

---

## Esposizione riassuntiva

L'elaborazione elettronica di immagini e video riveste un ruolo importante nello sviluppo di tecnologie per la sicurezza. Le telecamere per videosorveglianza sono molto diffuse come strumento di riduzione della criminalità e anche in ambito forense si usano tecniche di analisi elettronica di immagini. In questa tesi sono stati approfonditi due problemi legati ad applicazioni di sicurezza: la ricostruzione di documenti trasformati in strisciole di carta da una macchina distruggi-documenti e il miglioramento di video acquisiti in condizioni di scarsa illuminazione.

Per quanto riguarda la ricostruzione documenti il sistema che è stato sviluppato comprende, come prima fase, l'acquisizione delle strisce di carta con uno scanner. Dopo una pre-elaborazione ogni striscia è rappresentata da un'immagine digitale, mentre un'altra immagine, detta di maschera, contiene l'informazione necessaria a separare la striscia dallo sfondo di acquisizione. Il contenuto visivo delle strisce deve essere opportunamente codificato per la ricostruzione, mentre la forma del contorno dei pezzi, solitamente usata nella ricostruzione automatica dei puzzle o dei frammenti di opere d'arte, non fornisce in questo caso l'informazione necessaria. Dopo un primo tentativo di descrizione del contenuto visivo delle strisce per mezzo dei descrittori MPEG-7, si è passati a delle *feature* più specifiche. Sono state considerate importanti per la rappresentazione del contenuto visivo le seguenti feature: interlinea, tipo di carattere, numero di righe di testo, posizione della prima riga di testo, posizione dell'ultima riga di testo, che caratterizzano il testo stampato; indice di carta quadrettata, utile nel caso di documenti scritti a mano; presenza di

un evidenziatore, colore dell'inchiostro, colore della carta, energia dei contorni del testo, bordo della striscia, caratteristiche usate sia per il testo stampato che per quello scritto a mano. Sono stati sviluppati gli algoritmi che estraggono automaticamente suddette feature dall'immagine digitale delle strisce. Gli algoritmi sono stati esclusivamente progettati per far fronte alle specificità delle strisce.

Sulla base delle feature le strisce possono venire raggruppate in modo tale che le strisce che appartenevano alla stessa pagina nei documenti originali vengano assegnate allo stesso gruppo e ci siano idealmente tanti gruppi quante erano le pagine. Per questo scopo si è utilizzato un algoritmo di *clustering* gerarchico. Il numero di gruppi che deve essere formato è scelto in maniera automatica dall'algoritmo in un intervallo selezionato dall'utente. Il clustering è efficace nel migliorare le prestazioni della ricostruzione con l'aiuto del computer. Il clustering permette inoltre di ridurre il tempo di calcolo del computer durante l'interazione online con l'operatore umano. La ricostruzione al computer è modellizzata come un'operazione di ricerca di immagini. L'utente sceglie una striscia e le strisce più simili ad essa vengono selezionate dal computer e presentate sul monitor (ordinate per valori di somiglianza decrescenti). Tra di esse l'utente individua le strisce che si attaccano correttamente a quella in questione e le incolla virtualmente. Il processo viene ripetuto iterativamente finché la ricostruzione è completata.

In un scenario di ricostruzione totalmente automatica è il computer che deve individuare le strisce che si attaccano correttamente. In questo caso viene sfruttata l'informazione contenuta nei bordi della striscia, lungo i quali la striscia si attacca: viene estratta la successione di valori di grigio dei pixel sul bordo destro o sinistro. Il problema è modellizzato come un problema di ottimizzazione combinatoria e si è dimostrata la sua NP-Completezza. Poiché è NP-Completo, bisogna usare degli algoritmi sub-ottimali per la sua soluzione. Come primo algoritmo viene proposto un matching locale: dato un pezzo, quello che si abbina correttamente alla sua destra è il pezzo il cui bordo sinistro è il più simile al bordo destro del pezzo sotto esame. Tuttavia ci possono essere degli errori, dovuti al fatto che il bordo presenta del rumore causato dal processo di taglio della macchina distruggi-documenti e dall'acquisizione. Si è quindi esplorata una soluzione di matching globale, in cui il problema viene

modellizzato come un Problema di Assegnazione: ogni bordo sinistro deve essere assegnato ad un bordo destro in modo che la somiglianza complessiva sia massimizzata.

In conclusione i contributi originali sviluppati in questa tesi per quanto concerne il problema di ricostruzione documenti sono i seguenti:

1. la caratterizzazione del problema;
2. la progettazione di un certo numero di feature di basso livello che descrivono il contenuto visivo delle strisce; le feature vengono estratte automaticamente dal computer;
3. la definizione di un algoritmo per raggruppare le strisce appartenenti ad una stessa pagina;
4. la modellizzazione del problema come un problema di ottimizzazione combinatoria e la conseguente definizione di algoritmi sub-ottimali per la sua soluzione automatica.

