

Capitolo 8

Applicazioni e considerazioni conclusive

8.1 La telecamera tridimensionale

Il sistema di scansione a luce debolmente strutturata permette di scandire un oggetto e di ricavarne un'immagine tridimensionale da visualizzare sullo schermo di un computer o su altro display. L'utente che visualizza l'immagine ha la possibilità di spostare il suo punto di vista e quindi di cogliere i dettagli dell'oggetto, normalmente nascosti, se l'immagine è rappresentata da una semplice fotografia bidimensionale.

Le applicazioni di tale tecnica sono svariate. Per esempio, si può pensare di scandire con il sistema suddetto un gruppo di sculture e bassorilievi, metterli in un catalogo d'arte e visualizzarli sullo schermo; oppure si può costruire un archivio di autovetture incidentate i cui danni possono meglio essere valutati dalle compagnie di assicurazione; o si può costruire un catalogo di oggetti da vendere per corrispondenza via commercio elettronico sul Web, così che l'acquirente potrà esaminare meglio l'oggetto da acquistare.

L'oggetto chiave per poter ottenere immagini tridimensionali deve essere facile da usare e di basso costo. Si può pensare ad una telecamera digitale che contenga, al di là del necessario per svolgere la sua funzione di base, una serie di componenti che le permettano la scansione tridimensionale in tempo reale e l'immagazzinamento in memoria del risultato di tale scansione.

Le operazioni che devono essere eseguite sono:

- proiezione e spostamento dell'ombra sull'oggetto
- cattura di una sequenza di immagini

- digitalizzazione delle immagini e processamento in tempo reale di alcuni dei pixel che stanno sul bordo dell'ombra attivo nell'immagine di lavoro corrente, ottenendo l'informazione sulla profondita' per tutti i pixel processati.

Per tale operazione di calcolo in tempo reale, occorre un DSP che analizzi i pixel di due immagini (quella corrente I_{curr} e quella precedente I_{prev}) a due a due, e che in tempo reale riempia di dati (numeri reali) un gruppo di matrici $M_1, M_2, M_3, \dots, M_N$ di lavoro. Contemporaneamente l'informazione sulla profondita' (e quindi le altre 2 coordinate nello spazio) associata ai pixel processati deve essere memorizzata in 3 matrici X_o, Y_o, Z_o di numeri reali. Queste matrici vanno aggiornate immagine per immagine via via che tutti i pixel vengono processati. Quando finisce la scansione, cioe' quando l'operatore preme il comando di stop, si ottengono le coordinate spaziali dei punti dell'oggetto scandito memorizzate nelle tre matrici X_o, Y_o, Z_o .

Si consideri un'immagine di $N=320 \times 240 = 76,800$ pixel. Tale immagine non e' ad alta risoluzione, ma il numero N permette di fissare le idee. Per risoluzioni maggiori, e' sufficiente rifare i calcoli per N piu' grande. Per memorizzare I_{curr} e I_{prev} sono necessarie 2 matrici ciascuna di $N \times 8$ bit (le immagini sono a 256 livelli di grigio che possono essere codificati usando 8 bit). Quindi serve in totale una memoria di $2 \times N \times 8 = 1,228,800$ bit. Questa memoria deve essere a lettura e scrittura rapida e puo' inoltre essere volatile, dato che le informazioni devono restare in memoria solo per il tempo strettamente necessario per l'elaborazione da parte del DSP. Si tratta quindi essenzialmente di una memoria di tipo DRAM.

Si supponga di dover utilizzare un gruppo di matrici $M_1, M_2, M_3, \dots, M_{10}$ di lavoro necessarie per l'immagazzinamento dei dati di passaggio utili per l'elaborazione. Tali dati sono numeri reali ciascuno dei quali richiede $4 \times 8 = 32$ bit. Pertanto il numero complessivo di bit necessari per le tutte le 10 matrici e' $10 \times N \times 4 \times 8 = 24,576,000$ bit. Anche in questo caso si tratta di dati che possono essere conservati in una memoria DRAM. Considerando un taglio commerciale di memoria DRAM a basso costo, due componenti da 16M bit sono sufficienti sia per la memoria di immagine che per la memoria di lavoro. Le memorie DRAM sono attive solo durante la scansione e quindi possono essere spente o messe in stand-by quando la telecamera non scandisce, risparmiando potenza dalla batteria della telecamera. E' noto

infatti che le memorie DRAM sono veloci in scrittura ed in lettura, ma consumano una considerevole potenza in operazione.

