



UNIVERSITÀ DEGLI STUDI DI CAGLIARI

Facoltà

Corso di Laurea in Ingegneria Elettrica Elettronica

**Metodi numerici per l'integrazione dei vettori
normali nella tecnica del photometric stereo**

Relatore

Prof. Giuseppe Rodriguez

Candidato

Veronica Puggioni
(Matr. 70/87/65218)

Anno Accademico 2017/2018

Indice

| | |
|--|-----------|
| Introduzione | 3 |
| 1 Shape from Shading | 5 |
| 1.0.1 Modello Lambertiano | 6 |
| 1.1 Stereo Vision | 6 |
| 1.2 Photometric Stereo | 7 |
| 1.2.1 Mappa di Riflettanza | 8 |
| 1.3 Photometric Stereo con posizione della luce conosciuta | 9 |
| 2 Integrazione delle Normali | 11 |
| 2.1 Determinazione delle Normali | 11 |
| 2.2 Equazione di base | 12 |
| 2.3 Proiezione Ortografica | 12 |
| 2.4 Proiezione a prospettiva debole | 14 |
| 2.5 Proiezione Prospettica | 15 |
| 2.6 Problema differenziale | 17 |
| 3 Sperimentazione numerica | 21 |
| 3.1 Valutazione dell'errore | 21 |
| 3.1.1 Laplaciano Esatto | 21 |
| 3.1.2 Gradiente Numerico passo h e $2h$ | 22 |
| 3.1.3 Gradiente ricavato dalle immagini | 23 |
| 3.2 Superficie ricostruita con errore sperimentale | 25 |
| Bibliografia | 28 |

Introduzione

La computer vision si occupa dell'analisi di immagini ed è finalizzata a scoprire cosa è presente nella scena. Possiamo distinguere tra Visione Computazionale di basso livello e di alto livello. La prima si occupa di estrarre determinate proprietà fisiche dell'ambiente visibile, come profondità, forma tridimensionale, contorni degli oggetti, mentre quella di alto livello si occupa della estrazione delle proprietà delle forme, di riconoscimento e classificazione di oggetti, e sono di solito applicati a una porzione dell'immagine. Il compito di ricostruire la pura struttura della scena è difficile. Esistono molti metodi per l'acquisizione automatica della forma di un oggetto. Gli oggetti irradiano luce visibile; la fotocamera cattura questa luce, le cui caratteristiche dipendono dalla illuminazione della scena, dalla geometria e dalla riflettanza della superficie che viene analizzata al calcolatore con lo scopo di inferire la struttura 3D degli oggetti. La distinzione fondamentali tra le tecniche ottiche riguarda l'impiego o meno di sorgenti luminose, in particolare distinguiamo metodi *attivi* e *passivi*.

I primi irradiano la scena con radiazioni elettromagnetiche e sebbene abbiano risoluzioni più elevate non sempre sono applicabili, i metodi passivi invece si basano sull'analisi di immagini della scena così com'è, sono più economici e pongono meno vincoli di applicabilità ma hanno una risoluzione più bassa. Tutti i metodi attivi impiegano una o due fotocamere ed una sorgente di luce speciale e rientrano nella classe più ampia dei metodi ad illuminazione strutturata. Tra i metodi ottici passivi riconosciamo Shape from Shading e Photometric stereo, di cui sarà oggetto la tesi. Nel primo capitolo infatti vengono spiegate le tecniche con cui è possibile ricostruire una superficie tridimensionale, in particolare faremo riferimento al Photometric Stereo. Nel secondo capitolo sono affrontati i concetti di integrazione delle normali. Nell'ultimo capitolo vengono confrontati i diversi metodi per ricostruire una superficie.

Ringraziamenti

Ringrazio professor G. Rodriguez, capace di chiarire ogni mio dubbio, per tutta la disponibilità e l'aiuto fornitomi nella stesura di questa tesi, ringrazio inoltre la dott.ssa Anna Concas del Gruppo di Analisi Numerica di Cagliari (CaNA) per il suo importante aiuto nella preparazione della tesi.

Capitolo 1

Shape from Shading

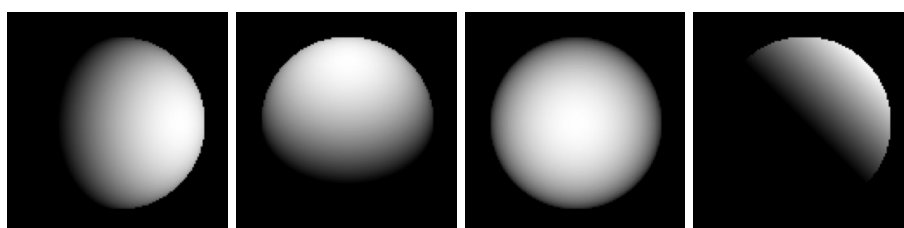


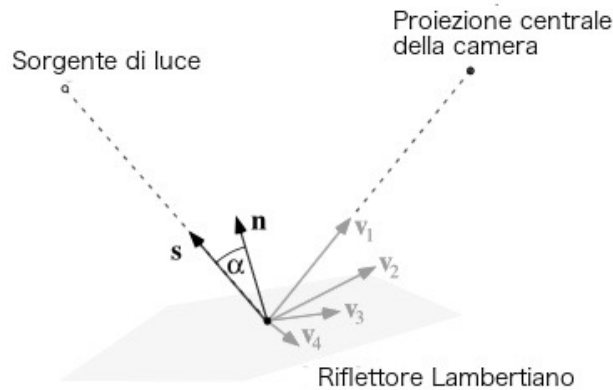
Figura 1.1: Sfere Lambertiane illuminati da diverse direzioni.

Lo *Shape from Shading* è un problema inverso della *Computer Vision*. Dal termine inglese che significa chiaroscuro, è la tecnica che cerca di ricostruire un oggetto tridimensionalmente da una sola immagine, basandosi sulla relazione tra l'intensità di colore registrata nell'immagine e l'orientazione delle normali alla superficie. Lo *Shape from Shading* si è rilevato essere un problema mal posto a causa della mancanza di unicità della soluzione, è infatti impossibile determinare in maniera univoca le normali alla superficie dell'oggetto da una sola immagine senza avere ulteriori informazioni, possono esistere infinite soluzioni che riportano alla stessa sorgente luminosa. Il problema richiede inoltre la conoscenza a priori di parametri, come colore, condizioni di bordo e valori di profondità al contorno, che raramente sono disponibili. Un modo per evitare questo problema è quello di tenere conto non di una sola immagine ma di più immagini, le quali possono essere prese in due modi:

1. Le immagini sono prese da diversi punti di vista sotto le stesse condizioni di luce.
2. Le immagini vengono prese da una stessa posizione, e ciò che cambia è la posizione della luce.

Gli algoritmi sono chiamati rispettivamente *Stereo Vision* e *Photometric Stereo*.

1.0.1 Modello Lambertiano



Uno dei modelli più semplici è il modello proposto intorno alla meta del '700 dal matematico Johann Heinrich Lambert. In una superficie di questo tipo la quantità di luce riflessa rimane costante una volta fissata la direzione della fonte luminosa, infatti essa è sempre uguale al coseno dell'angolo tra la normale alla superficie e la direzione dei raggi di luce.

$$I = \rho \vec{n} \cdot \vec{l} \quad (1.1)$$

in cui la costante ρ rappresenta l'albedo della superficie ed è un fattore di riflettanza tipico del materiale di cui è composta.

1.1 Stereo Vision

La Stereo Vision presuppone l'osservazione di una scena da parte di più punti di vista, molto spesso chiamata infatti *multiview*.

