

È l'ora della videosorveglianza basata su FPGA

Kambiz Khalilian,
Strategic marketing manager
Lattice Semiconductor

Le preoccupazioni crescenti per la sicurezza hanno costretto i governi e le istituzioni a effettuare notevoli investimenti in apparecchiature per la sorveglianza e per la sicurezza. Inoltre, le innovazioni tecnologiche nell'elaborazione delle immagini e dei video hanno rivoluzionato l'industria della videosorveglianza, che non è limitata alla sicurezza, ma include altri settori bancario, dei trasporti, dell'istruzione, della vendita al dettaglio, della sanità e dei videogiochi

Secundo ABI Research, i ricavi complessivi relativi al mercato della videosorveglianza aumenteranno da 16 miliardi di dollari nel 2008 a 29 miliardi di dollari nel 2015, con un CAGR del 9%. La videosorveglianza ha subito una trasformazione dalle videocamere e dai videoregistratori analogici con definizione standard, alle videocamere e ai video registratori digitali ad alta definizione con risoluzione dell'ordine dei megapixel, fino alle videocamere IP connesse in rete che trasmettono video in tempo reale via

Ethernet, per giungere ai video registratori di rete (NVR) basati su cloud. Anziché visualizzare dal vivo, registrare continuamente ed effettuare ricerche visuali sul materiale registrato, le videocamere e i videoregistratori intelligenti ora consentono la registrazione basata su eventi e l'attivazione su allarme, e le ricerche automatiche sul materiale registrato.

Questi progressi nella sicurezza e nella sorveglianza sono strettamente accoppiati con l'evoluzione nelle tecnologie