Le tre matrici finali X_o, Y_o, Z_o (destinate a memorizzare le coordinate spaziali definitive dei pixel processati e pertanto a contenere numeri reali) richiedono in totale $3 \times N \times 4 \times 8 = 7,372,800$ bit. Essendo tali matrici addette all'immagazzinamento di dati definitivi, la memoria deve essere del tipo non volatile, scrivibile e cancellabile elettricamente. Si tratta quindi di una memoria Flash, che non consuma potenza una volta che la memoria viene scritta e quindi riempita con i dati dell'immagine scandita. Un taglio commerciale a basso costo che fa allo scopo e' di 8M bit. Va da se' che per una risoluzione doppia in numero di pixel, occorre un DSP in grado di processare a velocita' doppia, una DRAM da 64M ed una Flash da 16M. Questi tagli sono da considerare oggi ancora a basso costo, aggirandosi il prezzo di tali componenti tra i 7 e 10 \$ ognuno, prezzi destinati comunque ad essere abbattuti drasticamente nel giro di pochissimi anni.

Sono comunque necessarie alcune considerazioni pratiche. Mentre per la memoria DRAM ed il DSP non si intravedono problemi di velocita' dei componenti rispetto alla velocita'di processamento richiesta dalla scansione, ne' e' necessaria una memoria maggiore nel caso si vogliano effettuare una serie di scansioni successive su diversi oggetti, la memoria Flash soffre di alcune limitazioni sia intrinseche per la peculiarita' di questa tecnologia, che dovute all'applicazione in esame. In particolare:

- la lentezza di scrittura (almeno 3 ordini di grandezza superiore a quella delle DRAM), per cui il processo di scrittura delle le matrici X Y Z potrebbe rallentare l'operazione in tempo reale.
- la necessita' di usare M memorie Flash da 8M (o da 16M per doppia risoluzione) se si vogliono immagazzinare nella telecamera M immagini prima di scaricarle in un PC per archivarle su Hard Disk, visualizzarle o comunque utilizzarle. Supponendo $M=100$, necessita una banco di memoria Flash da 800 Mbit, sicuramente non commerciale a bassi costi al giorno d'oggi, quando memorie Flash integrate in mini-schede per macchine fotografiche raggiungono densita' dell'ordine di 40 Mbyte a costi non trascurabili rispetto a quello della macchina fotografica digitale o della telecamera stessa.

Un'alternativa sarebbe quella di utilizzare come memoria non volatile un Hard Disk magnetico. Questa non e' una soluzione ideale ed elegante in quanto si tratta di un dispositivo poco adatto ad essere integrato in un'apparecchiatura portatile come una telecamera, contenendo delicati organi meccanici in movimento. In alcune macchine fotografiche digitali bidimensionali in commercio viene adoperato un floppy disk con relativo driver, con svantaggi di dimensioni e ancora di scarsa robustezza meccanica. La soluzione ideale resta un sistema di memoria Flash, in un'implementazione che consenta di ridurre gli inconvenienti sopramenzionati.

Per la lentezza della scrittura, e' indispensabile che i banchi di memoria Flash o l'architettura della memoria Flash stessa permetta un certo grado di parallelismo e di pipeline in scrittura. Aumentando il numero delle memorie Flash disponibili, si puo' ottenere una riduzione della velocita di scrittura a livello sistema, se non di singola memoria. Questo problema puo' essere aggirato conservando temporaneamente anche i valori delle matrici X Y Z in una memoria DRAM e poi effettuare il riversamento in memoria Flash una volta terminata la scansione. Con quest' approccio si aumenta la memoria DRAM necessaria di 7,372,800 bit. Tuttavia, sommando questo numero di bit a quello necessario per immagazzinare le matrici di lavoro M1, M2, M3, M10 di lavoro, restiamo al di sotto di 2x16 Mbit quindi non aumenta il numero di componenti DRAM richiesti. L'altro inconveniente, inevitabile in quest'approccio, e' il tempo di latenza necessario tra una scansione e la successiva. Infatti, in caso di riversamento a fine scansione, occorrera' attendere che tutta la memoria delle matrici X Y Z sia scritta nella Flash. Quest'operazione richiede di solito, per una Flash commerciale da 8 Mbit, alcuni secondi o qualche decina di secondi.