Nel caso in cui si utilizzino solo due telecamere si parla di *Visione Binoculare*. È dimostrato in un lavoro di A. Chambolle, riportato in [3], che il problema con due immagini è ben posto ed ammette soluzione unica. Il principio della visione stereoscopica si basa sul processo di triangolazione, mirato a mettere in relazione la proiezione di un punto della scena (*punti omologhi*) su due o più piani immagine delle telecamere, che compongono il sistema di visione. L'individuazione dei *punti omologhi* consente di ottenere una grandezza denominata *disparità* mediante la quale, attraverso la conoscenza di opportuni parametri del sistema stereoscopico, è possibile risalire alla posizione 3D del punto considerato.

1.2 Photometric Stereo

L'idea del Photometric Stereo è quella di ricostruire la profondità di una superficie, a partire da più immagini bidimensionali di un oggetto, da un solo punto di vista, facendo variare la posizione della sorgente luminosa. Essa si basa sul fatto che la quantità di luce riflessa da una superficie dipende dall'orientazione della luce che la illumina. Ogni frame viene infatti preso in circostanze luminose differenti, abbiamo bisogno inoltre di conoscere le proprietà di riflettanza della superficie, infatti l'ombreggiatura di una superficie dipende sia da come essa è illuminata sia dalle sue proprietà di riflettanza. Woodham [10] fu il primo ad introdurre la stereo fotometria, propose un metodo semplice ed efficiente ma applicabile solo a superfici Lambertiane. Supponendo dunque che la superficie sia tale e che abbia un albedo ρ , conoscendo le direzioni delle sorgenti luminose, sono necessarie due immagini per determinare l'orientazione della superficie in ogni suo punto. Otteniamo pertanto due valori di intensità dell'immagine $I_1(x, y)$ e $I_2(x, y)$ in ciascun punto (x, y) . In generale, i valori dell'immagine di ciascuna fonte di luce corrispondono a due punti sulla mappa di riflettanza (vedi Sezione 1.2.1), come segue:

$$\begin{aligned} I_1(x, y) &= R_1(p, q) \\ I_2(x, y) &= R_2(p, q) \end{aligned} \tag{1.2}$$

In questo modo possiamo determinare i parametri normali della superficie da due immagini. Definendo i due vettori che individuano le sorgenti di luce come $[p_1, q_1, -1]$ e $[p_2, q_2, -1]$ ed equazioni (1.2) come lineari e indipendenti, ci sarà una soluzione unica per p e q :

$$p = \frac{(I_1^2 r_1 - 1)q_2 - (I_2^2 r_2 - 1)q_1}{p_1 q_2 - q_1 p_2} \tag{1.3}$$

$$q = \frac{(I_2^2 r_2 - 1)p_1 - (I_1^2 r_1 - 1)p_2}{p_1 q_2 - q_1 p_2} \tag{1.4}$$

in cui $p_1/q_1 \neq p_2/q_2$; $r_1 = \sqrt{1 + p_1^2 + q_1^2}$ e $r_2 = \sqrt{1 + p_2^2 + q_2^2}$.

Se le equazioni (1.2) non sono lineari, non ci sono soluzioni, o si hanno più soluzioni. Nel caso di superficie Lambertiana dobbiamo introdurre un'altra immagine per eliminare tali ambiguità, questa immagine ci consente di stimare un altro parametro *l'albedo* della superficie, che risulta particolarmente utile nel caso la superficie non sia uniforme.

1.2.1 Mappa di Riflettanza

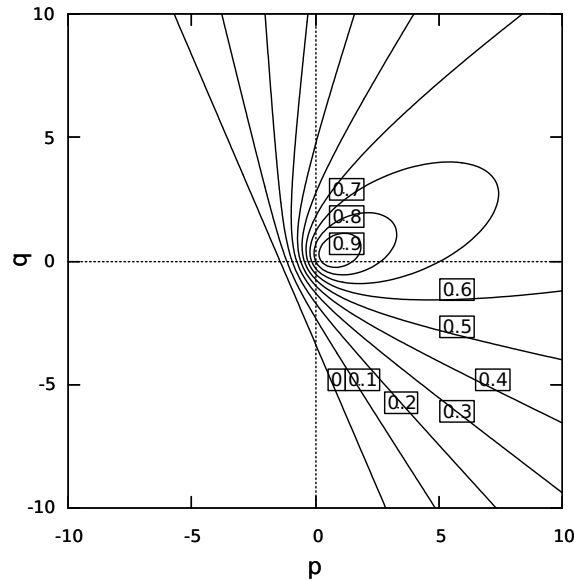


Figura 1.2: Grafico delle curve di livello di una mappa di riflettanza ottenuta con il modello di diffusione Lambertiana

Consideriamo l'immagine di un oggetto posto idealmente su un piano (x, y) , presa con proiezione ortografica, in modo tale da evitare distorsioni prospettiche, da un certo punto di vista con direzione \vec{v} perpendicolare al piano. L'immagine viene spesso rappresentata in maniera discreta da una matrice $(I_{i,j})$, dove i e j sono le coordinate orizzontali e verticali di un pixel e il valore $I_{i,j} \in \{0, \dots, 255\}$ denota l'intensità, o il tono di grigio, del pixel, ma è possibile dare anche un'interpretazione nel continuo introducendo le coordinate immagine (u, v) e pensandola come una funzione $I(u, v)$ a valori nell'intervallo $[0, 1]$. Si può assumere una scala appropriata del piano dell'immagine, tale che le coordinate (x, y, z) dei punti della superficie dell'oggetto vengono mappate sulle coordinate (u, v) dell'immagine in modo che si abbia $u = x$, $v = y$. Se la superficie dell'oggetto è rappresentabile esplicitamente da un'equazione del tipo

$$z = f(x, y)$$

allora la normale esterna alla superficie è data in ogni punto da

$$\vec{n} = \frac{(-p, -q, 1)}{\sqrt{1 + p^2 + q^2}}$$

dove

$$p = \frac{\partial f}{\partial x}, \quad q = \frac{\partial f}{\partial y}$$

Si supponga dunque che l'oggetto impresso nell'immagine sia illuminato da un'unica sorgente luminosa, posta ad una distanza abbastanza grande da poter considerare i raggi di luce incidenti sulla superficie tutti provenienti dalla stessa direzione

$$\vec{\ell} = (\ell_1, \ell_2, \ell_3)$$

Avendo fissato il punto di vista dell'osservatore, la funzione $\phi(\vec{n}, \vec{\ell}, \vec{v})$ che descrive in ogni punto le caratteristiche di riflettanza della superficie dipende solo dalla sua orientazione locale, data dalla normale esterna \vec{n} . In questo caso ϕ è una funzione dello spazio gradiente (p, q) e viene definita una *Mappa di Riflettanza* spesso indicata con la notazione $R(p, q)$. Per ogni punto (x, y) dell'immagine si ha dunque

$$I(x, y) = R(p, q)$$

Nella pratica una mappa di riflettanza può essere determinata empiricamente analizzando le immagini, derivata da modelli fenomenologici della riflettanza della superficie o da modelli analitici delle microstrutture del materiale cui la superficie è composta.

