

Dottorato di ricerca in Informatica e Matematica
XXIX ciclo

Progetto di ricerca

Dottorando: Dott. Michele Giuliano Fiorentino

Tutor: Dott.ssa Eleonora Faggiano

Coordinatore

Prof. Donato Malerba

Firma del dottorando _____

Firma del tutor _____

1) Titolo della ricerca:

La Computer Science per leggere ed innovare gli studi sulla prospettiva nella geometria, nell'arte, nella Scuola Italiana.

2) Area nella quale si inquadra la ricerca:

Geometria

3) Obiettivi della ricerca

La geometria si presenta come una tra le più antiche discipline che l'umanità ha prodotto e ad essa fanno riferimento diversi sistemi assiomatici: geometria metrica, geometria proiettiva, geometria differenziale. Gli obiettivi della ricerca vanno ad inquadrarsi nell'assiomatica della geometria proiettiva ponendo particolare attenzione sui tre seguenti aspetti:

1. Lo studio, con opportuni strumenti (macchine matematiche e software), delle problematiche affrontate nel campo della geometria proiettiva nel passaggio dal concreto al virtuale, analizzando le caratteristiche e le proprietà che si conservano in tale passaggio e viceversa l'utilizzo della geometria proiettiva nel design di tali software e artefatti;
2. Le conseguenze e quindi le ricadute in vari ambienti dei problemi legati alla geometria proiettiva (ad esempio nel disegno, nella visualizzazione delle figure e delle opere d'arte, nel recupero delle stesse, ecc.);
3. L'attenzione ai fondamenti matematici della prospettiva nell'ottica dell'insegnamento-apprendimento della geometria, nella scuola secondaria di secondo grado, così come suggerito dalle Indicazioni Nazionali¹: lo studente "...approfondirà le relazioni tra le conoscenze acquisite in ambito geometrico e le problematiche di rappresentazione figurativa e artistica."

4) Motivazioni della ricerca

Negli ultimi quarant'anni, nell'ambito della Computer Vision sono stati fatti enormi progressi per quanto riguarda tutte quelle tecniche di visione artificiale o computazionale volte a riprodurre le abilità della visione umana in dispositivi elettronici per l'acquisizione e la comprensione delle immagini. I campi di applicazione sono soprattutto ingegneristici e trovano particolare importanza nelle interazioni uomo-macchina e modellazione di oggetti.

Dal punto di vista teorico, ci sono importantissimi risultati raggiunti in un'area di questa disciplina che potremmo chiamare "Geometric Computer Vision". Questa include la descrizione del modo di apparire degli oggetti e dei cambiamenti che questi assumono quando vengono osservati da diversi punti di vista, come una funzione della forma degli oggetti stessi.

Con l'utilizzo delle tecniche della Computer Vision diventa interessante implementare i più importanti teoremi e risultati della geometria proiettiva per andare oltre e sfruttarne le potenzialità nel mondo della realtà. Questo approccio differente caratterizza una soluzione alternativa alle problematiche classiche affrontate nella geometria proiettiva. D'altra parte sin dagli anni '80 c'è stato un crescente interesse nella comunità che si occupava del riconoscimento degli oggetti, per

¹ Indicazioni Nazionali per la Scuola Secondaria di Secondo Grado

muoversi al di là del semplice approccio tramite modellizzazione manuale, a favore di una sempre più invadente automatizzazione dell'acquisizione. In questo si è sempre cercato di puntare notevolmente alla ricerca degli invarianti geometrici. È proprio in parallelo a questo sviluppo che nel 1992 Richard Hartley e Oliver Faugueras iniziarono ad applicare la teoria della geometria proiettiva alle relazioni tra le visioni prospettiche multiple.

A partire da quanto fin qui illustrato si vogliono analizzare e affrontare alcuni di tali problemi con modalità innovative.

5) Stato dell'arte

Le basi della ricerca sono suddivise in vari micro settori, ciascuno dei quali verrà elencato nel seguito.