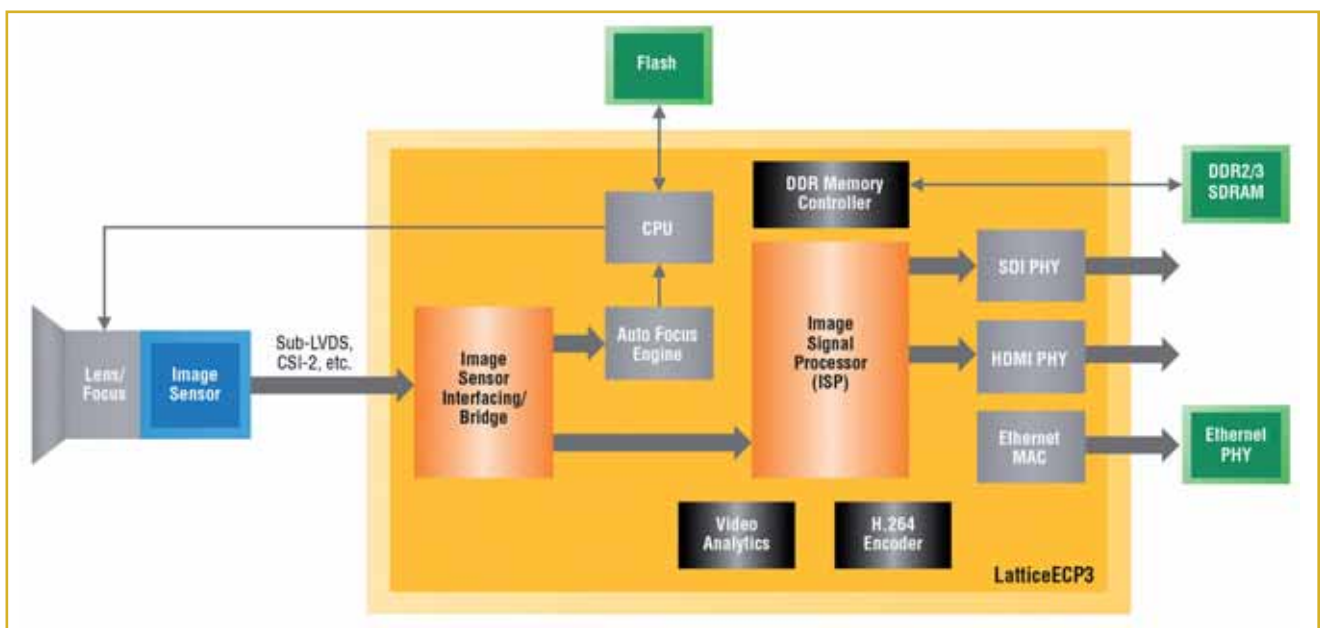


Fig. 1 - Architettura di una videocamera realizzata su FPGA

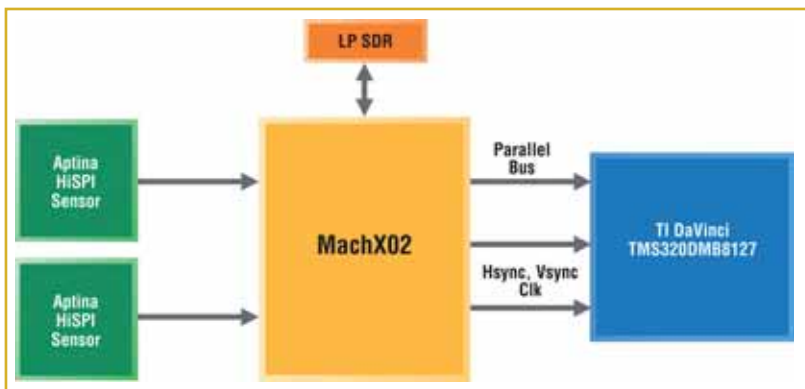


Fig. 2 - Applicazione di un bridge con l'interfaccia sensore usando un FPGA

dei sensori di immagine e nelle proprietà di elaborazione dei segnali di immagine e dei semiconduttori. Tali progressi includono le architetture di videocamera, i bridge fra sensori e le interfacce sensore, l'elaborazione dei segnali di immagine, l'elaborazione HDR (High Dynamic Range) e l'analisi del video.

Architetture delle videocamere

Una tipica videocamera consiste in un sensore di immagine, un processore di segnali di immagine e in determinato tipo di interfaccia di uscita. Oltre a questi blocchi funzionali, la videocamera potrebbe anche includere la compressione video o le funzioni di analisi del video. La figura 1 mostra un esempio di una videocamera mega pixel connessa in rete realizzata su un dispositivo FPGA LatticeECP3.

Interfacce con il sensore / Bridge con il sensore

Un processore di segnali di immagine (ISP) si interfaccia tipicamente con uno o più sensori di immagine. I sensori di immagine più comuni sono i sensori CMOS. I sensori con risoluzione e con velocità dei fotogrammi inferiori possono interfacciarsi a un ISP attraverso un'interfaccia parallela. I sensori con risoluzione e con velocità dei fotogrammi superiori richiedono interfacce ad alta velocità, che includono le interfacce MIPI, Aptina HiSPI o sub-LVDS. Gli ISP a basso costo supportano tipicamente solo un'interfaccia parallela.

Di conseguenza, è necessario usare una funzione bridge del sensore per convertire i segnali MIPI, HiSPI o sub-LVDS verso l'interfaccia parallela. È possibile usare degli FPGA a basso costo per convertire le interfacce MIPI, HiSPI o sub-LVDS verso un'interfaccia parallela. In figura 2, è usato un FPGA MachX02 di Lattice per convertire i segnali di misura HiSpi che provengono da due sensori di immagine Aptina (MT9034) in un formato paralle-

lo, che può quindi essere elaborato da un dispositivo di TI (DM8127).

Elaborazione dei segnali di immagine (ISP)

Un processore di segnali di immagine (ISP) consiste tipicamente in più algoritmi di elaborazione video, raggruppati sotto forma di una serie che esegue le operazioni di miglioramento delle immagini e le funzioni di conversione su un flusso video in ingresso.

La serie di elaborazione delle immagini potrebbe essere eseguita sotto forma di software che gira su un computer, o su un FPGA oppure su un processore dei segnali digitali (DSP) o su un prodotto standard specifico per un'applicazione (ASSP).