1.3 Photometric Stereo con posizione della luce conosciuta

Consideriamo una superficie Lambertiana, ossia una superficie che rispetta la legge del coseno di Lambert data da:

$$\rho(x, y) \langle n(x, y) \ell_t \rangle = I_t \quad t = 1, 2, \dots, q \quad (1.5)$$

in cui ρ rappresenta l'albedo della superficie nel punto di coordinate (x, y) , che tiene conto dell'assorbimento parziale della luce di quella parte della superficie, $I_t(x, y)$ l'intensità luminosa nello stesso punto, $\langle \cdot, \cdot \rangle$ rappresenta il prodotto scalare in \mathbb{R}^3 , in questa sezione la direzione della luce è conosciuta ed è rappresentata da ℓ_t , $t = 1, \dots, q$. Se l'albedo è costante l'oggetto si dice essere un riflettore lambertiano. Riassumendo dunque le condizioni per un modello di questo tipo sono le seguenti:

- Superficie Lambertiana ;
- La luce posta ad una distanza infinita ;

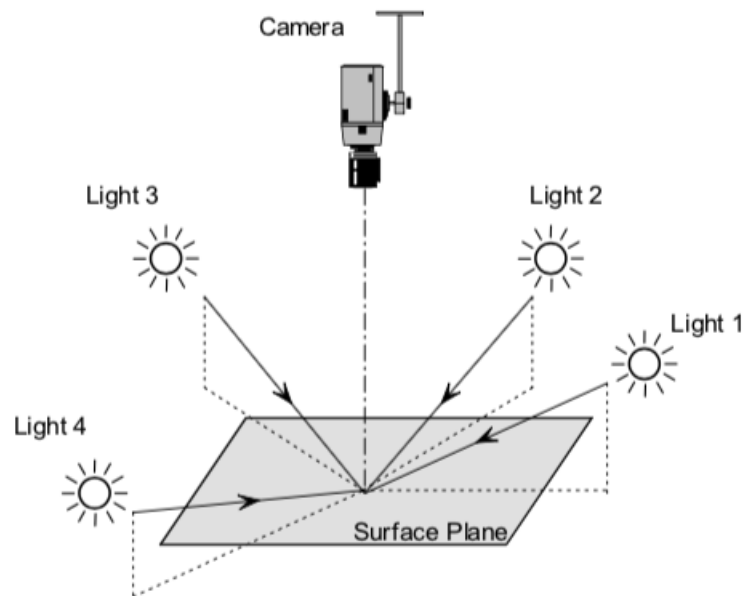


Figura 1.3: Schema grafico della tecnica del Photometric Stereo: diverse immagini di una superficie vengono scattate da uno stesso punto di vista facendo variare la posizione della sorgente luminosa.

- Nessuna parte dell'immagine è ombreggiata ;
- La camera è sufficientemente lontana dall'oggetto in modo tale da non deformarlo ;

Capitolo 2

Integrazione delle Normali

2.1 Determinazione delle Normali

L'approccio di Poisson per la ricostruzione 3D di una superficie Lambertiana segue diverse strategie, il primo passo consiste nel discretizzare la legge di Lambert, risolvendo un'equazione matriciale per determinare il campo normale della superficie. Il campo dei vettori normali che si ottiene è numericamente differenziato, che porta a una discretizzazione approssimativa dell'operatore di Laplace. Infine, risolvendo l'equazione differenziale parziale di Poisson, si ha il profilo 3D dell'oggetto osservato. Il vantaggio di questo metodo è che il calcolo può essere disaccoppiato per problemi più semplici. L'inconveniente è che richiede un numero maggiore di immagini rispetto alla formulazione di Hamilton-Jacobi. Applichiamo la discretizzazione alla (1.5), ignorando i pixel al contorno, in cui le condizioni Dirichlet impostano la soluzione a zero, e successivamente riorganizza i pixel per ordinamento lessicografico. Ponendo n_k il vettore normale alla superficie nell' k -esimo pixel, e con m_{kt} riflessione radiale nella piccola area vicino ai pixel che quando illuminati dalla direzione ℓ , la seguente relazione è valida per ogni punto di ogni immagine:

$$\rho_k \mathbf{n}_k^T \ell_t = m_{kt}, \quad k = 1, \dots, p \quad t = 1, \dots, q \quad (2.1)$$

in cui ρ rappresenta il valore dell'*albedo* per k -esimo pixel. Definendo le matrici

$$D = \text{diag}(\rho_1, \dots, \rho_p) \in \mathbb{R}^{p \times p} \quad N = [n_1, \dots, n_p] \in \mathbb{R}^{3 \times p}$$

$$L = [\ell_1, \dots, \ell_q] \in \mathbb{R}^{3 \times q} \quad M = [m_1, \dots, m_q] \in \mathbb{R}^{p \times q}$$

possiamo dunque riscrivere le equazioni in forma matriciale

$$DN^T L = M \tag{2.2}$$

Quando conosciamo la posizione della luce:

$$\tilde{N}^T = ML^\dagger$$

in cui L^\dagger è la matrice pseudo inversa di L [1]. Quindi le matrici D e N possono essere calcolate dalla scomposizione $ND = \tilde{N}$, normalizzando la colonna di \tilde{N} .

2.2 Equazione di base

Supponiamo di conoscere per ogni punto $\bar{\mathbf{a}} = [u, v]^T$ della superficie, la normale esterna $\mathbf{n}(u, v) = [n_1(u, v), n_2(u, v), n_3(u, v)]^T$. Integrare il campo delle normali \mathbf{n} equivale alla ricerca della funzione $z = (x, y)$ tale che in ogni punto (x, y) del dominio, il vettore normale coincida con quello identificato dal campo.

2.3 Proiezione Ortografica

Fissiamo una camera 3D in un sistema xyz in cui \mathbf{c} è l'origine del centro ottico, e cz coincide con l'asse ottico. (Figura: 2.1)

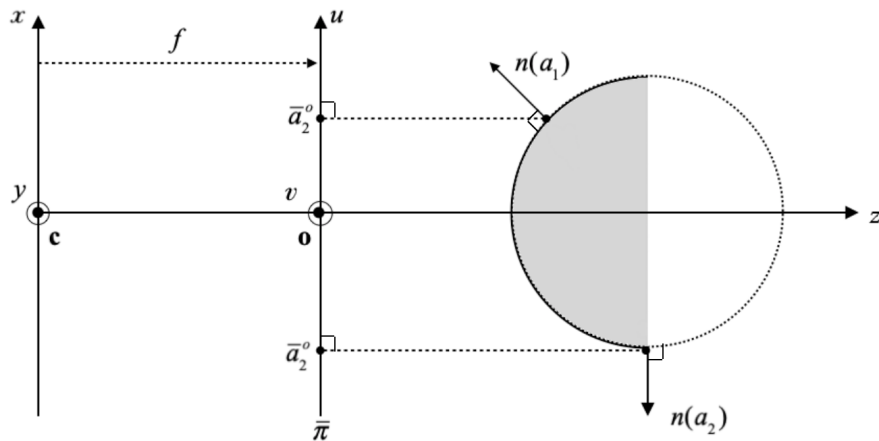


Figura 2.1: Proiezione ortografica, a_1 e a_2 sono coniugati di \bar{a}_1° e \bar{a}_2° rispettivamente. La parte visibile della superficie è evidenziata in grigio. Il punto si trova sul contorno di occlusione.

L'origine delle coordinate del pixel sono date dall'intersezione \mathbf{o} dell'asse ottico \mathbf{cz} e del piano $\bar{\pi}$, che è coincidente con il piano focale $z = f$, dove f rappresenta la lente focale. Supponiamo una proiezione ortografica del punto 3D \mathbf{a}

$$x(u, v) = u, \quad y(u, v) = v \quad (2.3)$$

Dalla normalizzazione del prodotto vettoriale delle derivate parziali $\partial_u \mathbf{a}, \partial_v \mathbf{a}$ e scegliendo il segno di n in modo tale che risulti verso la camera si ottiene :

$$\mathbf{n} = \frac{1}{\sqrt{1 + \|\nabla_z\|^2}} \begin{bmatrix} \partial_u z \\ \partial_v z \\ -1 \end{bmatrix} \quad (2.4)$$

∇_z^2 rappresenta il gradiente della mappa di profondità z . Vediamo che $n_3 < 0$, se $n_3 > 0$ il punto non è visibile, i punti sono visibili e la proiezione ricade nel contorno di occlusione se $n_3 = 0$. Anche se \mathbf{n} è facilmente determinabile nel contorno di occlusione, il calcolo di z non è sempre possibile, poiché \mathbf{n} è sia parallela a $\bar{\pi}$ e sia ortogonale al contorno. Consideriamo ora i punti che non appartengono al contorno di occlusione:

$$\nabla_z = [p, q]^T \quad (2.5)$$

con

$$p = -\frac{n_1}{n_3}, \quad q = -\frac{n_2}{n_3} \quad (2.6)$$

L'equazione (2.5) mostra che integrare il campo delle normali, calcolare z a partire da \mathbf{n} , equivale all'integrazione del campo vettoriale $[p, q]^T$, la cui soluzione è:

$$z(u, v) = z(u_0, v_0) + \int_{(r,s)=(u_0,v_0)}^{(u,v)} [p(r, s)dr + q(r, s)ds] \quad (2.7)$$

L'equazione non dipende dal percorso di integrazione nel momento in cui q e p soddisfano il vincolo di integrabilità $\partial_v p = \partial_u q$ (Teorema di Schwartz). Se non viene soddisfatto, l'Integrale (2.6) dipende dagli estremi di interazione. Se non ci sono punti (u_0, v_0) in cui z sia conosciuto, segue dalla (2.6) che $z(u, v)$ è calcolabile a meno di una costante additiva. La costante può essere scelta in modo da minimizzare l'errore quadratico medio in z .