5.1 Le basi storiche della Computer Vision

La geometria proiettiva è stata utilizzata nella computer vision sin dal principio. Per esempio nel 1965 Roberts [1] nella sua tesi di dottorato ha usato la matrice di trasformazione proiettiva dal 3D al 2D per rappresentare la visualizzazione 3-dimensionale dei poliedri. Ha sviluppato un algoritmo lineare per determinare i parametri matriciali ed è stato in grado di visualizzare oggetti immersi nello spazio 3-dimensionale su piani (2-dimensionali) con l'uso della geometria proiettiva.

Nel corso degli anni, la proiezione prospettica è stata analizzata in molti modi differenti: la maggior parte degli studi parte dalle equazioni della proiezione centrale e quindi continua con alcune semplificazioni e nuovi gruppi di parametri che costituiscono il risultato teorico principale. Un esempio esplicativo è il problema della determinazione del minimo numero di punti di corrispondenza necessari per calcolare le trasformazioni tra due viste prospettiche dello stesso insieme di punti contenuti nello spazio 3-dimensionale.

5.2 Il background: Geometria proiettiva e trasformazioni

Un'interessante introduzione al piano geometricoproiettivo per la computer vision e quindi utile ai nostri scopi è in Mundy e Zisserman [2]. Qui vengono sviluppati i concetti di base e fondamentali per i passi successivi. Un approccio più formale è invece quello di Semple e Kneebone [3].

Una descrizione delle proprietà ed entità dello spazio proiettivo P_3 si deve a Hartley e Zisserman [4]: alcune di queste proprietà non sono altro che una generalizzazione di quelle del piano proiettivo (ad esempio le rette parallele che si incontrano ad l_∞ qui sono i piani paralleli che si intersecano a π_∞). Anche per questa parte è rilevante il lavoro di Mundy e Zisserman, nonché quello precedente di Hilbert e Cohn-Vossen [5] per alcune proprietà dettagliate di curve e superfici.

5.3 Camera Geometry e Single View Geometry

Uno strumento di acquisizione (ad esempio una fotocamera) è una funzione tra il mondo reale 3D (spazio degli oggetti) e un'immagine 2D: esso può essere pensato come una funzione,

$$X = (X, Y, Z)^T \mapsto x = (x, y)^T$$

Ci sono diverse tipologie di strumenti e quelli probabilmente più interessanti sono quelli a proiezione centrale. Le caratteristiche di questa tipologia di strumenti sono esaminate con i risultati

della geometria proiettiva: le entità geometriche (ad esempio il centro di proiezione e il piano dell'immagine) possono essere dedotte molto facilmente a partire dalla matrice di trasformazione.

Una gerarchia di queste tipologie è delineata in Aloimonos [6], invece Mundy e Zisserman [2] generalizzano alcune proprietà negli strumenti di acquisizione di tipo affine (dove la matrice di trasformazione ha delle proprietà particolari). Faugeras invece sviluppa le proprietà degli strumenti di acquisizione di tipo proiettivo in [7].

Si sono studiate le caratteristiche nella proiezione di punti e rette: Faugeras e Mourrain [8], e Faugeras e Papadopoulos [9] sviluppano la proiezione delle rette utilizzando le coordinate di Plücker; Koenderink [10], [11] e Giblin e Weiss [12] forniscono molte proprietà sulla generazione del contorno e sul contorno apparente, e la loro relazione con la geometria differenziale delle superfici.

5.4 Two-View Geometry

Nella two-view geometry si ampliano i punti di vista (l'obiettivo è quello di raggiungere l'*n*-view geometry): le due viste possono essere ottenute simultaneamente come in un impianto stereo o acquisite sequenzialmente, per esempio con uno strumento di acquisizione che si muove relativamente alla scena. Queste due situazioni sono geometricamente equivalenti. La geometria epipolare è la geometria intrinseca tra le due view, è indipendente dalla struttura della scena e dipende soltanto dai parametri interni dello strumento di acquisizione: la geometria epipolare è la cosiddetta geometria della visione stereoscopica. Essa si preoccupa di analizzare i vincoli e le relazioni geometriche che collegano due immagini 2-dimensionali della stessa scena 3-dimensionale, catturata da due strumenti di acquisizione con posizione e orientamento differenti. Anche questa geometria intrinseca può essere sviluppata con l'utilizzo di opportune matrici di trasformazione [13] [14] [15] [16].

