerazione diversi aspetti: da una parte è importante considerare il modello di codifica utilizzato perché le peculiarità di quest'ultimo sono certamente molto legate alle caratteristiche geometriche della point cloud che ne facilitano la codifica. D'altra parte, però, sarebbe scorretto basare le misure suddette unicamente sul codificatore scelto, in quanto la misura che questa tesi si propone di definire dovrebbe descrivere la complessità intrinseca dell'oggetto, indipendentemente dalla codifica utilizzata. Di conseguenza, è utile anche vagliare delle ipotesi basate sulla struttura e le caratteristiche della point cloud in sé al fine di determinare una misura quanto più generale possibile.

Nel seguito, viene proposta una descrizione approfondita delle principali misure candidate a valutare la complessità di una point cloud, evidenziandone i motivi che hanno portato ad effettuare tali ipotesi.

3.3.1 Distanza dai punti più vicini

La prima misura che viene proposta è la distanza da un numero predefinito di punti vicini, che consiste nell'andare a valutare la distanza euclidea tra il punto della point cloud preso come riferimento e i punti ad esso contigui. È possibile, in seguito, ottenere una misura univoca di valutazione della complessità dell'oggetto facendo un'analisi statistica dei risultati ottenuti: l'approccio da utilizzare può essere variabile e diverse alternative verranno vagliate successivamente.

La distanza dai punti più vicini risulta essere una metrica di valutazione della densità di una point cloud. Infatti, un contenuto può essere definito denso nel momento in cui è costituito da una grande mole di punti concentrati in uno spazio ristretto: questo comporta che la distanza

media tra i punti e conseguentemente la misura proposta, debbano necessariamente avere un valore ridotto nel caso di point cloud dense.

Tale misura viene spesso utilizzata nella letteratura, in particolare nella trasformazione e nell'analisi dei contenuti: infatti, esistono diverse misure della struttura di una point cloud basate sull'insieme di punti più vicini. Inoltre, è necessario sottolineare che la definizione della distanza dal punto più vicino non è così banale come nel caso dei contenuti bidimensionali e questo è determinato prevalentemente da due fattori: da una parte bisogna considerare che, come sottolineato nella descrizione iniziale delle point cloud, la distribuzione dei punti non è uniforme e questo comporta che ci possano essere delle zone estremamente ricche di punti mentre altrove questi siano molto lontani tra loro. Dall'altra parte, è necessario ricordare anche che nella rappresentazione matriciale dei punti, generalmente non è vero che punti con indici vicini corrispondano a punti fisicamente vicini, definiti come punti aventi coordinate simili, e questo comporta che la definizione dell'insieme dei punti più vicini non risulta essere un problema di immediata soluzione.

In generale, è ragionevole pensare che punti molto vicini tra loro siano più facili da codificare, in quanto è più facile definire le relazioni spaziali e le informazioni da essi portate sono molto correlate: questo è dimostrato anche dal fatto che il codificatore descritto, nel caso in cui il contenuto sia eccessivamente sparso, effettua la scelta di inviare direttamente le coordinate dei punti, in quanto qualsiasi modello di codifica non risulta efficiente in questa circostanza.

Inoltre, l'idea di andare a considerare la distanza dai punti più vicini come metrica di valutazione della complessità di una point cloud può essere giustificata anche facendo riferimento al principio di codifica octree: se la distanza tra i punti è molto ridotta, il segnale che si ottiene in uscita dal quantizzatore sarà molto sparso e quindi più facile da codificare e trasmettere con un numero limitato di bit. Infatti, punti molto vicini vengono approssimati allo stesso modo dalla funzione round precedentemente menzionata e questo implica che possono essere descritti attraverso un unico voxel.

Definite le ragioni che hanno portato alla scelta di andare a valutare la distanza dal punto più vicino come metrica per stimare la complessità di un contenuto, è necessario andare a definire una strategia per determinare tale distanza in maniera efficiente. Infatti, l'algoritmo di forza bruta prevede di andare a calcolare per ciascun punto di riferimento la distanza da tutti gli altri punti della point cloud e successivamente ordinare i risultati in ordine crescente di distanza, andando a definire i punti più vicini. Tuttavia, le point cloud hanno, generalmente, diverse centinaia di migliaia di punti e questo implica che la procedura precedentemente descritta non sia applicabile nella pratica, in quanto risulta computazionalmente troppo costosa.

Si è, quindi, reso necessario utilizzare un algoritmo più efficiente, che permetta di perseguire lo stesso scopo in minor tempo.

Il punto di partenza è quello di servirsi di un kd-tree (k dimensional tree), che è una struttura dati utilizzata per organizzare i punti in uno spazio di dimensione k ed è particolarmente utile per accelerare le operazioni di ricerca e vicinato nel caso di contenuti multidimensionali.

L'idea principale di un kd-tree è quella di creare una struttura ad albero bilanciato in cui i punti sono suddivisi in modo ricorsivo in sottoinsiemi in base alle loro coordinate. In particolare, il numero di nodi definiti a ciascun livello è variabile, con l'obiettivo di fare in modo che ciascuno di essi includa un numero limitato di punti. La costruzione di un kd-tree, quindi, comporta il partizionamento dei punti in modo ricorsivo fino al livello in cui ogni foglia dell'albero rappresenta un unico punto o un insieme molto piccolo di punti. Questo processo crea una struttura ad albero bilanciata, in cui i punti vicini nello spazio sono solitamente raggruppati insieme.

La ricerca dei punti più vicini è strutturata come segue:

- **Costruzione del kd-tree:** Innanzitutto, il kd-tree viene costruito utilizzando il dataset di punti disponibili. Questo consiste nell'andare a definire ricorsivamente l'albero, suddividendo i punti in base alle coordinate in modo bilanciato.
- **Ricerca del punto di riferimento:** La ricerca dei K punti più vicini inizia dalla radice dell'albero kd. Il punto di riferimento viene confrontato con la coordinata di suddivisione del nodo corrente per determinare in quale sottoalbero continuare la ricerca.
- **Attraversamento dell'albero:** L'algoritmo procede attraversando l'albero in modo ricorsivo in base alla posizione del punto cercato rispetto al nodo corrente. Se il punto si trova a sinistra della coordinata di suddivisione, l'algoritmo prosegue nel sottoalbero sinistro, altrimenti, prosegue nel sottoalbero destro.
- **Calcolo della distanza:** Durante l'attraversamento dell'albero, vengono calcolate le distanze tra il punto di riferimento e i punti interni ai nodi visitati. Queste distanze sono calcolate utilizzando una metrica di distanza appropriata, ad esempio la distanza euclidea. Le distanze precedentemente definite vengono utilizzate come raggio di ricerca e permettono di determinare gli altri nodi che contengono almeno un punto che ricade all'interno dello spazio di ricerca (3.4).
- **Mantenimento dei K punti più vicini:** Durante l'attraversamento dell'albero, viene mantenuto un insieme di dimensione K che tiene traccia dei punti più vicini finora trova-

ti. Inizialmente, l'insieme è vuoto. Man mano che l'algoritmo visita i nodi dell'albero, i punti vengono inseriti o rimossi dall'insieme in base alla loro distanza dal punto di riferimento, mantenendo solo i K punti più vicini che, al termine della ricerca, sono restituiti.