2.4 Proiezione a prospettiva debole

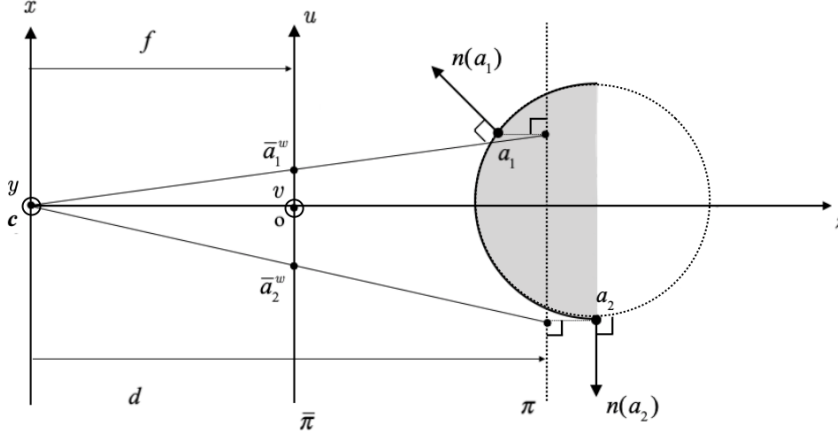


Figura 2.2: Proiezione ortografica weak: a_1 e a_2 sono coniugati di \bar{a}_1^p e \bar{a}_2^p rispettivamente (a_2^w giace nel contorno di occlusione). Il piano π è coniugato di $\bar{\pi}$, che si suppone corrisponda alla media della superficie.

La proiezione a prospettiva debole (Figura: 2.2) presuppone che la camera sia posizionata su un piano π con equazione $z = d$, supponendo che corrisponda alla posizione media della superficie. Ogni punto visibile \mathbf{a} si proietta prima ortogonalmente su π e poi in modo prospettico su $\bar{\pi}$, con \mathbf{c} centro delle proiezioni. Assumendo quindi di avere una "weak-perspective":

$$x(u, v) = \frac{d}{f}u, \quad y(u, v) = \frac{d}{f}v \quad (2.8)$$

denotando con m l'ingrandimento dell'immagine f/d avremo che :

$$\mathbf{n} = \frac{1}{\sqrt{1 + m^2 \|\nabla_z\|^2}} \begin{bmatrix} m \partial u z \\ m \partial v z \\ -1 \end{bmatrix} \quad (2.9)$$

per i punti che non giacciono nel contorno di occlusione la coppia di PDE diventa:

$$\nabla_z = \frac{1}{m} [p, q]^T \quad (2.10)$$

il che spiega perché la "weak-perspective projection" è anche chiamata "proiezione ortografica scalata". Per cui analogamente si ha:

$$z(u, v) = z(u_0, v_0) + \frac{1}{m} \int_{(r,s)=(u_0,v_0)}^{(u,v)} [p(r, s)dr + q(r, s)ds] \quad (2.11)$$

2.5 Proiezione Prospettica

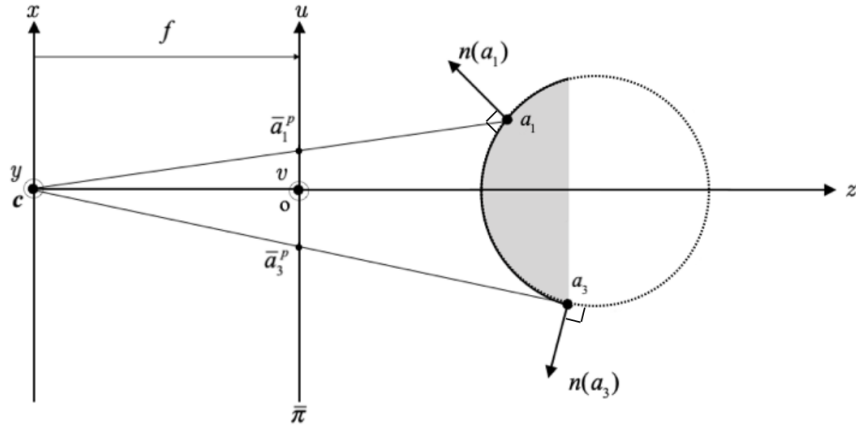


Figura 2.3: Proiezione Prospettica, a_1 e a_3 sono coniugati di \bar{a}_1^p e \bar{a}_3^p rispettivamente (a_3^p giace sul contorno di occlusione).

Nella Proiezione Prospettica (Figura: 2.2) d va sostituito con $z(u, v)$, questa è una delle maggiori differenze con la "weak-perspective":

$$x(u, v) = \frac{z(u, v)}{f}u, \quad y(u, v) = \frac{z(u, v)}{f}v \quad (2.12)$$

Il prodotto vettoriale $\partial_u a \times \partial_v a$ in questo caso risulta più complicato, rispetto al caso precedente:

$$\begin{aligned} \partial_u a \times \partial_v a &= \det \begin{bmatrix} \partial_u x & \partial_u y & \partial_u z \\ \partial_v x & \partial_v y & \partial_v z \\ i & j & k \end{bmatrix} \\ &= \left[\left(\frac{\partial_u z}{f} v \right) (\partial_v z) - \left(\frac{\partial_u z}{f} v + \frac{z}{f} \right) (\partial_v z) \right] i - \left[\left(\frac{\partial_u z}{f} u + \frac{z}{f} \right) (\partial_v z) + \right. \\ &\quad \left. - \left(\frac{\partial_v z}{f} u \right) (\partial_u z) \right] j + \left[\left(\frac{\partial_u z}{f} u + \frac{z}{f} \right) \left(\frac{\partial_v z}{f} v + \frac{z}{f} \right) - \left(\frac{\partial_v z}{f} u \right) \left(\frac{\partial_u z}{f} v \right) \right] k \end{aligned}$$

raggruppando a fattore comune $\frac{z}{f^2}$:

$$\frac{z}{f^2} \begin{bmatrix} -f \partial_u z \\ -f \partial_v z \\ z + u \partial_u z + v \partial_v z \end{bmatrix} \quad (2.13)$$

il vettore trovato è parallelo a $\mathbf{n}(u, v)$ se

$$\frac{z}{f^2} \begin{bmatrix} -f\partial_u z \\ -f\partial_v z \\ z + u\partial_u z + v\partial_v z \end{bmatrix} \times n = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

risolvendo i minori otteniamo un sistema sovradimensionato a tre equazioni:

$$\begin{aligned} fn_3\partial_u z + n_1[z + u\partial_u z + v\partial_v z] &= 0 \\ fn_3\partial_v z + n_2[z + u\partial_u z + v\partial_v z] &= 0 \\ n_2\partial_u z - n_1\partial_v z &= 0 \end{aligned} \quad (2.14)$$