In figura 3 è mostrato un tipico flusso di elaborazione delle immagini. I dati grezzi del sensore sono ricevuti dal sensore di immagine CMOS. Dopo la linearizzazione, i pixel difet-

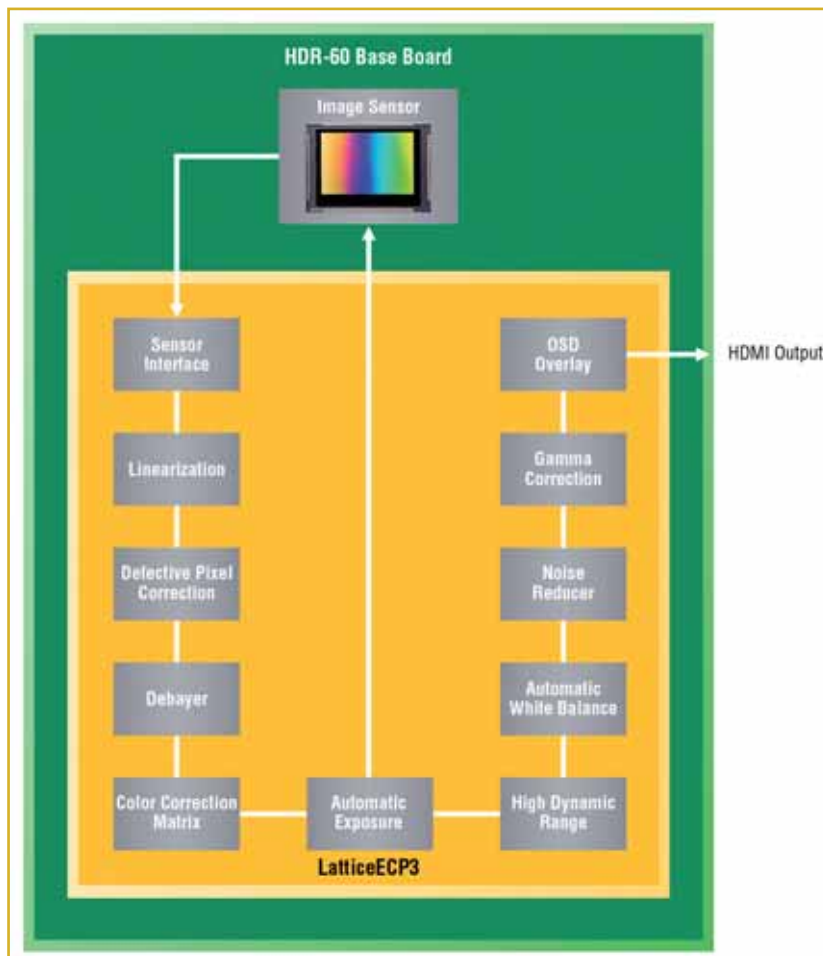


Fig. 3 - Flusso del processore dei segnali di immagine (ISP) eseguito su un FPGA



Fig. 4 - Un esempio di elaborazione HDR

tati del sensore sono corretti in base al valore dei pixel a essi adiacenti. Il sensore di immagine fornisce tipicamente un'immagine grigia, che è convertita in rosso, verde e blu usando un filtro bayer (operazione detta de-bayering). Per eliminare le interferenze fra i pixel rossi, verdi e blu, viene usata una matrice di correzione del colore (CCM).

Il blocco di auto esposizione (AE) regola automaticamente il tempo di esposizione per compensare le condizioni di illuminazione variabili. Gli algoritmi HDR migliorano il contrasto nelle regioni chiare e scure di un'immagine (gli algoritmi HDR saranno spiegati in maggiore dettaglio nella prossima sezione). Il bilanciamento automatico del bianco (AWB) effettua delle correzioni a tutti i colori in base a un punto di riferimento bianco nella scena. Un riduttore del rumore rimuove il rumore spesso visibile come punti casuali di colori ovviamente errati in un'area altrimenti colorata uniformemente. Il rumore aumenta con

la temperatura e con i tempi di esposizione. La correzione della gamma ridistribuisce i livelli di tonalità nativi della videocamera che sono percettivamente uniformi e visibili all'occhio umano, facendo quindi un uso più efficiente di una determinata profondità dei bit.