Il secondo problema che è stato studiato durante il dottorato è il miglioramento di immagini e video acquisiti in condizioni di illuminazione scarsa o non uniforme. Sono state considerate sia immagini a bassa dinamica che immagini ad alta dinamica. Nel secondo caso il miglioramento dell'immagine viene combinato a una riduzione della dinamica, come spiegato qui di seguito. Le immagini ad alta dinamica sono immagini capaci di acquisire un alto range di luminosità. Il Sistema Visivo Umano ha un comportamento ad alta dinamica: quando si guarda verso una finestra da una stanza si possono distinguere sia i dettagli interni che quelli esterni. I sistemi di acquisizione non hanno sempre la stessa capacità e la fotografia risultante potrebbe essere saturata nel bianco nella parte esterna o troppo scura nella parte interna. Esistono oggi delle tecniche per l'acquisizione di immagini ad alta dinamica che usano metodi diversi: si combinano delle foto di una stessa scena acquisite con diversi parametri di esposizione, oppure si usano dei sensori ad alta dinamica, quali il sensore CMOS logaritmico o il sensore CMOS lineare-logaritmico. I display comuni, tuttavia, presentano bassa dinamica, e per la visualizzazione delle immagini bisogna quindi operare una riduzione della

dinamica. L'algoritmo per la riduzione della dinamica che è stato considerato in questa tesi è il Locally Adaptive dynamics Reduction (appartenente alla serie LARx). Rispetto alla letteratura esistente esso ha il vantaggio di essere computazionalmente leggero e perciò particolarmente adatto ad applicazioni in tempo reale, quali video sorveglianza e assistenza alla guida. Alla stregua di molti altri algoritmi in letteratura esso si basa sulla teoria Retinex, secondo la quale, quando osserviamo un oggetto, l'immagine formata nell'occhio è il prodotto dell'illuminazione per la riflettanza dell'oggetto. E' l'illuminazione la parte che presenta un'alta dinamica, mentre la riflettanza corrisponde ai dettagli dell'oggetto e ha bassa dinamica. Stando a questa teoria, per migliorare l'aspetto delle immagini basta comprimere la dinamica dell'illuminazione e mantenere inalterata o evidenziare la riflettanza. La separazione dell'immagine in riflettanza e illuminazione è tuttavia un problema *ill-posed* e in letteratura sono state proposte diverse soluzioni. L'algoritmo LAR stima l'illuminazione usando un filtro passa-basso *edge preserving*. La sua implementazione, facendo uso di un Filtro Ricorsivo Razionale, ha come risultato quello di produrre un operatore computazionalmente leggero.

In video sorveglianza il problema affrontato riguarda l'acquisizione in condizioni di scarsa illuminazione. L'algoritmo LAR per sequenze video (LARS) è stato ottimizzato per questa applicazione, in cui il numero di parametri dell'algoritmo che devono venire manualmente settati dall'utente non esperto deve essere basso e potrebbe essere richiesta l'elaborazione in tempo reale. Poiché a basse luminosità le telecamere sono piuttosto rumorose, e poiché il rumore diventa molto evidente dopo l'elaborazione, è stato sviluppata una nuova versione dell'algoritmo per far fronte a tale problema.

Nell'applicazione di assistenza alla guida una telecamera ad alta dinamica è montata sullo specchietto retrovisore dell'auto. In questo caso il sensore ad alta dinamica risulta molto utile in quanto l'illuminazione può cambiare rapidamente mentre l'auto si muove, per esempio quando l'auto entra in una galleria o quando si ha una situazione di contro-sole. Con l'idea di realizzare l'elaborazione direttamente sulla telecamera è stata sviluppata un'implementazione dell'algoritmo LARS su hardware a basso costo di tipo FPGA. Inoltre la coerenza temporale è stata tenuta in particolare considerazione per evitare fastidiosi scintillii.

Sebbene in letteratura esistano numerosi algoritmi per la riduzione della dinamica, il problema di come valutare le loro prestazioni non è ancora stato risolto. Solitamente si fa una valutazione qualitativa di tipo soggettivo e in tal modo il giudizio sulla bontà dell'algoritmo è strettamente legato al gusto personale dell'osservatore. Si sono sviluppate in questa tesi due nuove misure di qualità, ovvero il tool basato sulla matrice delle co-occorrenze e la misura di contrasto locale. Entrambe prendono come riferimento l'immagine ad alta dinamica e giudicano buono un algoritmo che dà in output un'immagine, seppure a bassa dinamica, ma con caratteristiche simili. Il tool che utilizza la matrice delle co-occorrenze descrive la distribuzione spaziale delle immagini ad alta e bassa dinamica per mezzo di una rappresentazione visiva e di alcune feature numeriche. La seconda misura di qualità è invece basata unicamente sulla caratteristica di contrasto locale. Sulla base del contrasto locale, infatti, si possono separare i contributi dovuti rispettivamente a rumore, dettagli e parti omogenee dell'immagine. Un algoritmo che è in grado di migliorare il contrasto in particolare nella parte contenente i dettagli ha buone prestazioni secondo tale misura. Un terzo metodo per misurare le prestazioni di algoritmi di miglioramento dell'immagine è quello di mettere a punto un *setup* sperimentale per acquisire una stessa scena variando le condizioni di illuminazione. La scena acquisita con le migliori condizioni di illuminazione viene tenuta come immagine di riferimento, mentre le immagini acquisite con illuminazione scarsa vengono elaborate dagli algoritmi di cui si vuole valutare la qualità. Un algoritmo risulta quindi tanto migliore quanto più è in grado di dare in output immagini a bassa dinamica che hanno una bassa distanza dall'immagine di riferimento.

In conclusione i principali contributi originali della presente tesi nel campo del miglioramento di immagini e video sono i seguenti:

1. l'algoritmo LARS è stato migliorato ed ottimizzato per le applicazioni di videosorveglianza e aiuto alla guida (ed è stata proposta un'implementazione su FPGA);
2. sono state sviluppate tre nuove misure oggettive di qualità per stabilire le prestazioni di algoritmi di riduzione della dinamica.