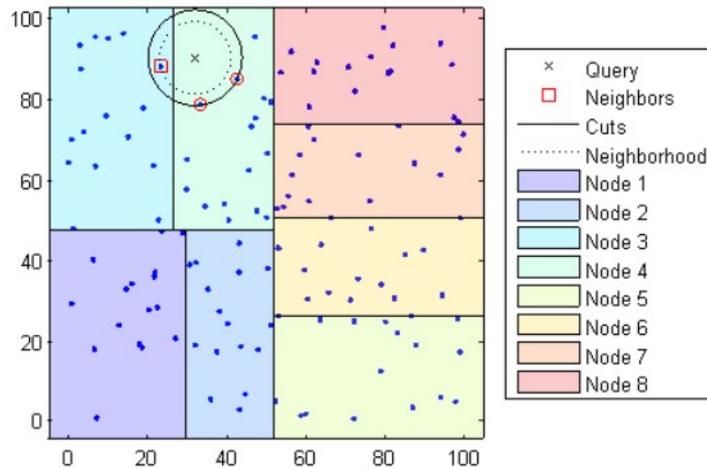


Figura 3.4: Ricerca dei punti più vicini

Definite le distanze dai punti più vicini è necessario andare a determinare una misura univoca per il singolo punto e successivamente una metrica di valutazione della complessità di una point cloud sulla base delle distanze suddette.

La prima proposta è quella di andare a considerare la distanza media da punto più vicino: quindi, per ogni punto si tiene in considerazione unicamente la più piccola delle distanze precedentemente calcolate e il risultato finale è ottenuto mediando su tutti i punti della point cloud. Tale approccio presenta principalmente due criticità: la prima deriva dal fatto che la misura risulta completamente priva di significato nel momento in cui la maggior parte dei punti presenta la stessa distanza minima, situazione che si può verificare nel caso in cui il dataset preso in considerazione sia stato precedentemente voxelizzato. In secondo luogo, è necessario sottolineare che andando a considerare un unico punto non è detto che si abbia una buona misura della densità di una point cloud e, a questo scopo, sarebbe più adatto andare a esaminare un numero di punti maggiore.

In questo caso, la misura di complessità sarebbe definita come la media su tutti i punti della point cloud della distanza media dai K punti più vicini (K fissato). Andando a considerare un numero di punti maggiore, si complica leggermente la complessità dell'algoritmo precedente-

mente descritto, tuttavia ciò permetterebbe di avere una misura più generale della densità del contenuto e potrebbe permettere anche di risolvere il problema precedentemente esposto.

Un'alternativa è quella di andare a considerare la deviazione standard della distanza dal punto più vicino (o da un insieme di punti vicini). La deviazione standard dei punti comporta una distribuzione più irregolare degli stessi e questo implica una maggiore difficoltà di codifica. Per esempio, nel caso del codificatore descritto, il fattore di scala utilizzato è costante per tutti i punti della point cloud, così come il numero di suddivisioni octree e questo comporta che, in presenza di una deviazione standard che si discosta molto dallo zero, i risultati ottenuti non possano avere la stessa qualità in tutto il contenuto.

3.3.2 Curvatura

La curvatura di una point cloud è una misura di quanto il suo contenuto differisce dall'essere localmente planare in ciascun punto che lo costituisce. È possibile, anche in questo caso, ottenere una misura univoca di valutazione della complessità dell'oggetto facendo un'analisi statistica dei risultati ottenuti, come verrà illustrato in seguito.

La scelta di considerare la curvatura come predittore della complessità di una point cloud deriva principalmente da alcune considerazioni fatte per quanto concerne il modello di codifica trisoup. Come descritto nel capitolo dedicato, la codifica trisoup si basa sulla suddivisione ricorsiva dello spazio di definizione ma include anche un'approssimazione planare locale del contenuto: infatti, ciascun voxel viene approssimato con il piano che meglio descrive la disposizione dei punti al suo interno e in fase di decodifica i punti interni a tale volume vengono estratti casualmente dalla superficie precedentemente definita.

Questo comporta che point cloud aventi una curvatura inferiore saranno compresse con dei risultati migliori in quanto, sebbene sia possibile utilizzare la codifica trisoup indipendentemente dalle caratteristiche del contenuto, i punti che vengono ricostruiti in fase di decodifica in questa circostanza risultano essere una buona approssimazione di quelli inizialmente a disposizione. Al contrario, la distorsione introdotta dalla compressione è notevolmente superiore nel caso di contenuti difficilmente approssimabili con delle superfici planari, in quanto i punti decodificati differiranno maggiormente da quelli originali.

Inoltre, è ragionevole considerare la curvatura come misura della complessità di una point cloud, in quanto valori della curvatura molto elevati tipicamente suggeriscono la presenza di una grande variabilità geometrica, che implica la presenza di bordi, angoli o regioni di transi-

zione, mentre a cospetto di valori di curvatura inferiori generalmente si hanno aree più piatte o superfici meno complicate.

Nella letteratura, la curvatura viene utilizzata in diverse misure di valutazione della distorsione di contenuti compressi. In particolare, in [4] viene proposto un metodo per definire la qualità della point cloud compressa in riferimento a quella originale: per ciascun punto viene calcolata la curvatura sulla base dei punti appartenenti alla sfera di vicinanza e lo stesso viene effettuato nella point cloud di riferimento.

Definita la corrispondenza tra ciascun punto e il punto più vicino appartenente al contenuto non compresso, i valori della media e della varianza ottenuti vengono confrontati con l'obiettivo di determinare una misura univoca di distorsione per il singolo punto. Tali valori di distorsione locale possono, poi, essere combinati in modo da definire una singola misura di distorsione globale.

In maniera analoga in [5] viene proposto un metodo di valutazione della qualità percettiva per 3D meshes che considera come principale fattore la rugosità del contenuto, definita come differenza pesata tra la curvatura gaussiana del punto di riferimento e quella dei suoi vicini. In entrambi i casi risulta che le metriche proposte basate sul calcolo della curvatura siano particolarmente abili nel definire la qualità percettiva del risultato. Infatti, sembrano essere maggiormente correlate con i risultati degli esperimenti soggettivi rispetto alle altre metriche di valutazione della distorsione.

Il fatto che la curvatura sia un ottimo elemento per definire la difficoltà di compressione di un contenuto fa in modo che diventi anche un buon candidato per predire la complessità geometrica di una point cloud: infatti, se il contenuto è caratterizzato da una superficie molto irregolare e quindi presenta valori di curvatura molto elevati, è probabile che le perdite introdotte con la distorsione siano elevate in quanto la superficie non riesca ad essere approssimata bene.

Giustificata la scelta di valutare la curvatura di una point cloud come misura della sua complessità, viene in seguito fornita una definizione per tale metrica e proposto un procedimento operativo per il suo calcolo.

Come accennato in precedenza, la curvatura descrive per ciascuna parte di una superficie quanto questa si discosti dall'essere un piano. Di conseguenza, la principale difficoltà che si riscontra nell'andare a definire la curvatura di una point cloud è il fatto che questa si presenta come un insieme di punti indipendenti tra loro e non è definita una superficie da valutare. La soluzione è quella di approssimare localmente l'insieme dei punti con una superficie e questo vuol dire che, fissato il raggio di vicinanza, i punti appartenenti alla sfera di tale raggio vengono

descritti attraverso la superficie di ordine prestabilito, per esempio una superficie quadratica, che minimizzi l'errore di approssimazione.