Una funzione di sovrapposizione consente di sovrapporre testo e grafica a un video per visualizzare un menu o le impostazioni della videocamera. Ulteriori moduli funzionali possono essere aggiunti in base ai requisiti specifici. Questi possono spaziare dagli algoritmi di compressione, come H.264 o JPEG, alle funzioni intelligenti di analisi video, come la rilevazione del movimento, la rilevazione di oggetti o il riconoscimento facciale.

Elaborazione HDR

La qualità di un'immagine dipende da vari fattori, inclusa l'illuminazione. In base alle condizioni di illuminazione,

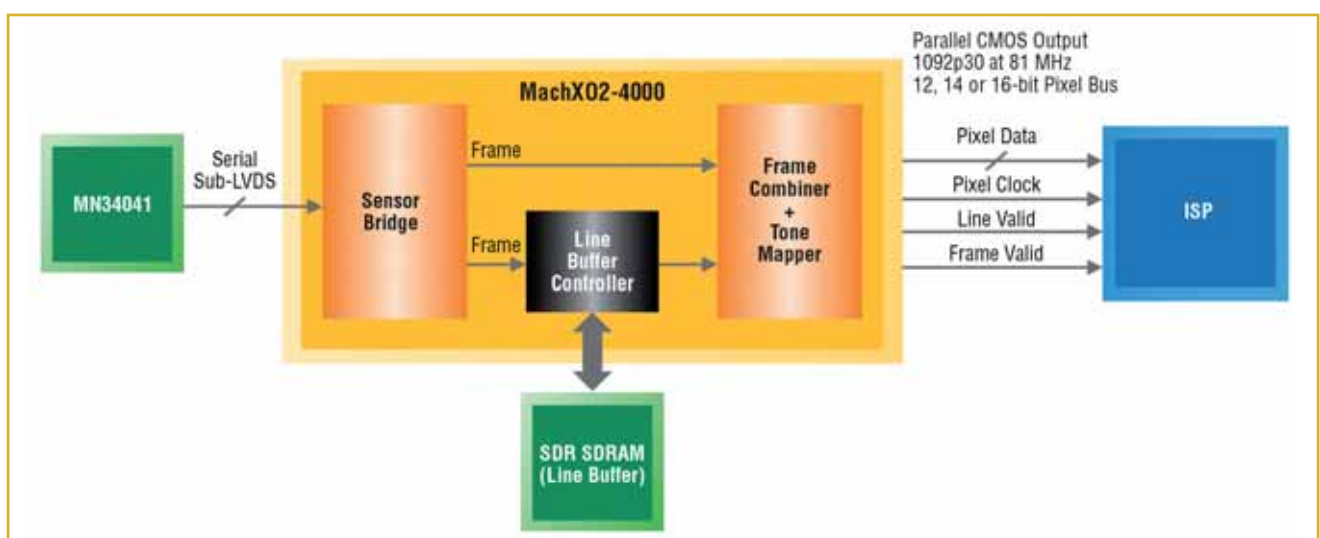


Fig. 5 - Collegamento di interfaccia ed elaborazione HDR basata su FPGA

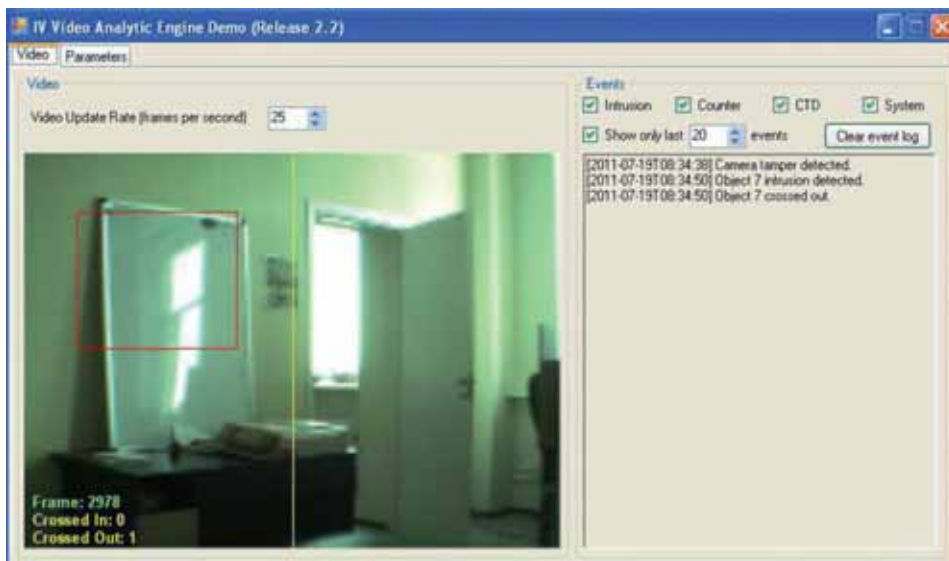


Fig. 6 - Schermata di una demo di analisi video

alcune aree di una scena potrebbero apparire troppo scure e alcune troppo luminose. I dettagli che mancano potrebbero essere molto importanti in applicazioni dove è necessario prendere decisioni critiche.

L'intervallo dinamico (chiamato anche massimo contrasto di un'immagine) descrive il rapporto fra il punto più luminoso di un'immagine e il punto più scuro della stessa immagine. L'occhio umano è in grado di vedere i dettagli in entrambe queste aree, ma gran parte dei sensori di immagine presentano un intervallo dinamico limitato e non sono in grado di catturare i dettagli in entrambe le aree della stessa immagine.

Allo scopo di rendere visibili i dettagli nelle aree a elevato contrasto, è necessario catturare più immagini della stessa scena con tempi di esposizione diversi. È richiesto un tempo di esposizione breve per catturare immagini molto luminose ed è necessario un tempo di esposizione elevato per catturare le aree scure.

