

Mario Vento^(*)

La motion detection è video analisi intelligente?

(*) CEO di A.I. Tech srl

Motion detection è video analisi intelligente? E prima ancora: cosa significa intelligenza?

Quando si parla di video analisi *intelligente* ci si avvicina ad un territorio i cui confini sono vaghi e imprecisi: sarà che il termine *intelligente* è abusato, che sul suo significato non vi è un'unitarietà di interpretazione, che faccia scena inserirlo per aggettivare il nome di un prodotto, che un po' di intelligenza non guasti mai. Quindi cosa significa intelligenza?

Dal punto di vista tecnico, mi piace dare la definizione che offro ai discenti del corso di Visione e Modelli per la Visione Artificiale da me tenuto all'Università degli Studi di Salerno. «Se desideriamo appurare se un software sia intelligente e contiamo di poter applicare il test di Turing, ci fermiamo prima di partire!». Per chi non lo sapesse, in via semplificata, il test di Turing ha successo per un software allorché, mettendo dietro una tenda un essere umano e un software, questi risultino indistinguibili a un utente che porge domande e riceve risposta da uno o dall'altro. Affascinante? Coinvolgente? Certo, ma l'utilità pratica del test è decisamente limitata, e non è difficile comprenderlo. E ancora «una video analisi è intelligente quando il relativo software è realizzato con metodologie non procedurali». Paroloni da accademia? Forse sì, ma non vi spavento. I software moderni usano tecniche risolutive che, mediante reti neurali o altre diavolerie informatiche, apprendono da esempi. Un attimo di pazienza e vi offro un esempio che spero possa farci uscire dal vago. Se desideriamo imparare a riconoscere funghi mangerecci da quelli velenosi possiamo procedere in due modi: un esperto ci insegna le proprietà da osservare per capire, passo dopo passo, se un fungo di identità ignota sia velenoso o meno - questo modo di procedere è quello che è realizzato nel software tradizionale, che è guidato da un algoritmo. Questo altro non è che una codifica, adeguata ad un calcolatore, della conoscenza dell'esperto che ci ha messo a disposizione il procedimento per distinguere, nelle suddette categorie, i funghi. Un metodo alternativo è invece quello di osservare tanti funghi, sia buoni che velenosi, e conoscerne la categoria, di imparare nel tempo a distinguerli. «Learning from examples» è il modello di software intelligente. In questo caso non vi è conoscenza trasferita da un esperto e codificata in un programma, bensì un software più evoluto (e quindi detto intelligente) che impara a risolvere problemi, valutando esempi di problemi già risolti. Questo modello, praticabile solo da qualche anno, successivamente all'avvento di calcolatori potenti, sta conquistando giorno dopo giorno fette di mercato.

Ma attenti agli abusi: se l'aggettivo *intelligente* non è affiancato dall'uso di questi modelli di software, ci troviamo di fronte ad un abuso. E ora, dopo questa interessante prolusione tecnica, con spunti di spicciola filosofia, torniamo alla domanda sottesa dal titolo: la motion detection è video analisi intelligente? Sapendo cosa si intende per «intelligente», proviamo a rispondere, ma preventivamente comprendiamo a che serve e come opera un algoritmo di **motion detection**.

A COSA SERVE?

Un tale algoritmo ha l'obiettivo di rilevare se all'interno della scena ci sono oggetti in movimento, come persone, auto o altro. In tal caso genera un allarme usato per attivare un'azione, che può consistere nel notificare un SMS di avviso su un telefonino o una e-mail magari corredata con alcuni fotogrammi selezionati tra quelli che hanno generato l'allarme; in tal modo il proprietario del sistema può verificare (anche da lontano) cosa sta accadendo. La maggior parte delle telecamere digitali sono corredate in fabbrica di un algoritmo di motion detection di cui si può configurare la sensibilità, ovvero determinare di quanto deve variare la scena per generare l'allarme.

COME FUNZIONA?