Sapendo che il sistema è omogeneo e che $z > 0$ possiamo effettuare un cambio di variabile:

$$\tilde{z} = \ln(z) \quad (2.15)$$

che rende (2.12) lineare rispetto a $\partial_u \tilde{z}$ e $\partial_v \tilde{z}$:

$$\begin{aligned} [fn_3 + un_1]\partial_u \tilde{z} + vn_1\partial_v \tilde{z} &= -n_1 \\ un_2\partial_u \tilde{z} + [fn_3 + vn_2]\partial_v \tilde{z} &= -n_2 \\ n_2\partial_u \tilde{z} - n_1\partial_v \tilde{z} &= 0 \end{aligned} \quad (2.16)$$

Questo sistema, se il rango è minore di 2, risulta non invertibile ovvero se i tre determinanti sono uguali a zero:

$$\begin{aligned} fn_3[un_1 + vn_2 + fn_3] &= 0 \\ -n_1[un_1 + vn_2 + fn_3] &= 0 \\ -n_2[un_1 + vn_2 + fn_3] &= 0 \end{aligned} \quad (2.17)$$

Poiché n_1, n_2 e n_3 non possono annullarsi contemporaneamente, \mathbf{n} è un vettore di lunghezza unitaria, allora deve valere $un_1 + vn_2 + fn_3 = 0$. Sapendo che $[u, v, f]^\top$, sono le coordinate del punto $\overline{\mathbf{a}^p}$ nella cornice 3D $\mathbf{c}xyz$, questo succede solo se $\overline{\mathbf{a}^p}$ è nascosto nel contorno di occlusione. E' evidente che per un dato oggetto, il contorno di occlusione dipende dal modello di proiezione che viene assunto. Per i punti dell'immagine che non giacciono sul contorno di occlusione, il sistema (2.16) è facilmente invertibile, il che ci da la seguente coppia di equazioni differenziali parziali (PDE):

$$\nabla \tilde{z} = [\tilde{p}, \tilde{q}]^\top \quad (2.18)$$

nel quale

$$\begin{aligned}\tilde{p} &= -\frac{n_1}{un_1 + vn_2 + fn_3} \\ \tilde{q} &= -\frac{n_2}{un_1 + vn_2 + fn_3}\end{aligned}\quad (2.19)$$

Quindi in proiezione prospettica, integrare un campo normale equivale ad integrare un campo vettoriale $[\tilde{p}, \tilde{q}]^\top$, la soluzione dell'equazione(2.18) è

$$\tilde{z}(u, v) = \tilde{z}(u_0, v_0) + \int_{(r,s)=(u_0,v_0)}^{(u,v)} [\tilde{p}(r, s)dr + \tilde{q}(r, s)ds] \quad (2.20)$$

che dalla (2.15) risulta :

$$z(u, v) = z(u_0, v_0) \exp \left\{ \int_{(r,s)=(u_0,v_0)}^{(u,v)} [\tilde{p}(r, s)dr + \tilde{q}(r, s)ds] \right\} \quad (2.21)$$

Notiamo che la lunghezza della lente f deve essere nota, così come la posizione del punto principale \mathbf{o} , poiché da esso dipendono le coordinate u e v . Dalla somiglianza delle varie equazioni possiamo constatare che può essere esteso qualsiasi metodo di integrazione normale, a condizione che siano noti i parametri della camera. Possiamo quindi limitarci a risolvere la seguente coppia di PDE, che consideriamo come *Modello del problema*

$$\nabla z = \mathbf{g} \quad (2.22)$$

dove per (z, \mathbf{g}) si intende $(z, [p, q]^\top)$, $(z, m[p, q]^\top)$, $\mathbf{o}(\tilde{z}, [\tilde{p}, \tilde{q}^\top])$, a seconda che il modello di proiezione sia rispettivamente *ortogonale*, *weak perspective*, *prospettico*.

2.6 Problema differenziale

Una volta calcolate le normali alla superficie sorge il problema di determinare il valore di profondità $z(x, y)$ per ottenere una risoluzione completa. Consideriamo una superficie discretizzata con una griglia XY , e avente a disposizione per ogni vertice (x, y) della griglia un vettore normale noto $n(x, y) = (n_x, n_y, n_z)$ nella forma

$$n(x, y) = n(p, q) = \frac{1}{\sqrt{(1 + p^2 + q^2)}}(-p, -q, 1) \quad (2.23)$$

$p = \frac{\partial z}{\partial x}$ e $q = \frac{\partial z}{\partial y}$ sono le componenti del gradiente della funzione di profondità $z(x, y)$ della superficie.

$$p = -\frac{n_x}{n_z}, \quad q = -\frac{n_y}{n_z} \quad (2.24)$$

La relazione tra normale e gradiente può essere sfruttata per recuperare la funzione z . scrivendo q e p in termini della profondità z utilizzando un rapporto incrementale

$$p \simeq \frac{z(x+h, y) - z(x, y)}{h}, \quad q \simeq \frac{z(x, y+h) - z(x, y)}{h} \quad (2.25)$$

si ottiene un sistema lineare di equazioni che può essere risolto per recuperare i valori di z . Per costruire il sistema lineare si ha bisogno di una matrice, griglia regolare:

$$u_x = \frac{u_{i+1,j} - u_{i,j}}{h} \quad u_y = \frac{u_{i,j+1} - u_{i,j}}{h}$$

Per essere risolvibile ha bisogno di condizioni al bordo di tipo Neumann o Dirichlet che richiedono ulteriori informazioni sulla soluzione cercata, in particolare sono difficili da imporre su domini di forma più complessa di un rettangolo. Il problema principale di questa integrazione diretta è dato dal fatto che le normali ottenute dalla Stereo Fotometria sono calcolate in maniera indipendente l'una dall'altra. Ciò può determinare inconsistenze quando si prova ad integrarle, poiché le immagini utilizzate per recuperarle potrebbero contenere rumore, che va ad intaccare i vincoli di integrabilità.

$$\frac{\partial p}{\partial y}(x, y) = \frac{\partial q}{\partial x}(x, y) \quad (2.26)$$

per superare questi inconvenienti sono state sviluppate tecniche basate sul calcolo delle variazioni (Metodo di Horn [8]) che cerca la funzione che minimizza il funzionale

$$\int \int (z_x - p)^2 + (z_y - q)^2 dx dy \quad (2.27)$$

che ha associata l'equazione di Eulero-Lagrange $\Delta z = p_x + q_y$, che per essere risolvibile ha bisogno delle seguenti condizioni al contorno $(z_x, z_y) \cdot n_c = (p, q) \cdot n_c$ in cui n_c indica la normale esterna al dominio. Una volta che si ha il campo dei vettori normali alla superficie, distinguiamo le prime due componenti del vettore n_k , ottenuto normalizzando la terza componente. Nello stesso modo otteniamo una griglia approssimata del Laplaciano $f(x, y) = u_{xx} + u_{yy}$. Il profilo tridimensionale dell'oggetto può essere ricavato risolvendo l'equazione differenziale a derivate parziali

$$\Delta u(x, y) = f(x, y) \quad (2.28)$$

Δ rappresenta l'operatore di Laplace. Dopo aver imposto le condizioni al contorno di Dirichlet, si presuppone che l'oggetto osservato si trova su una superficie nera

piatta, l'equazione di Poisson è condizionata da uno schema alle differenze finite del secondo ordine. Consideriamo l'equazione bidimensionale di Poisson (2.28) in un rettangolo $[-A/2, A/2] \times [-B/2, B/2]$, con le seguenti condizioni al contorno:

$$\begin{aligned} u(x, -B/2) = u(x, B/2) = 0, \quad x \in [-A/2, A/2], \\ u(-A/2, y) = u(A/2, y) = 0, \quad y \in [-B/2, B/2] \end{aligned} \quad (2.29)$$