Immagini con tempi di esposizione diversi possono essere combinate in un'unica immagine, aumentando l'intervallo dinamico dell'immagine stessa. La cattura di una scena con diversi tempi di esposizione e la combinazione di queste immagini in un'unica immagine è detta elaborazione a elevato intervallo dinamico (HDR), nota anche come elaborazione ad ampio intervallo dinamico (WDR).

Per l'elaborazione HDR, vengono usati degli algoritmi di mappatura delle tonalità globali e locali per creare una singola immagine con un intervallo dinamico esteso, la quale collega i dettagli nelle aree a elevato contrasto della stessa immagine. L'intervallo dinamico è spesso espresso in dB. Una buona immagine HDR presenta un intervallo dinamico di 90 o superiore. Per esempio, il sensore HDR

Aptina 720p (MT9M034) offre un intervallo dinamico di 120 dB.

Gli algoritmi di elaborazione HDR sono a intensità di calcolo molto elevata, specialmente quando si usano sensori megapixel a velocità di immagine superiori. Queste richiedono un'enorme quantità di potenza di elaborazione, che spesso non è ottenibile da un processore di segnali di immagine (ISP) comunemente disponibile in commercio.

La realizzazione di tali algoritmi complessi e ad alta intensità di risorse è effettuata al meglio usando un FPGA. Lo stesso dispositivo FPGA può essere usato per

l'esecuzione di ulteriori algoritmi di elaborazione delle immagini.