A questo punto, esistono diversi approcci per valutare la curvatura di una superficie ma tutti si basano sul calcolo delle curvature principali (Figura 3.5). Dato un punto appartenente ad una superficie e noto il vettore normale alla superficie in questo punto, è possibile andare a definire numerosi piani che contengano il vettore normale e che quindi siano perpendicolari alla superficie. Ognuno di essi la interseca definendo una curva caratterizzata da una certa curvatura: le curvature principali rappresentano il massimo e il minimo tra questi valori.

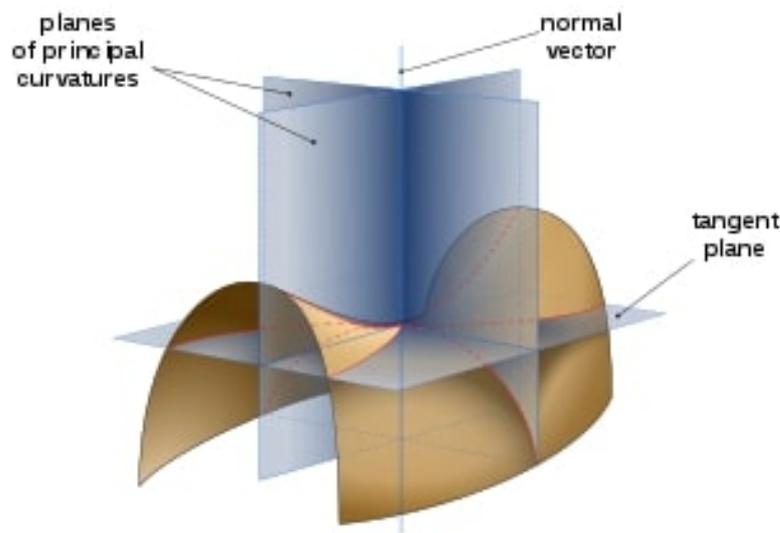


Figura 3.5: Curvature principali di una superficie

In maniera più formale, è possibile ottenere una stima delle curvature principali andando a calcolare la matrice Hessiana dei vettori normali alla superficie nelle vicinanze del punto di interesse. Calcolati gli autovalori di tale matrice, le curvature principali risultano essere il massimo e il minimo tra questi valori.

La curvatura media viene definita come la media tra le curvature principali. Quindi, dette k_{min} e k_{max} le curvature principali, la curvatura media può essere calcolata come segue:

$$m_c = \frac{k_{min} + k_{max}}{2}. \quad (3.2)$$

La curvatura media fornisce informazioni per quanto riguarda l'andamento generale della superficie nelle vicinanze del punto d'interesse: valori positivi implicano che la superficie sia prevalentemente convessa, valori negativi corrispondono a superfici perlopiù concave e infine una curvatura media approssimativamente nulla si ha in presenza di regioni piane.

In modo analogo, possiamo definire la curvatura gaussiana come il prodotto delle due curvature principali, ossia:

$$g_c = k_{min} \cdot k_{max} \quad (3.3)$$

dove k_{min} e k_{max} assumono lo stesso significato dell'equazione precedente.

La curvatura gaussiana definisce intrinsecamente la curvatura della superficie di interesse: valori positivi corrispondono ad una superficie localmente sferica, mentre valori negativi fanno riferimento ad una superficie localmente iperbolica. Come nel caso precedente, valori vicini allo zero si riferiscono a regioni perlopiù piatte.

In questa tesi è stata effettuata la scelta di considerare la curvatura gaussiana come misura della rugosità del contenuto e in particolare, al fine di determinare tale misura, è stato utilizzato il software CloudCompare che offre diverse funzionalità per l'elaborazione e l'analisi di contenuti tridimensionali.

La procedura utilizzata, descritta in [2], non differisce da quella precedentemente illustrata, andando a effettuare il fitting di una superficie quadratica per approssimare localmente l'insieme dei punti appartenenti alla sfera di vicinanza. Il numero di punti da considerare dipende dal raggio scelto dall'utente: è necessario un numero di punti superiore a sei per poter determinare la superficie quadratica suddetta, quindi nel caso in cui la sfera di vicinanza non contenga un numero di punti sufficientemente elevato, il calcolo non può essere effettuato.

La scelta del raggio di vicinanza non è affatto banale in quanto un valore troppo ridotto comporta che per una grande quantità di punti la metrica suddetta non possa essere calcolata a causa dell'impossibilità di andare a determinare la superficie quadratica e, di conseguenza, la misura complessiva della rugosità del contenuto viene compromessa poiché la statistica è fatta su un numero di punti troppo ridotto. Dall'altra parte, andando a considerare un raggio troppo elevato si includono un numero eccessivo di punti nella stima della superficie quadratica, rendendola estremamente rumorosa e poco adatta a descrivere l'andamento locale della superficie nelle vicinanze del punto di riferimento.

Un'ulteriore difficoltà si presenta nel tentativo di determinare un raggio univoco che permetta di confrontare la curvatura dei diversi contenuti. Infatti, il confronto dovrebbe essere effettuato facendo in modo che il numero di punti considerati per il fitting sia approssimativamente lo stesso e non sia né troppo ridotto, non permettendo di determinarne il valore, né troppo elevato, inficiando la misura. Tuttavia questo è fortemente influenzato dalla densità dei punti, che varia a seconda del contenuto, rendendo difficile definire una misura univoca per il raggio.

A titolo esemplificativo, sono riportati in Tabella 3.5 le percentuali di punti per cui è stato

possibile effettuare la misura considerando un raggio di 2.225. Come si può notare tale valore permette di calcolare la curvatura per un numero sufficientemente elevato di punti nel caso di thaidancer, mentre per le altre point cloud la percentuale di punti validi è troppo ridotta per poter considerare la misura significativa.

Point Cloud	Punti Validi
Boxer	2.568%
Longdress	1.52%
Loot	2.126%
Redandblack	1.314%
Soldier	1.355%
Thaidancer	96.42%

Tabella 3.5: Percentuale di punti validi per il calcolo della curvatura con raggio 2.225

Lo stesso software fornisce una misura alternativa per la valutazione della rugosità della superficie che viene determinata sulla base della distanza del punto di riferimento dal piano che meglio approssima l'insieme dei punti appartenenti alla sfera di vicinanza, definita come sopra.

Tale misura, che può essere effettuata solo se la sfera di vicinanza contiene almeno 3 punti, risulta andare a descrivere quanto la superficie sia localmente planare. Infatti, una superficie con curvatura approssimativamente nulla potrà essere descritta da un piano in maniera molto più precisa di quanto non possa esserlo una superficie con una curvatura molto elevata, per cui la distorsione introdotta con l'approssimazione è notevolmente superiore. Tuttavia, questo tipo di misura, sebbene sia computazionalmente più semplice da calcolare, riesce a rilevare in maniera più marginale la complessità della superficie, quindi i risultati attesi in questa circostanza sono peggiori.

Al fine di determinare una misura complessiva di valutazione della rugosità del contenuto, è possibile considerare la media dei valori ottenuti per i singoli punti oppure la deviazione standard degli stessi risultati. Entrambe le misure risultano delle valide proposte per valutare la complessità del contenuto, anche se descrivono aspetti leggermente diversi. In presenza di un'elevata deviazione standard si ha che l'andamento della superficie è estremamente irregolare in quanto il valore della curvatura si modifica molto spostandosi all'interno del contenuto, cosa che invece non è assicurata per una point cloud con una curvatura media elevata.