5.5 Three-viewed *n*-view geometry

Con l'introduzione dei tensori trifocali (aventi proprietà analoghe a quelle delle matrici di trasformazione della single view e della 2-view geometry) si dà l'avvio ad una nuova geometria. La scoperta dei tensori trifocali può essere attribuita a [17] e [18] i quali li utilizzano per la ricostruzione della scena in caso di strumenti di acquisizione calibrati. Più tardi è stato mostrato che queste caratteristiche potevano essere applicate anche nel caso di strumenti non calibrati [19].

Il passaggio all'*n*-view geometry è naturale ed è una generalizzazione dei casi precedenti.

Uno dei risultati più importanti della *n*-view geometry non calibrata è quella della ricostruzione della struttura della scena esclusivamente a partire dalle corrispondenze punto-immagine. I metodi per la ricostruzione affine a partire dalla traslazione sono apparsi in Moons et al. [20].

6) Approccio al problema

Si studieranno dapprima a fondo i risultati teorici più profondi della teoria della geometria proiettiva. In questo studio verranno analizzati e scelti alcuni dei risultati che troveranno maggiore applicazione nel passaggio dal concreto al virtuale ed eventualmente trattati alcuni problemi ancora aperti.

In seguito si approfondiranno le funzionalità dei software e le possibilità di implementare delle nuove funzionalità non ancora presenti per studiare al meglio le caratteristiche dei problemi della geometria proiettiva. Con l'ausilio grafico e quindi, con la visualizzazione dei problemi ancora aperti, sarà possibile osservare sotto un'altra lente i possibili risultati.

Infine è prevista una ricerca sperimentale in ambito didattico che permetta di raggiungere le competenze richieste dalle Indicazioni Nazionali e che quindi fornisca un controllo delle potenzialità educative e esemplificative dei suddetti software.

Alla base dell'approccio al problema ci sarà sempre un percorso circolare che non verrà mai abbandonato in ciascuno dei tre obiettivi su indicati: passare dal teorema al software, studiarne le conseguenze visivamente e ritornare al teorema per una sua riformulazione o aggiunta di nuovi risultati.

7) Ricadute applicative

Le ricadute applicative si possono sintetizzare in due aspetti principali:

1. Ingegneristico - Artistico
2. Didattico

Nel campo ingegneristico-artistico, le potenzialità offerte dai software possono essere fondamentali per la rivalutazione e riproduzione digitale di opere d'arte e strutture architettoniche. In questo modo, con la potenza offerta dalle moderne tecnologie sarà possibile progettare o riprogettare opere di ammodernamento, messa in sicurezza o ristrutturazione delle stesse.

Per quanto riguarda invece l'ambito didattico, i problemi legati alla prospettiva si manifestano sin dai primissimi anni scolastici. Ad esempio nella "rappresentazione della realtà come si vede". Già si è fatto uso in passato di artefatti per la realizzazione di tali rappresentazioni: l'utilizzo di un prospettografo, ad esempio, sebbene abbia di certo facilitato tale problema presenta alcune limitazioni legate alla distanza tra il vetro e l'oculare che non può superare la lunghezza del braccio dell'utilizzatore. Non solo: il disegno realizzato può man mano coprire l'oggetto ed impedire il progredire del lavoro.

La necessità di superare questo tipo di problemi è alla base della ricerca che intende sviluppare gli strumenti software adeguati.

8) Riferimenti bibliografici

[1] L.G. Roberts, Machine Perception of three-dimensional solids, *Optical and Electro-optical Information Processing*, editor Tippett, J. Et al., MIT Press, pagg. 159-197, 1965

[2] J. Mundy, A. Zisserman. *Geometric Invariance in Computer Vision*. MIT Press, 1992

[3] J. G. Semple, G. T. Kneebone. *Algebraic Projective Geometry*. Oxford University Press. 1979

[4] R. Hartley, A. Zisserman, Multiple view Geometry in computer Vision second edition, Cambridge University Press, 2003

[5] D. Hilbert, S. Cohn-Vossen. *Geometry and the Imagination*. Chelsea, NY, 1956.