La figura 5 mostra un'esecuzione dell'algoritmo HDR con il sensore di immagine MN34041 di Panasonic. Due fotogrammi da 1080p con diversi tempi di esposizione sono catturati a 60 fotogrammi al secondo (fps). I due fotogrammi catturati sono sincronizzati facendo passare uno dei fotogrammi attraverso un buffer di linea. Dopo la sincronizzazione, i due fotogrammi sono combinati ed elaborati in un unico fotogramma da 1080p a 30 fotogrammi al secondo. Nell'esempio qui sopra, l'interfacciamento di sensore, il controllore del buffer di linea, il combinatore di fotogrammi e gli algoritmi di mappatura delle tonalità sono realizzati su un singolo FPGA MachXO2 di Lattice. Il fotogramma elaborato viene quindi passato a un processore di segnali di immagine per una successiva elaborazione.

Analisi del video

L'analisi del video è la capacità di analizzare automaticamente il video per individuare e determinare eventi. L'analisi video è usata in un'ampia gamma di applicazioni, inclusa la sicurezza e la sorveglianza, e in campo automotive, del controllo del traffico e nella vendita al dettaglio. La rilevazione del movimento è uno degli algoritmi di analisi del video più popolari, in cui il movimento viene individuato relativamente a una scena fissa di sfondo. Intellivision Inc. ha sviluppato una serie completa di analisi video intelligenti e l'ha trasportata nel kit di sviluppo per videocamere HDR-60 basate su FPGA Lattice ECP3, consentendo la progettazione di videocamere intelligenti per la sicurezza. La serie include la rilevazione intelligente del movimento su video, l'individuazione intelligente di intru-

sioni e la rilevazione di manomissioni alla videocamera. La rilevazione intelligente del movimento su video è in grado di differenziare fra un animale e una persona, inviando allarmi solo quando è necessario. L'individuazione intelligente di intrusioni invia allarmi quando qualcuno apre una porta o entra in uno spazio delimitato. Con la funzione di rilevazione di manomissioni alla videocamera, è possibile essere allertati e agire di conseguenza se una videocamera è stata spostata, bloccata o messa fuori fuoco.

La figura 6 mostra l'interfaccia di controllo basata su Ethernet per la demo di analisi video basata sul kit di sviluppo per videocamere HDR-60.

La linea gialla rappresenta la demo del contatore e l'utente può posizionarla ovunque sullo schermo in senso orizzontale, verticale o diagonale. Se un oggetto attraversa la linea, allora tale oggetto che passa da una parte all'altra dell'immagine viene contato. Questa funzionalità può essere usata per contare il numero delle persone che entrano o escono da un luogo. Il riquadro rosso rappresenta un'area ristretta e le dimensioni e il posizionamento di tale riquadro possono essere programmati per rispondere ai requisiti di una determinata applicazione. Una volta che un oggetto entra nella scena, esso viene rilevato. Se esso entra nell'area ristretta rappresentata dal riquadro rosso, viene impostato

un allarme. Il rilevamento e l'allarme automatici possono sostituire costosi monitoraggi dal vivo delle immagini delle videocamere. Una terza funzionalità, che non è visualizzata in figura 6, è la rilevazione delle manomissioni alla videocamera. Anche quando la videocamera viene coperta o spostata verrà inviato un segnale di allarme.

Gli FPGA favoriranno lo sviluppo di applicazioni complesse

Continua a crescere la rete di sensori di immagine con risoluzioni superiori e con una maggiore velocità dei fotogrammi.

La combinazione di questi sensori di immagine con algoritmi avanzati di elaborazione dei segnali di immagine, come l'HDR e l'analisi video intelligente, stanno schiudendo nuove aree di applicazione, ancora non immaginate. I processori esistenti disponibili in commercio non possono più stare al passo con i requisiti di prestazioni delle nuove applicazioni complesse. I dispositivi FPGA programmabili costituiscono la scelta perfetta per l'interfacciamento simultaneo con più sensori di immagine ad alta risoluzione, unendoli insieme in formati panoramici o sovrapposti ed eseguendo gli algoritmi HDR e di analisi intelligente del video sullo stesso chip. ■

The power of brain

www.fieramilanomedia.it