3.3.3 Densità

Data una point cloud voxelizzata con L bit, definiamo il LoD $k < L$ come il risultato del sottocampionamento di un fattore 2^{L-k} ; questo significa che corrisponde al contenuto ottenuto interrompendo la suddivisione octree a profondità k . Per ciascun LoD definiamo:

- n_k il numero di punti
- v_k il numero di voxel validi, intesi come il numero di voxel che potrebbero essere pieni al livello k , ossia unicamente i figli di voxel pieni al livello $k - 1$. Ne segue che $v_1 = 8$ mentre $v_k = 8 \cdot n_{k-1}$
- V_k il numero totale di voxel. Ne segue che $V_k = 2^{3k}$

Fatte queste precisazioni, è possibile introdurre il concetto di densità locale, che può essere intesa come la probabilità di riempimento al livello k dato il numero di punti al livello $k - 1$ ed è definita come segue:

$$d_k = \frac{n_k}{v_k}. \quad (3.4)$$

Analogamente, la densità globale è definita come la probabilità di riempimento del livello k e può essere calcolata come segue:

$$d_k = \frac{n_k}{V_k}. \quad (3.5)$$

È possibile definire anche una relazione tra le due grandezze appena introdotte, infatti:

$$D_k = \frac{n_k}{V_k} = \prod_{i=1}^k \frac{n_i}{v_k} = \prod_{i=1}^k d_i \quad (3.6)$$

Illustrate queste grandezze, è possibile effettuare diverse considerazioni legate alla complessità del contenuto di riferimento. Innanzitutto, è doveroso notare che per k sufficientemente elevato, la densità locale è strettamente legata alla distanza media tra i punti. Infatti, se i punti sono molto vicini tra loro si avrà che i voxel validi al livello $k \gg 1$ saranno quasi tutti pieni. Quindi, le considerazioni che si possono fare per quanto riguarda la densità locale sono molto simili a quelle che sono già state fatte nel caso della distanza media.

Lo stesso, invece, non si può dire nel caso della densità globale in quanto questa non tiene in considerazione soltanto ciò che riguarda il livello k , ma anche tutti quelli precedenti. Di conseguenza, se la distribuzione dei punti è tale che il numero di voxel pieni nei primi LoD

è molto piccolo, questo influenzerà in maniera considerevole la densità globale del contenuto portando ad avere valori molto piccoli anche nel caso di point cloud caratterizzate da una distanza media ridotta.

Al contrario, ci si aspetta che in presenza di valori di densità globale ridotta per k sufficientemente piccolo si abbiano dei contenuti non complessi da comprimere in quanto, dovendo codificare con un byte ogni nodo pieno incontrato, in questa circostanza si abbatte la mole di dati da trasmettere a parità di qualità raggiunta.

Al fine di ottenere le misure di densità locale e globale descritte è stato usato, anche in questo caso, il software CloudCompare, che permette di ricreare la struttura octree e fornisce le statistiche relative al numero di nodi riempiti a ciascun livello. Stando così le cose, le misure di complessità proposte consistono nella densità locale e globale media, ottenute considerando un numero prefissato di livelli di dettaglio.

Capitolo 4

Risultati sperimentali

In questo capitolo verranno presentati i risultati ottenuti per il dataset di riferimento. Il capitolo è strutturato come segue: per ciascuna delle misure precedentemente illustrate viene qui riportata la correlazione con la misura di valutazione della complessità di compressione e, se ritenuto opportuno, il grafico dei punti in quanto, trattandosi di un dataset molto limitato, è possibile che la correlazione non descriva perfettamente la bontà dei risultati ottenuti il cui andamento risulta più facile da valutare visivamente.

Per quanto riguarda la misura di complessità considerata, è stato precedentemente dimostrato che la scelta è quasi completamente irrilevante nel caso del dataset considerato e, di conseguenza, è possibile procedere con una qualsiasi delle misure proposte. Quindi, è stata presa la decisione di considerare l'integrale effettuato sul supporto comune tra le varie curve in modo da avere la garanzia di non integrare fuori dalla regione di definizione.

4.1 Distanza dai punti più vicini

Nella Tabella 4.1 sono presentati i risultati relativi alle misure di predizione della complessità geometrica del contenuto basate sulla distanza tra i punti.

	Distanza media (1 punto)	Distanza media (5 punti)	Distanza media (10 punti)	Distanza STD (1 punto)	Distanza STD (5 punti)	Distanza STD (10 punti)
Octree	0.9974	0.9886	0.9880	0.9769	0.9759	0.9581
Trisoup	0.8129	0.7668	0.7647	0.7275	0.8236	0.9092

Tabella 4.1: Correlazione tra le misure di complessità legate alla distanza tra i punti e la misura di difficoltà di compressione

Si può notare che tutte queste misure funzionano estremamente bene nel predire la complessità nel caso in cui la codifica utilizzata sia octree, mentre i risultati ottenuti per la codifica trisoup, seppure siano comunque buoni, risultano leggermente peggiori. Questo era prevedibile in quanto, come illustrato nella descrizione della metrica proposta, essa si adatta perfettamente alla codifica octree. Infatti, nel caso di point cloud molto dense, è possibile ridurre maggiormente la quantità di informazione in fase di quantizzazione, senza alterare notevolmente la qualità del contenuto. Questo sicuramente influisce in parte anche nelle prestazioni della codifica trisoup, essendo quest'ultima caratterizzata da un certo numero di suddivisioni octree, ma la qualità raggiunta del contenuto compresso è prevalentemente influenzata dall'approssimazione planare eseguita.

In generale, si può notare che all'aumentare del numero di vicini considerati per il singolo punto diminuisce progressivamente la correlazione con la misura di complessità di compressione. Questo comporta che, sulla base dei risultati ottenuti, sembrerebbe più adatto considerare un numero di punti ridotto, anche se, come precedentemente sottolineato, questa scelta potrebbe risultare molto meno consona in alcune circostanze estranee al dataset preso in considerazione.

L'unica eccezione alle osservazioni appena fatte è la deviazione standard della distanza media tra i punti vicini nel caso in cui la codifica utilizzata sia trisoup. Come si può vedere, in questa circostanza la correlazione aumenta all'aumentare del numero di vicini considerati ed è stato provato che esaminando un numero di punti ancora più grande i risultati sono sempre migliori (0.9430 nel caso in cui si considerino 20 vicini, 0.9636 se i punti valutati sono 30). Tuttavia, risulta difficile avanzare delle ipotesi che spieghino questa singolarità basandosi unicamente sui risultati raccolti sul dataset corrente.

Infine, è interessante notare che, a primo impatto, sia la deviazione standard che la media risultano essere delle ottime alternative per determinare una misura univoca di valutazione dell'intera point cloud. In entrambi i casi, i risultati ottenuti sono molto buoni e paragonabili tra loro: la media ha delle prestazioni leggermente migliori nel caso della codifica octree mentre, per quanto riguarda trisoup, i risultati sono altalenanti a causa della singolarità precedentemente illustrata.

Tuttavia, osservando la disposizione dei punti in Figura 4.1, si può notare che la distanza media dal punto più vicino risulta estremamente più adatta a descrivere la complessità di una point cloud in quanto l'andamento è molto più lineare.