in cui $u_{i,0}, u_{i,s+1}, u_{j,0}$ e $u_{r+1,j}$ rappresentano i valori al bordo. Discretizzando l'equazione di Poisson tramite il metodo alle differenze finite, con passo di dimensione h , otteniamo un sistema lineare:

$$u_{i-1,j} + u_{i,j-1} - 4u_{i,j} + u_{i,j+1} + u_{i+1,j} = \tilde{f}_{i,j} \quad (2.30)$$

per $i = 1, \dots, r$ e $j = 1, \dots, s$ con $\tilde{f}_{i,j} = h^2 f_{i,j}$. Unendo i punti $u_{i,j}$ dalle colonne, otteniamo un sistema pentagonale di dimensione $p = r \times s$

$$I_s \mathbf{u}_{i-1} + T \mathbf{u}_i + I_s \mathbf{u}_{i+1} = \tilde{\mathbf{f}}_i \quad i = 1, \dots, r,$$

in cui

$$T = \begin{bmatrix} -4 & 1 & & & \\ 1 & -4 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & 1 & -4 \end{bmatrix} \in \mathbb{R}^{s \times s}, \quad \mathbf{u}_i = \begin{bmatrix} u_{i,1} \\ u_{i,2} \\ \vdots \\ u_{i,s} \end{bmatrix} \in \mathbb{R}^s, \quad \tilde{\mathbf{f}}_i = \begin{bmatrix} h^2 f_{i,1} \\ h^2 f_{i,2} \\ \vdots \\ h^2 f_{i,s} \end{bmatrix} \in \mathbb{R}^s$$

I_s rappresenta la matrice identità di dimensione s . Il sistema può essere riscritto come

$$A \mathbf{u} = \mathbf{f}$$

con

$$A = \begin{bmatrix} T & I_s & & & \\ I_s & T & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & I_s & T \end{bmatrix} \in \mathbb{R}^{rs \times rs}, \quad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_r \end{bmatrix} \in \mathbb{R}^{rs}, \quad \mathbf{f} = \begin{bmatrix} h^2 f_1 \\ h^2 f_2 \\ \vdots \\ h^2 f_r \end{bmatrix} \in \mathbb{R}^{rs} \quad (2.31)$$

Può essere risolto con qualsiasi metodo iterativo generale diretto o preconditionato adatto per matrici sparse di grandi dimensioni [7], o in particolare, tramite il metodo

di Poisson [2, 4] L'*integrazione delle normali* è un classico problema di ricostruzione 3D, che consiste nel calcolare la forma tridimensionale di una superficie a partire da un certo numero di norme. Solitamente è un problema ben posto, la risoluzione, anche nel caso in cui si conoscano tutti i pixel dell'immagine, non è sempre banale. Non esiste infatti un metodo che sia del tutto soddisfacente, spesso la difficoltà può essere dovuta al fatto che possono soddisfare contemporaneamente più requisiti.

Capitolo 3

Sperimentazione numerica

In questo capitolo verranno mostrati i test numerici, svolti su Matlab, per valutare le prestazioni dei metodi utilizzati per ricostruire una superficie 3D.

3.1 Valutazione dell'errore

Una volta noto il campo delle normali, possiamo vedere che la relazione (2.28) può essere risolta in quattro modi differenti:

1. Laplaciano esatto
2. Gradiente numerico con errore di ordine h^2
3. Gradiente numerico con errore di ordine h
4. Gradiente ricavato dalle immagini

verificheremo mediante dei test che tre di questi sono perfettamente equivalenti dal punto di vista dell'errore.

3.1.1 Laplaciano Esatto

Come abbiamo visto nel capitolo precedente, l'operatore di Laplace Δ è ottenuto dopo aver imposto le condizioni di Dirichlet (3.1), discretizzando l'equazione di Poisson con un metodo alle differenze finite (2.31). Per ricostruire la superficie abbiamo scelto una funzione test f della quale sono note le derivate parziali p e q ,

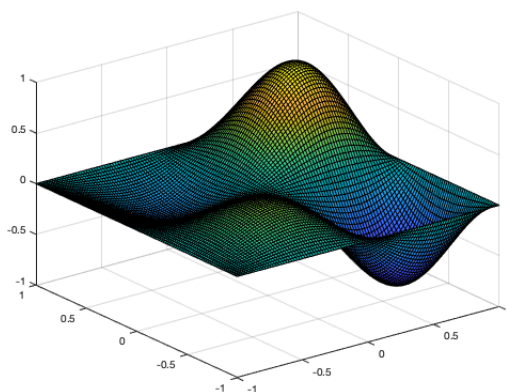


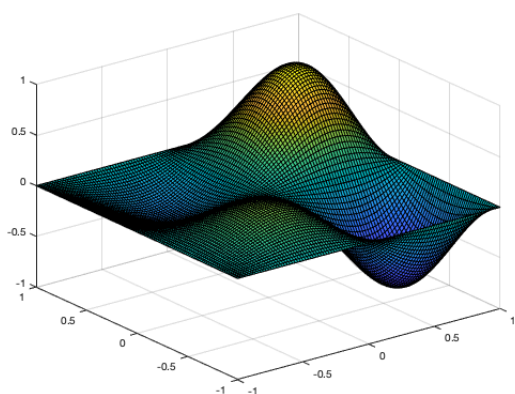
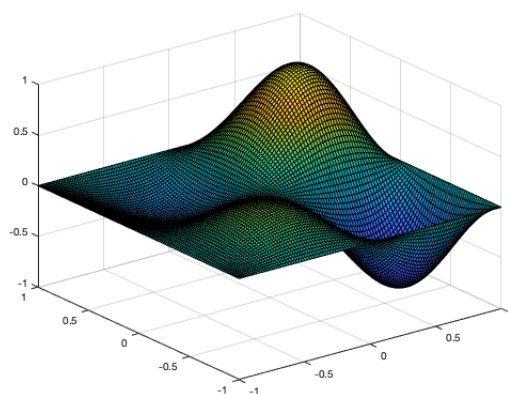
Figura 3.1: Superficie ricostruita con Laplaciano Esatto

abbiamo discretizzato il dominio di definizione di f con un passo h :

$$\Delta u = \frac{\partial p}{\partial x} + \frac{\partial q}{\partial y} \quad (3.1)$$

l'equazione differenziale viene risolta su dominio rettangolare. Cerchiamo in un punto generico il valore u_j che approssima u_i . In questo caso l'errore è dato esclusivamente dalla risoluzione dell'equazione differenziale.

3.1.2 Gradiente Numerico passo h e $2h$

(a) Gradiente Numerico passo $2h$ (b) Gradiente Numerico passo h

Se le derivate parziali non sono note, vengono approssimate con il rapporto differenziale utilizzando in un caso la formula alle differenze centrate:

$$\frac{\partial p}{\partial x} = \frac{p(x+h, y) - p(x-h, y)}{2h} \quad \frac{\partial q}{\partial y} = \frac{q(x, y+h) - q(x, y-h)}{2h} \quad (3.2)$$

e nell'altro caso la formula delle differenziale in avanti :

$$\frac{\partial p}{\partial x} = \frac{p(x+h, y) - p(x, y)}{h} \quad \frac{\partial q}{\partial y} = \frac{q(x, y+h) - q(x, y)}{h} \quad (3.3)$$

l'errore si comporta come $O(h^2)$ nella (3.2), e come $O(h)$ nella (3.3). Ci aspettiamo dunque che l'errore in questi due casi sia maggiore del caso in cui il laplaciano è calcolato in modo esatto, poiché all'errore dovuto alla risoluzione dell'equazione differenziale si aggiunge quello del calcolo delle derivate parziali.