Se comunque si desidera poter scandire e conservare nella telecamera M immagini prima del riversamento dell'intera memoria Flash su un Hard Disk di un PC, diventa in ogni caso indispensabile aumentare le dimensioni globali della memoria Flash residente nella telecamera.

Secondo la legge empirica di Moore, un componente raddoppia la propria velocita' di calcolo e una memoria raddoppia la sua densita' di integrazione ogni 18 mesi. Nelle memorie Flash e' comunque possibile, nel medio termine, un'accelerazione della legge di Moore se verranno industrializzate memorie Flash a Singolo Elettrone.

Com'è noto, in una memoria Flash le cariche sono immagazzinate in un transistor FAMOS con floating gate. A seconda se la floating gate sia svuotata o piena di un numero E di elettroni, si associano a tali stati i valori 1 e 0 del bit. Il numero di elettroni E si aggira intorno ai 300,000 nelle celle delle memorie Flash commerciali. Il circuito di lettura della memoria deve discriminare tra 0 e 300,000 elettroni, ovvero tra due tensioni di soglia del transistor FAMOS distanti di circa 4-5 V. Tuttavia, recenti scoperte sul controllo del numero di elettroni E immagazzinati sulla floating gate, hanno condotto alla commercializzazione di memorie Flash multilivello, in cui nella floating gate delle celle di memoria vengono conservati pacchetti di elettroni di taglio più piccolo. Semplificando, nelle memorie INTEL di tipo Strataflash, si immagazzinano quattro pacchetti con un numero approssimativo di 0, 100K, 200K e 300K elettroni. Un opportuno amplificatore di lettura integrato nel chip è in grado di discriminare i 4 livelli, ovvero 4 tensioni di soglia distanziate di circa 1 V, cui vengono associate le combinazioni di bit 00, 01, 10 ed 11, ottenendo in pratica l'immagazzinamento di 2 bit per cella.

Più recenti ricerche hanno dimostrato che sia possibile una quantizzazione spinta fino ad un singolo elettrone per cella di memoria. La cella di memoria Flash così realizzata è non solo più piccola, in quanto per l'intrappolamento del singolo elettrone è necessario che la floating gate abbia dimensioni paragonabili alla lunghezza d'onda dell'elettrone stesso, ma offre anche il potenziale di immagazzinare un numero finito e controllato di elettroni, quindi di ottenere una potente Flash multilivello. Da celle Flash sperimentali risulta che la differenza in tensione di soglia tra una cella contenente un elettrone ed una cella scarica è di circa 56 mV, sufficiente per essere discriminata da un opportuno amplificatore di lettura.

In conclusione, studi più approfonditi sulle Memorie Flash multilivello ed a singolo elettrone renderanno fattibili ed a basso costo telecamere per la ripresa di immagini tridimensionali.

8.2 Conclusioni

Il sistema di scansione presentato in questo contesto risente di tutti i difetti tipici della scansione basata su metodi attivi. In particolare la ricostruzione di superfici riflettenti e' particolarmente critica. Il metodo a luce debolmente strutturata, essendo basato essenzialmente sulla proiezione di un'ombra, e' difficilmente applicabile nel caso di superfici che esibiscono bassi valori di contrasto. Ne e' un esempio un oggetto nero o di colore scuro. In tal caso e' molto difficile distinguere l'ombra dalla superficie stessa, rendendo inattuabili tutti i passi che conducono alla ricostruzione tridimensionale.

La forza del metodo, come piu' volte accennato, sta nella possibilita' di poter mettere a punto un sistema di scansione pratico, efficiente ed economico. Inoltre come visto nel precedente paragrafo, il metodo e' agevolmente implementabile in hardware senza costi eccessivi adoperando l'attuale tecnologia.

Sviluppi futuri del presente lavoro possono essere molteplici. Sarebbe interessante per esempio avere la possibilita' di integrare le ricostruzioni di uno stesso oggetto visto da angolazioni differenti. In tal modo verrebbero eliminate sensibilmente le aree non scandite e si potrebbero sfruttare le informazioni ridondanti per migliorare l'accuratezza del modello. Altri sforzi possono essere rivolti nella fase di calibrazione. Sarebbe estremamente comodo poter valutare dinamicamente la posizione della telecamera e del punto luce (insieme ai vari parametri intrinseci) durante la scansione. In tal modo si eviterebbe di dover svolgere tale operazione all'inizio e si darebbe la possibilita' di muovere la telecamera durante la ricostruzione, fornendo cosi' maggiore flessibilita' all'operatore.

Applicazioni e considerazioni conclusive