Questo è testimoniato anche dal fatto che, se si esclude dalla statistica thaidancer, che è l'unica point cloud del dataset avente delle caratteristiche leggermente diverse, i risultati

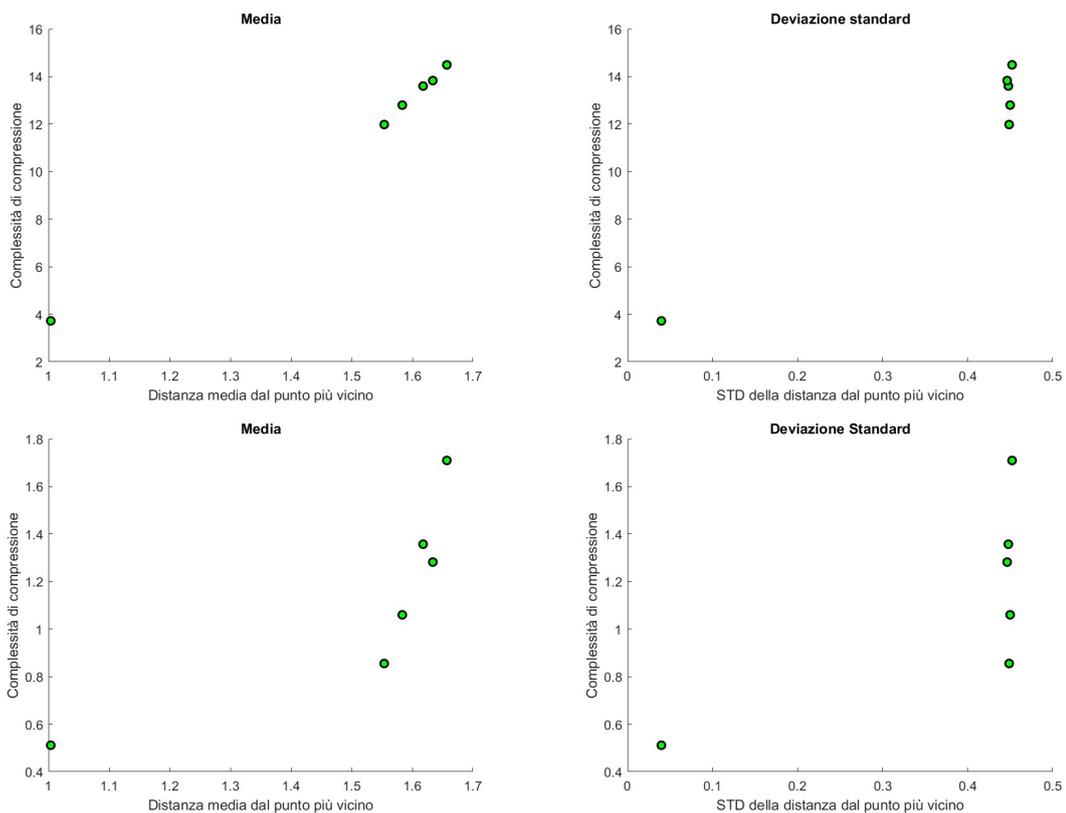


Figura 4.1: Scattering dei punti relativi alla media (a sinistra) e alla deviazione standard (a destra) della distanza dal punto più vicino per codifica octree (sopra) e trisoup (sotto)

relativi alla deviazione standard peggiorano drasticamente (Tabella 4.2), evidenziando la non ottimalità della misura. Al contrario, per quanto riguarda la distanza media, i risultati sono concordi con quelli dell'intero dataset.

	Distanza media	Distanza STD
Octree	0.9975	0.2307
Trisoup	0.9559	0.4206

Tabella 4.2: Correlazione tra la distanza dal punto più vicino e la misura di difficoltà di compressione, esclusa thaidancer

Tuttavia, andando a considerare un numero di punti molto superiore l'andamento della deviazione standard migliora notevolmente, come mostra la Figura 4.2.

Quindi, si può concludere che la deviazione standard della distanza non risulta una buona alternativa nel caso in cui venissero considerati un numero molto ridotto di vicini mentre, in caso contrario, i risultati ottenuti sono notevolmente migliori. Invece, per quanto riguarda la media, si rileva l'andamento opposto, per cui la bontà della predizione decresce all'aumentare del numero di punti considerati.

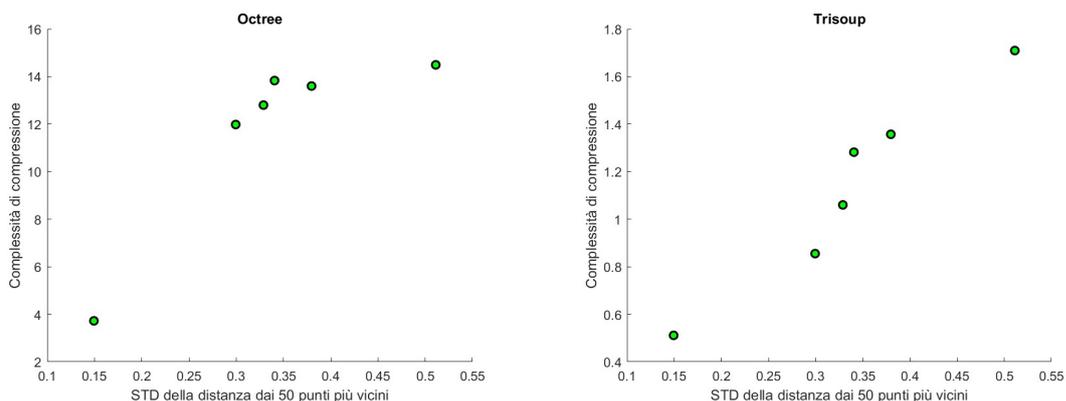


Figura 4.2: Scattering dei punti relativi alla deviazione standard della distanza media dai 50 punti più vicini per codifica octree (destra) e trisoup (sinistra)

4.2 Curvatura

Nella Tabella 4.3 sono presentati i risultati relativi alle misure di predizione della complessità geometrica del contenuto basate sulla rugosità della superficie.

	Curvatura media (raggio 3)	Curvatura STD (raggio 3)	Rugosità (raggio 3)	Curvatura media (raggio 4.5)	Curvatura STD (raggio 4.5)	Rugosità (raggio 4.5)
Octree	0.9888	0.8529	-0.0300	0.4404	0.2547	-0.5828
Trisoup	0.8931	0.8983	0.4037	0.8359	0.6962	-0.0949

Tabella 4.3: Correlazione tra le misure di complessità legate alla rugosità del contenuto e la misura di difficoltà di compressione

È interessante osservare che i risultati ottenuti relativamente alle misure legate alla curvatura gaussiana della superficie variano molto a seconda del raggio di vicinanza considerato. Infatti, la correlazione più che raddoppia passando da un raggio di 4.5 ad un raggio di 3: questo, come evidenziato nella descrizione della procedura di calcolo, è determinato dal fatto che l'inclusione di un numero eccessivo di punti nella definizione della superficie quadratica porta ad una stima estremamente rumorosa, ottenendo una superficie che approssima male l'andamento locale dei punti.

Escludendo thaidancer dalla statistica, si ottengono risultati notevolmente migliori (Tabella 4.4) a testimonianza del fatto che per point cloud con una distanza media tra i punti più elevata, la scelta di un raggio pari a 4.5 non compromette la bontà dell'approssimazione, in quanto il numero di punti considerati rimane ridotto.

Quindi, la decisione finale relativamente al raggio di vicinanza da utilizzare è ricaduta su un raggio pari a 3, in quanto, per il dataset in considerazione, risulta essere un buon compromesso tra la bontà dell'approssimazione della superficie e la possibilità di definire tale superficie per

	Curvatura media (raggio 4.5)	Curvatura STD (raggio 4.5)
Octree	0.9396	0.8453
Trisoup	0.9809	0.9340

Tabella 4.4: Correlazione della curvatura calcolata con raggio 4.5, escluso thaidancer

un numero sufficientemente elevato di punti: infatti, facendo questa scelta si ha che per ciascuna point cloud vengono considerati almeno il 90% dei punti (più precisamente la percentuale di validità minima è 92,6%).