3.1.3 Gradiente ricavato dalle immagini

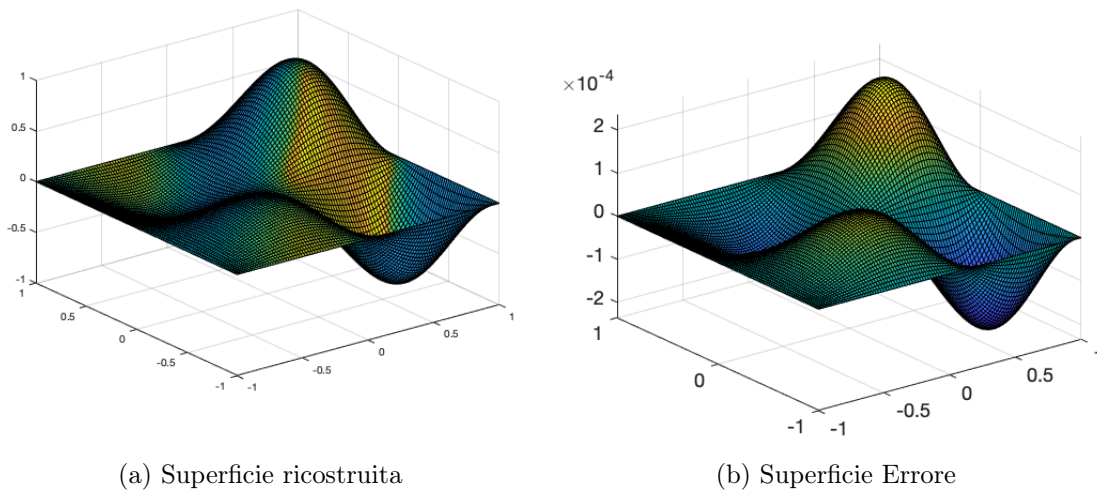


Figura 3.2: Superfici ricostruite mediante gradiente ricavato dalle immagini

In questo caso viene illustrato l'approccio di Poisson, per valutare l'accuratezza raggiungibile nella situazione ideale, quando tutte le ipotesi dei metodi sono soddisfatte, scegliamo una disposizione di q sorgenti luminose posizionate ad angoli regolari attorno all'oggetto e generiamo un insieme di immagini digitali applicando il modello diretto (2.1) a una superficie rappresentata da una funzione esplicita di due variabili $z(x, y)$. Nel documento [5] è possibile trovare una descrizione dettagliata

ta del test in oggetto. Nelle figure possiamo vedere la superficie ricostruita (3.2a) e la superficie dell'errore (3.2b).

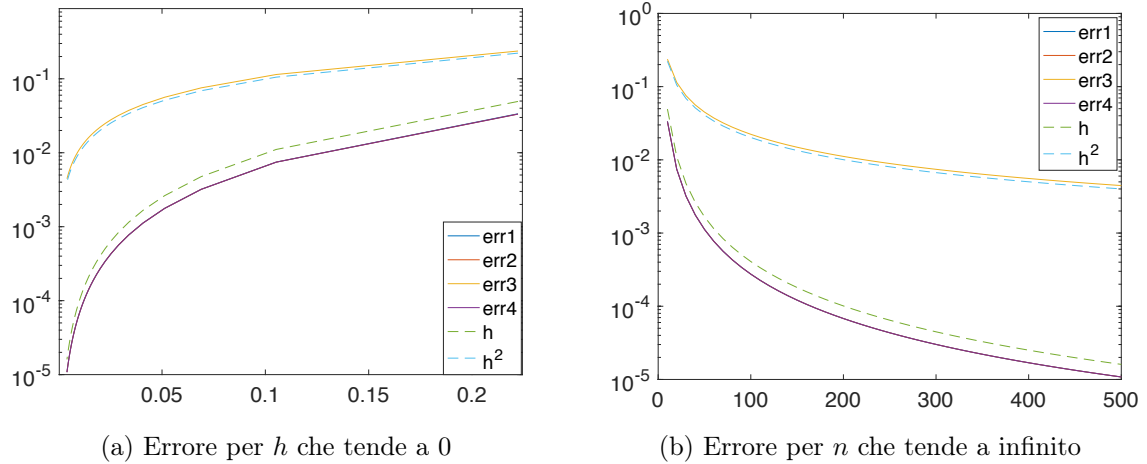


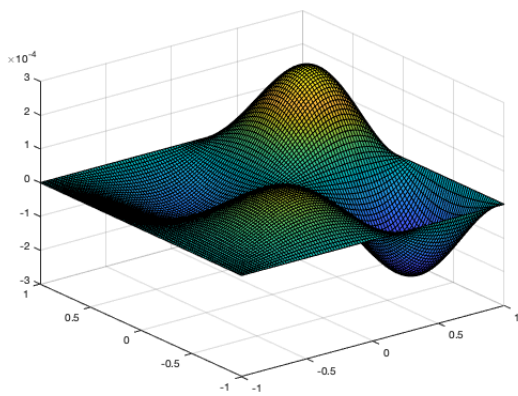
Figura 3.3: Andamento errore per h che tende a zero e per n che tende a infinito dei quattro metodi.

Nelle (Figura: 3.3a) ed (Figura: 3.3b) è possibile vedere come si comporta l'errore attribuendo ad h un valore da 0 a 500. Dal grafico (Figura: 3.3a) notiamo come i gli errori dei primi due casi e dell'ultimo siano sovrapposti, mentre il terzo metodo come ci aspettavamo tende ad h . Ciò che deduciamo è quindi che i metodi 1, 2, e 4 sono perfettamente equivalenti.

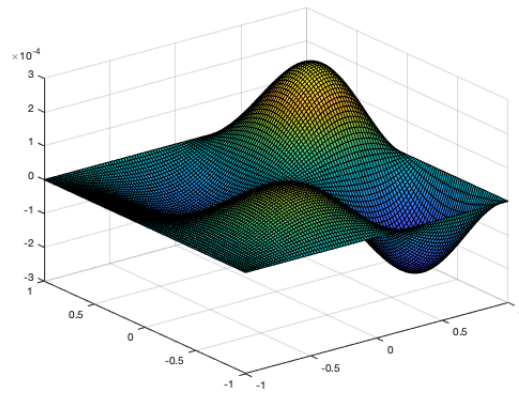
3.2 Superficie ricostruita con errore sperimentale

In questo paragrafo valuteremo come la ricostruzione dell'immagine viene distorta inserendo una perturbazione ds nella matrice delle immagini M . Abbiamo visto nel precedente paragrafo le superfici ricostruite nei vari metodi, ora verificheremo come esse si comportano con una perturbazione di

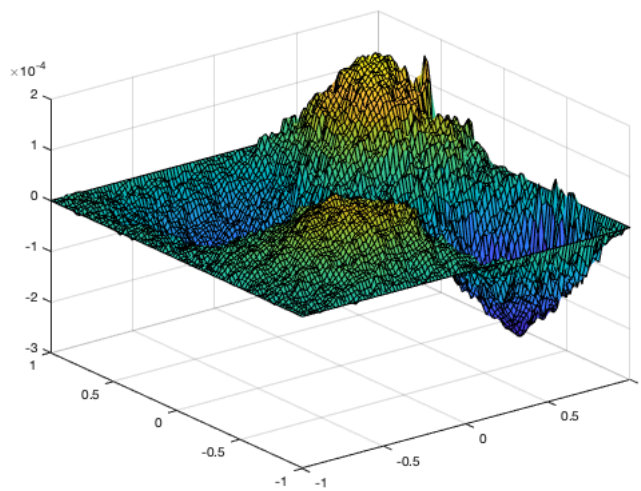
$ds = 10^{-3}$ e $ds = 10^{-2}$, per errori più piccoli di $ds = 10^{-3}$ le distorsioni della superficie non sono percepibili. Dalle figure (3.4) e (3.6) possiamo vedere come i primi due metodi (3.4a, 3.4b, 3.6a, 3.6b) non subiscano variazioni nella ricostruzione della superficie in quanto, i valori non vengono presi dalla matrice delle immagini come invece accade nell'ultimo metodo. E' possibile anche vedere come l'andamento dell'errore segua l'andamento della superficie in esame.