[6] J. Y. Aloimonos. Perspective Approximations. *Image and Vision Computing*, 8(3): pagg. 177-192. Agosto 1990

[7] O. D. Faugeras. *Three-Dimensional Computer Vision: a geometric Viewpoint*. MIT Press, 1993.

- [8] O. D. Faugeras, B. Mourrain. *On the geometry and algebra of point and line correspondences between N images*. In Proc. International Conference on Computer Vision. Pagg:485-492. Springer-verlag, 1994.
- [9] O.D. Faugeras, T. Papadopoulo. *Grassman-Cayley algebra for modelling systems of cameras and the algebraic equations of the manifold of trifocal tensors*. Technical Report 3225, INRIA, Sophia-Antipolis, Francia, 1997.
- [10] J.J. Koenderink. *What does the occluding contour tell us about solid shape?* *Perception*, 13:321-330, 1984.
- [11] J.J. Koenderink. *Solid Shape*. MIT Press, 1990.
- [12] P. Giblin, R. Weiss. *Reconstruction of surfaces from profiles*. In Proc. 1st International Conference on Computer Vision, London, pagg:136-144. Londra, 1987.
- [13] H.C. Longuet-Higgins. *A computer algorithm, for reconstructing a scene from two projections*. *Nature*, 293:133-135. Settembre 1981.
- [14] T.S. Huang, O.D. Faugeras. *Some properties of the E-matrix in two view motion estimation*. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 11, pagg:1310-1312, 1989.
- [15] S. J. Maybank, *Theory of reconstruction from image motion*. Springer-Verlag, Berlin, 1993.
- [16] B. K. P. Horn. *Relative orientation*. *International Journal of Computer Vision*, 4, pagg:59-78, 1990.
- [17] M.E. Spetsakis, J. Aloimonos. *A multi-frame approach to visual motion perception*. *International Journal of Computer Vision*, 16(3), pagg:245-255, 1991.
- [18] J. Weng, N. Ahuja, T.S. Huang. *Closed-form solution and maximum likelihood: A robust approach to motion and structure estimation*. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1988.
- [19] R. I. Hartley. *Projective reconstruction from line correspondence*. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 1994.
- [20] T. Moons, L. Van Gool, M Van Diest, E. Pauwels. *Affine reconstruction from perspective image pairs*. In J.Mundy, A. Zisserman, and D. Forsyth, editors, *Application of Invariance in Computer Vision*, LNCS 825, Springer-Verlag, 1994.

9) Fasi del progetto

Il progetto si suddivide nelle seguenti fasi:

FASE 0: Studio approfondito della letteratura esistente

FASE 1: Analisi dei software

FASE 2: Implementazione delle caratteristiche con l'ausilio di software

FASE 3: Sperimentazione didattica

FASE 4: Analisi dei risultati

FASE 5: Raccolta dati e conclusione della ricerca

10) Valutazione dei risultati.

I risultati della ricerca potranno essere valutati in maniera sperimentale, verificando l'attendibilità dei prodotti: tale verifica si baserà sulla rilevanza e validità dei teoremi nelle applicazioni che verranno sviluppate.

Inoltre nel passaggio tra il mondo concreto e quello virtuale, la "geometric computer vision" ci offre degli strumenti di controllo per verificare la validità degli algoritmi utilizzati per la digitalizzazione degli scenari osservati. I nuovi algoritmi che eventualmente verranno creati saranno valutati in maniera positiva se passeranno i test già utilizzati.

Infine gli obiettivi di natura didattica verranno valutati direttamente in classe durante la sperimentazione: le fasi saranno addestramento, apprendimento e verifica. Attraverso l'analisi dei protocolli prodotti dai ragazzi si potrà valutare l'efficacia dei metodi.

11) Eventuali referenti esterni al Dipartimento

Gli eventuali nominativi saranno precisati in seguito