Nondimeno, anche in questa circostanza la correlazione ottenuta escludendo thaidancer è leggermente migliore, nonostante la differenza sia molto meno consistente del caso precedente. Questo può essere determinato dal fatto che il raggio rimane troppo elevato per fornire una buona approssimazione della superficie, se si prendono in considerazione contenuti molto densi. Tale problema può essere risolto unicamente modificando l'approccio di calcolo della curvatura gaussiana: fissando il numero di punti appartenenti alla sfera di vicinanza, piuttosto che il raggio si potrebbe ottenere un'approssimazione la cui precisione sarebbe indipendente dalla densità del contenuto.

È doveroso notare anche che i risultati ottenuti risultano maggiormente correlati con la complessità di compressione relativa alla codifica octree. Questo è sorprendente per quanto sottolineato nella descrizione della metrica proposta, ovvero il fatto che è ragionevole supporre che un contenuto con una curvatura ridotta codificato con trisoup mantenga una qualità elevata, in quanto l'approssimazione planare effettuata comporta l'introduzione di una distorsione inferiore.

Tuttavia, questa difformità potrebbe essere determinata dal problema di definizione del raggio, precedentemente esposto. Infatti, per entrambi i valori valutati, la correlazione ottenuta nel caso di codifica trisoup supera quella che si ha per la codifica octree nel caso in cui non venga considerata la point cloud più densa, come si evince confrontando le Tabelle 4.3 e 4.4.

Inoltre, dai risultati riportati nella Tabella 4.3, risulta evidente che la misura della rugosità disponibile su CloudCompare non è affatto adatta a descrivere la complessità geometrica di una point cloud. Infatti, sebbene fosse prevedibile che la semplicità della soluzione proposta la rendesse meno efficace, i risultati ottenuti mostrano una completa incapacità nel predire la complessità del contenuto, tanto da ottenere in alcuni casi una correlazione negativa che, se fosse attendibile, implicherebbe una maggiore difficoltà di compressione per point cloud aventi una superficie più piatta.

Sicuramente i risultati sono influenzati, anche in questo caso, dalla scelta del raggio di vicinanza ma, sebbene ci sia un leggero miglioramento non considerando thaidancer, non è tale da rendere la misura una valida alternativa per determinare la complessità. Infatti, osservando la Figura 4.3, si può notare come l'andamento dei punti non rispecchi minimamente la complessità di compressione.

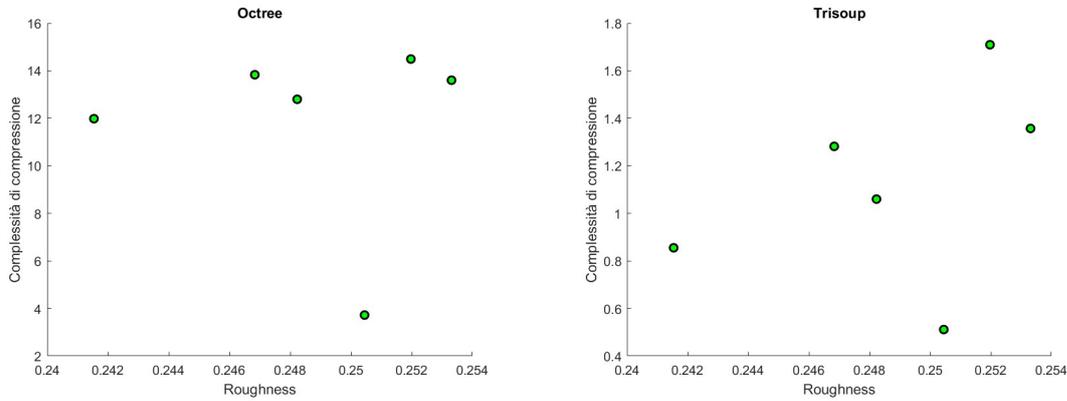


Figura 4.3: Scattering dei punti relativi alla rugosità per codifica octree (sinistra) e trisoup (destra)

Infine, alcune considerazioni possono essere effettuate riguardo alla scelta tra la media e la deviazione standard. Anche in questo caso, come osservato per la distanza tra i punti, a primo impatto i risultati ottenuti sono paragonabili, quindi è necessario osservare la distribuzione dei punti per valutare se una delle due misure sia preferibile rispetto all'altra.

Dalla Figura 4.4, si deduce che, così come per la distanza dal punto più vicino, la distribuzione relativa alla curvatura media è più lineare, tuttavia le differenze non sono così nette da portare a invalidare la misura che calcola la deviazione standard della curvatura.

Nel complesso, possiamo concludere che le misure legate alla curvatura risultano essere delle valide alternative per descrivere la complessità geometrica del contenuto. Tuttavia, per poter valutare la bontà di tali proposte è necessario andare a definire una nuova procedura di calcolo, che permetta di esaminare tutti i contenuti a disposizione nelle stesse condizioni.

4.3 Densità

Nella Tabella 4.5 sono presentati i risultati relativi alle misure di predizione della complessità geometrica del contenuto basate sulla probabilità di occupazione dei voxel. Nella definizione delle misure è stata usata la stessa simbologia introdotta nella descrizione della metrica, per cui le misure indicate con D fanno riferimento alla densità globale del contenuto, mentre nei casi in cui viene utilizzato d , si rimanda alla densità locale. Il risultato finale è ottenuto mediando