(a) Laplaciano Esatto

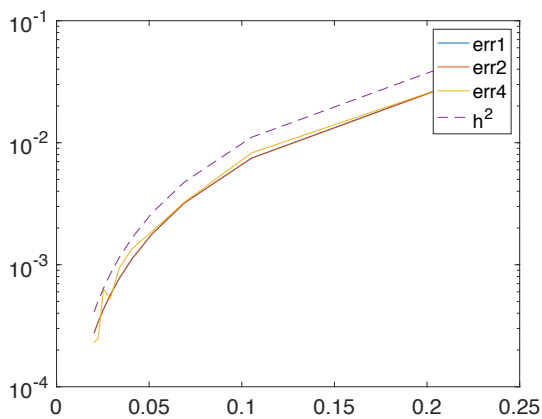


(b) Gradiente numerico

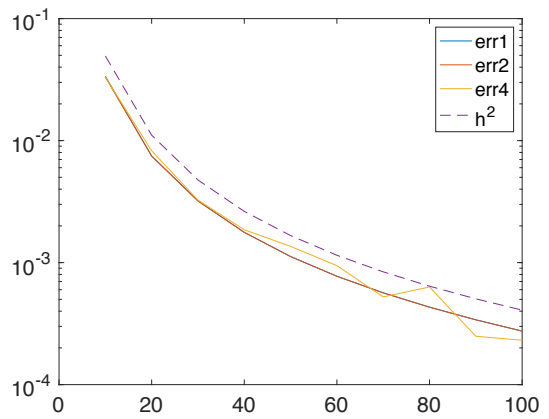


(c) Gradiente Ricavato dalle immagini

Figura 3.4: Superfici ricostruite con $ds = 10^{-3}$

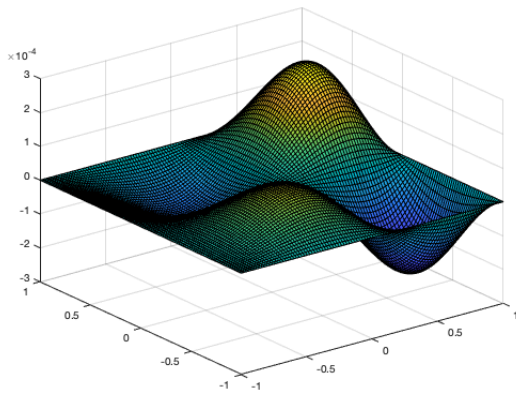


(a) Errore per h che tende a zero

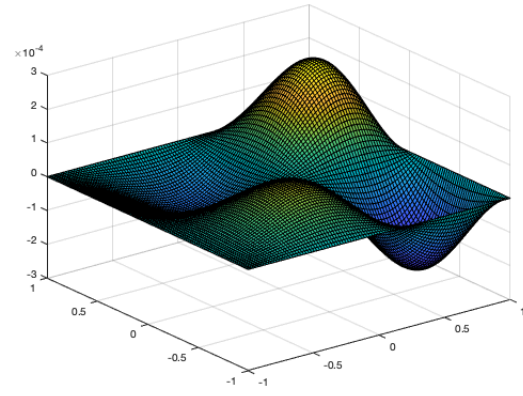


(b) Errore per n che tende a infinito

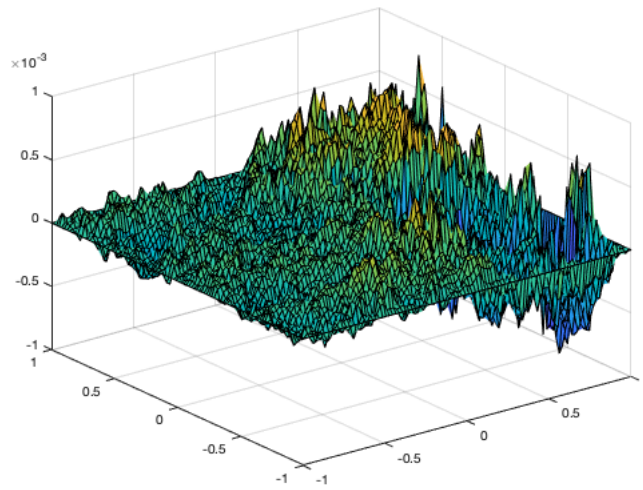
Figura 3.5: Errore relativo a $ds = 10^{-3}$



(a) Laplaciano Esatto

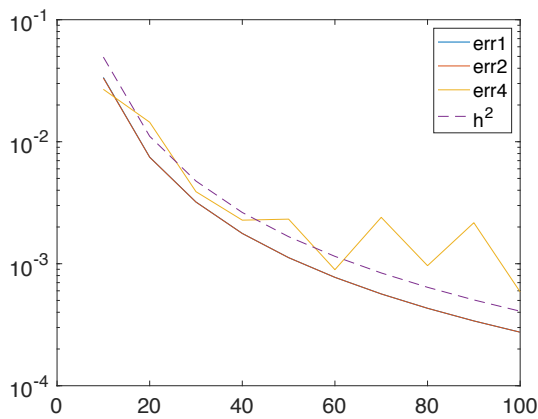


(b) Gradiente numerico

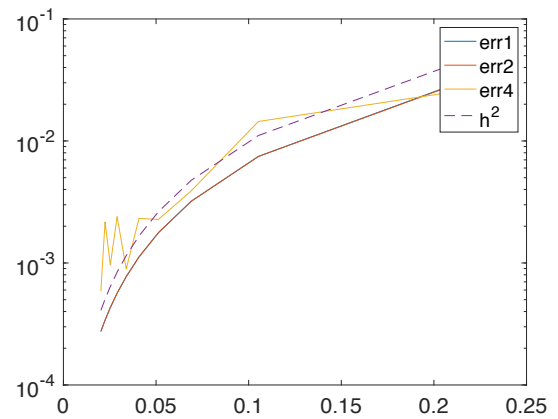


(c) Gradiente Ricavato dalle immagini

Figura 3.6: Superfici ricostruite con $ds = 10^{-2}$



(a) Errore per h che tende a zero



(b) Errore per n che tende a infinito

Figura 3.7: Errore relativo a $ds = 10^{-2}$

Bibliografia

- [1] Å. Bjöck, *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.
- [2] B. L. Buzbee, G. H. Golub, and C. W. Nielson, *On direct methods for solving Poissons equations*, SIAM Journal Numer. Anal., 7 (1970), pp. 627–656.
- [3] A. Chambolle *A uniqueness result in the theory of stereo vision: coupling shape from shading and binocular information allows unambiguous depth reconstruction*. Annales de l’I.H.P Analyse non linéaire 11(1):1–16, 1994
- [4] T. F. Chan and D. C. Resasco, *A domain-decomposed fast Poisson solver on a rectangle*, SIAM J. Sci. Stat. Comput., 8 (1987), pp. s14–s26.
- [5] A. Concas, R. Dessi, C. Fenu, G. Rodriguez, M. Vanzi, *Identifying the lights position in photometric stereo under unknown light*, arXiv:1809.06933 [math.NA], 2018.
- [6] Andrea Fusiello, *Visione computazionale. Tecniche di ricostruzione tridimensionale*, Franco Angeli, 2018.
- [7] G. H. Golub and C. F. Van Loan, *Matrix Computations*, , The John Hopkins University Press, Baltimore, third ed., 1996.
- [8] Berthold KP Horn and Michael J Brooks *The variational approach to shape from shading*, volume 33. Elsevier, 1986.
- [9] Yvain Queau, Jean-Denis Durou, Jean-Francois Aujol, *Normal Integration : A Survey*
- [10] Robert J Woodham. *Photometric method for determining surface orientation from multiple images*. Optical engineering, 19, 1980.