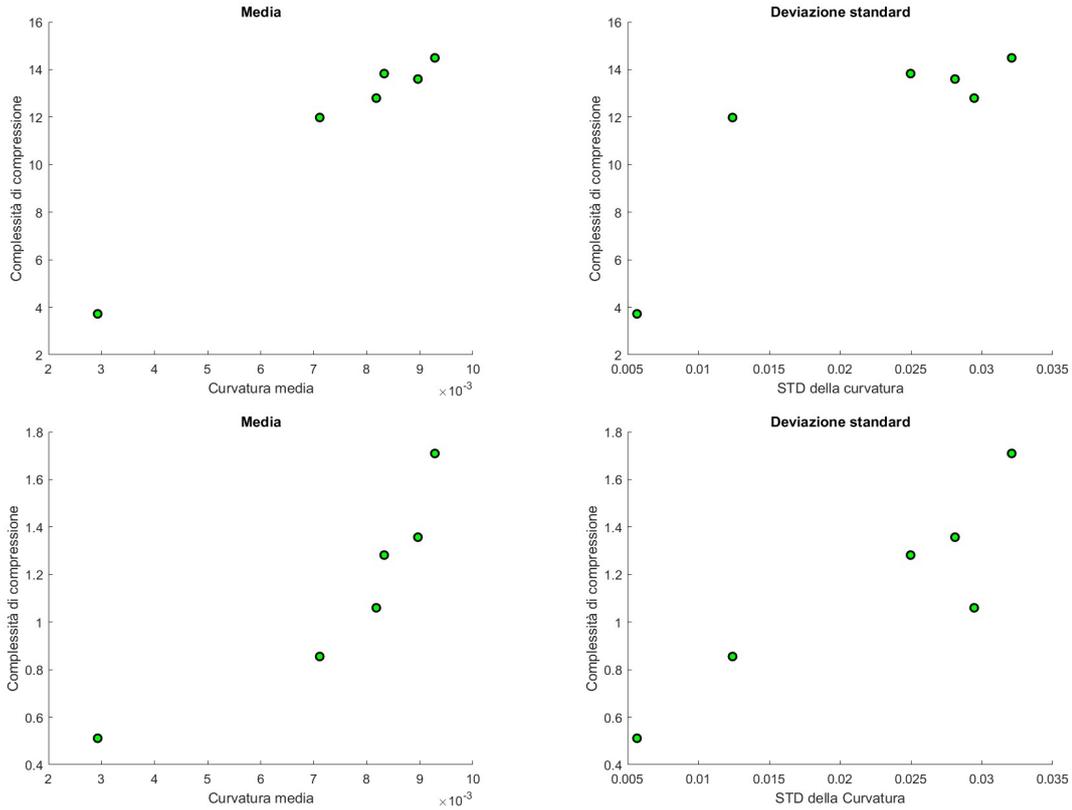


Figura 4.4: Scattering dei punti relativi alla media (a sinistra) e alla deviazione standard (a destra) della curvatura per codifica octree (sopra) e trisoup (sotto)

i valori della probabilità ottenuti per l'intervallo di livelli indicato a pedice. È stato considerato come massimo livello di suddivisione quello a profondità 12, essendo tutti i contenuti del dataset a disposizione inizialmente voxelizzati a questo stesso livello.

	$D_{1...12}$	$D_{1...5}$	$D_{1...3}$	$d_{1...12}$	$d_{9...12}$	$d_{11...12}$
Octree	0.0242	0.0291	0.0043	-0.2054	-0.5704	-0.1335
Trisoup	-0.2475	-0.2508	-0.3027	-0.2974	-0.8596	-0.5995

Tabella 4.5: Correlazione tra le misure di complessità legate alla densità del contenuto e la misura di difficoltà di compressione

In generale, dai valori di correlazione ottenuti si riscontra una pessima abilità nel predire la complessità geometrica del contenuto. Tuttavia, ai fini di giustificare tale risultato è necessario osservare il grafico dei punti, con l'obiettivo di acquisire maggiori informazioni relative all'andamento della complessità in funzione della probabilità media di occupazione dei voxel. Per non appesantire la trattazione, nel seguito vengono riportati unicamente i grafici relativi alla codifica octree, in quanto la misura proposta dovrebbe risultare più significativa in questa circostanza.

Per quanto riguarda le misure relative alla densità globale del contenuto (Figura 4.5), l'ipotesi fatta in fase di presentazione della misura sembrerebbe essere smentita. Infatti, la correla-

zione ottenuta è estremamente bassa e dai grafici si rileva che, anche se consideriamo unicamente le prime divisioni octree, tendenzialmente la complessità diminuisce per contenuti con una densità globale superiore, contrariamente a quanto previsto. Infatti, l'ipotesi avanzata affermava che una bassa probabilità di occupazione nelle prime suddivisioni avrebbe dovuto garantire una complessità inferiore, andando ad abbattere il numero di byte da trasmettere.

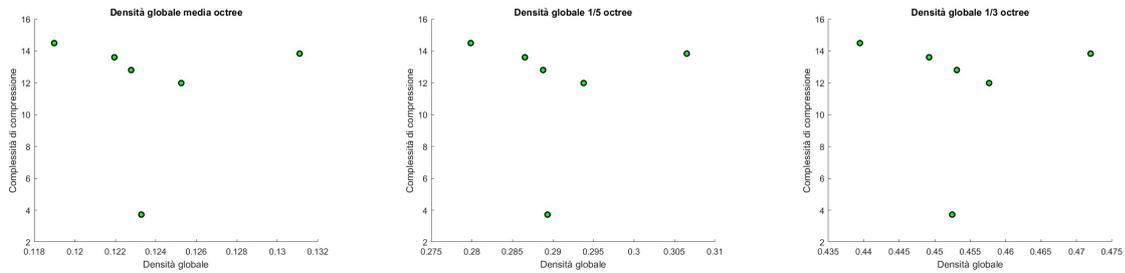


Figura 4.5: Scattering dei punti relativi alla densità globale dei contenuti: (1) considera la densità calcolata su tutti i livelli, (2) valuta unicamente i primi 5 livelli e (3) è relativa ai primi 3 livelli

Coerentemente con quanto ci si aspettava, la misura ottenuta mediando su tutte le suddivisioni octree risulta essere quella meno informativa. Infatti, la probabilità di occupazione dei primi livelli è influenzata da aspetti differenti rispetto alla stessa probabilità valutata ad una risoluzione maggiore e, di conseguenza, mediando su tutti i livelli si ottiene un risultato poco significativo. Ciononostante, anche le altre misure legate alla densità globale non sembrano essere dei buoni candidati a predire la qualità del contenuto, in quanto la disposizione dei punti non rispecchia l'andamento della complessità di compressione.

In Figura 4.6, vengono riportati i grafici relativi alle misure di densità locale. Innanzitutto dobbiamo sottolineare che la correlazione è negativa e questo è concorde con l'ipotesi fatta, in quanto alte probabilità si riferiscono a point cloud molto dense che, quindi, presentano una complessità di compressione limitata.

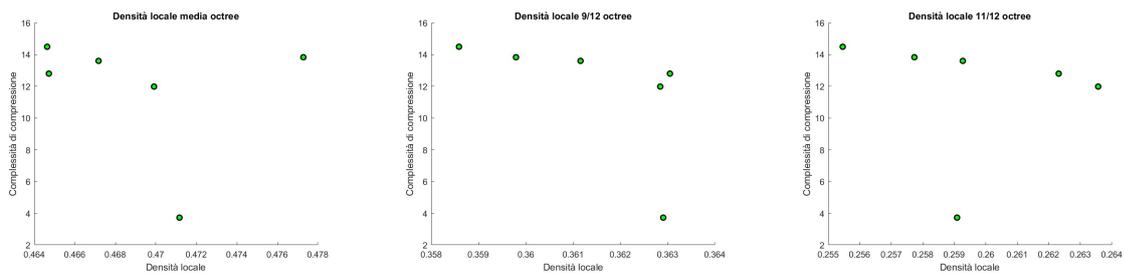


Figura 4.6: Scattering dei punti relativi alla densità locale dei contenuti: (1) considera la densità calcolata su tutti i livelli, (2) valuta unicamente gli ultimi 4 livelli e (3) è relativa agli ultimi 2 livelli

In secondo luogo, è evidente che, così come per la densità globale, coinvolgendo un numero eccessivo di risultati nella statistica, si ottiene una metrica meno significativa. Invece, se il numero di suddivisioni considerate è limitato, si può notare che l'andamento dei punti è abbastanza lineare, con un'unica eccezione rappresentata dalla misura relativa a thaidancer. Questo può essere giustificato dal fatto che, come sottolineato nella descrizione del dataset, la point cloud suddetta è definita approssimativamente su 1/8 dello spazio occupato dal resto del dataset, di conseguenza la risoluzione ottenuta a parità di livello di suddivisione non è la stessa, in quanto le celle hanno una dimensione differente.

Il problema può essere risolto valutando la probabilità di occupazione a parità di dimensione delle celle, senza fissare il livello di suddivisione. Per il dataset considerato tale obiettivo può essere agevolmente raggiunto in quanto, 5 delle point cloud che vi appartengono sono definite sullo stesso spazio mentre per quanto concerne il contenuto rimanente, viene raggiunta la dimensione delle celle desiderata con un livello di anticipo. Infatti, facendo tale modifica alla misura proposta i risultati migliorano notevolmente, come riportato nella Tabella 4.6.

	$d_{9...12}$	$d_{11...12}$
Octree	-0.9844	-0.9849
Trisoup	-0.7516	-0.7529

Tabella 4.6: Correlazione tra le misure di complessità legate alla probabilità di occupazione fissata la dimensione della cella e la misura di difficoltà di compressione

A questo punto, sono doverose le seguenti osservazioni: innanzitutto, è possibile notare che, coerentemente con quanto evidenziato nella descrizione della misura, questa rappresenta una stima della densità del contenuto e ciò è testimoniato dal fatto che la correlazione con la distanza media dal punto più vicino è quasi totale (-0.9947, nel caso in cui si considerino unicamente celle di dimensione 1 e 2). In secondo luogo, bisogna sottolineare che la misura proposta è difficilmente generalizzabile perché non vi è la garanzia di riuscire ad ottenere celle che abbiano sempre la stessa dimensione.

Capitolo 5

Conclusioni

In sintesi, lo scopo di questo lavoro di tesi è quello di determinare delle misure generali e computazionalmente semplici da calcolare per individuare la complessità geometrica di una point cloud.

Lo studio condotto ha permesso di valutare l'affidabilità di diverse misure, che si basano prevalentemente sull'analisi della densità del contenuto e della curvatura della superficie e ciò che si può concludere è che, dai risultati ottenuti, questi aspetti sembrano influenzare notevolmente la qualità raggiunta a seguito della compressione. In particolare, la correlazione con la misura di complessità di compressione utilizzata come banco di prova è, in generale, molto elevata, raggiungendo un valore massimo di 0.9974 nel caso di misure basate sulla densità e di 0.9888 andando a valutare la rugosità del contenuto.

Come sottolineato più volte nel corso della trattazione, la bontà dei risultati è certamente influenzata dalle caratteristiche del dataset utilizzato: il numero limitato di point cloud e l'estrema somiglianza tra di esse, sia per quanto riguarda il tipo di contenuti che per quanto concerne la quantità di punti, fanno in modo che alcuni aspetti non siano facili da valutare. Questo comporta la necessità di provare le stesse misure su un dataset più ampio e vario, con l'obiettivo di stabilire la loro validità in un contesto più generale. Senza questi ulteriori test, non è possibile definire quanto siano generalizzabili i risultati ottenuti poiché non c'è modo di capire se le misure proposte si adattino unicamente alle caratteristiche del dataset in questione.

È necessario anche osservare che l'obiettivo di questo studio più difficile da realizzare è la definizione di misure indipendenti dal metodo di codifica utilizzato, con lo scopo di determinare la complessità intrinseca del contenuto. Ciò che si evince dai risultati ottenuti è che le caratteristiche che influenzano la complessità di compressione sono strettamente legate ai principi di funzionamento del codificatore e questo comporta che i risultati ottenuti differisca-

no in maniera significativa a seconda del metodo utilizzato, rendendo difficile lo sviluppo di una metrica univoca.

5.1 Sviluppi futuri

Oltre alla precedentemente illustrata necessità di valutare le misure proposte su un numero maggiore di contenuti con l'obiettivo di estenderne la validità, lo studio effettuato in questa tesi ha diversi possibili sviluppi futuri. Innanzitutto, potrebbero essere affinate le procedure di calcolo delle misure proposte, al fine di risolvere le problematiche emerse in fase di valutazione dei risultati ottenuti. Nello specifico, sarebbe utile trovare un metodo alternativo per il calcolo della curvatura, che non preveda la definizione del raggio di vicinanza per determinare la superficie quadratica che approssimi localmente i punti, in modo da valutare tutti i contenuti nelle stesse condizioni.

Tuttavia, prima di procedere in questa direzione, è doveroso fare una precisazione: lo sviluppo di procedure di calcolo più complesse potrebbe portare a ottenere risultati migliori in termini di predizione della complessità del contenuto ma, al tempo stesso, aumenterebbe il costo computazionale, rendendo queste metriche non più immediate da calcolare. Quindi, è necessario determinare il giusto compromesso tra la complessità di calcolo e la bontà dell'approssimazione. Infatti, come sottolineato nell'introduzione, nessuna misura permetterà di ottenere una stima migliore della complessità di quanto non sia possibile fare confrontando a posteriori l'oggetto compresso con quello originale. Quindi, il vantaggio derivante dallo sviluppo di tali metriche è quello di ottenere una misura facile e immediata.

In secondo luogo, potrebbe essere interessante andare ad esaminare le misure proposte alla luce dei risultati ottenuti per un numero maggiore di point cloud con l'obiettivo di spiegare alcuni fenomeni per cui non è stato possibile trovare una giustificazione. Per esempio, un numero superiore di dati raccolti potrebbe aiutare a spiegare l'incremento della correlazione corrispondente all'aumento del numero di punti considerati nel caso in cui la misura proposta sia la deviazione standard della distanza dai vicini e a giustificare i suoi ottimi risultati se la codifica utilizzata è trisoup.

Infine, i test potrebbero essere riprodotti considerando dei metodi di codifica differenti. Infatti, essendo le misure proposte fortemente influenzate dalla codifica utilizzata, è necessario provarne il funzionamento per metodi differenti da quelli impiegati. In particolare, sarebbe interessante valutare la loro validità per metodi di codifica learning-based.

Acronimi

RGB *Red-Green-Blue*

LoD *Level of Detail*

GPCC *Geometry based Point Cloud Compression*

VPCC *Video based Point Cloud Compression*

MPEG *Moving Picture Expert Group*

DCM *Direct Coding Mode*

PSNR *Peak Signal-to-Noise Ratio*

MSE *Mean Square Error*

kd *k dimensional*

Bibliografia

- [1] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai. An overview of ongoing point cloud compression standardization activities: video-based (v-pcc) and geometry-based (g-pcc). *APSIPA Transactions on Signal and Information Processing*, 9:e13, 2020. [doi:10.1017/ATSIP.2020.12](https://doi.org/10.1017/ATSIP.2020.12).
- [2] Zvi Har’el. Curvature of curves and surfaces – a parabolic approach. 1995.
- [3] Maja Krivokuća, Philip A. Chou, and Patrick Savill. “8i Voxelized Surface Light Field (8ivslf) Dataset”, 2018. ISO/IEC JTC1/SC29 WG11 (MPEG) input document m42914.
- [4] Gabriel Meynet, Julie Digne, and Guillaume Lavoué. Pc-msdm: A quality metric for 3d point clouds. In *2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX)*, pages 1–3, 2019. [doi:10.1109/QoMEX.2019.8743313](https://doi.org/10.1109/QoMEX.2019.8743313).
- [5] Kai Wang, Fakhri Torkhani, and Annick Montanvert. A fast roughness-based approach to the assessment of 3d mesh visual quality. *Computers & Graphics*, 36:808–818, 11 2012. [doi:10.1016/j.cag.2012.06.004](https://doi.org/10.1016/j.cag.2012.06.004).
- [6] Honghai Yu and Stefan Winkler. Image complexity and spatial information. In *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)*, pages 12–17, 2013. [doi:10.1109/QoMEX.2013.6603194](https://doi.org/10.1109/QoMEX.2013.6603194).