La maggior parte degli algoritmi disponibili sulle telecamere sono molto semplici e operano su un principio elementare: calcolano su tutta l'immagine la differenza tra il valore dei pixel omologhi tra frame successivi. In assenza di oggetti in movimento, tutti i pixel dell'immagine non cambiano valore (teoricamente!) e quindi la somma di tutte le variazioni è nulla. Quando una persona o un oggetto entra nella scena, i pixel ad esso associati nel frame precedente hanno il valore dello sfondo e in quello corrente il colore della persona; si genera quindi per tutti questi pixel una variazione la cui somma viene confrontata con una soglia di sensibilità. A titolo di esempio si veda la **Figura 1** (nella pagina successiva). È quindi importante evidenziare che l'algoritmo non *ricosce* le persone o gli oggetti e di questi deduce il movimento, come invece fanno gli algoritmi di video analisi intelligente: il moto è desunto dal fatto che ci sono state variazioni di valori di pixel in quantità tale da presumere che ci sia stato un cambiamento. Insomma si tratta di un algoritmo *cieco*!

CHE PROBLEMI CREA?

Il semplice principio di funzionamento determina diversi problemi. Anzitutto bisogna fissare la soglia sulla base della dimensione presunta del soggetto o dell'oggetto che si vuole rilevare, in maniera tale che in sua presenza scatti l'allarme. Per come funziona l'algoritmo questo non dipende solo dalla dimensione: se entra un soggetto vestito di chiaro su sfondo mediamente scuro, a parità di dimensione, esso crea una differenza comples-

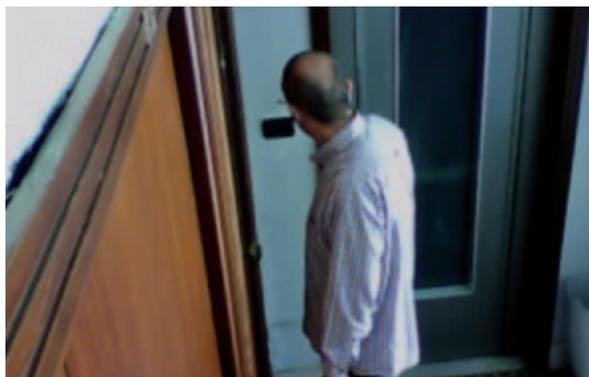
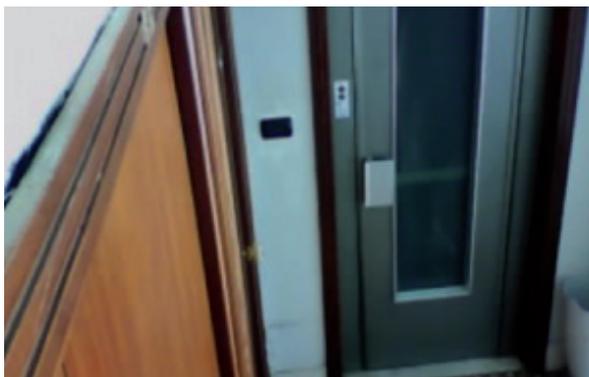


Figura 1: La telecamera (in modalità day) inquadra la porta su di un pianerottolo. L'ingresso di una persona determina correttamente l'allarme del motion detector. Si noti che la persona in questa inquadratura occupa circa 1/7 dell'area totale.

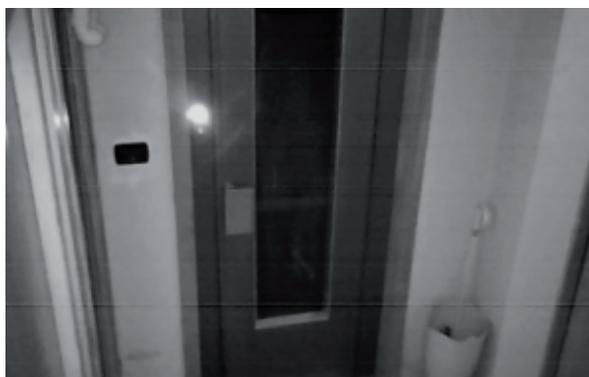
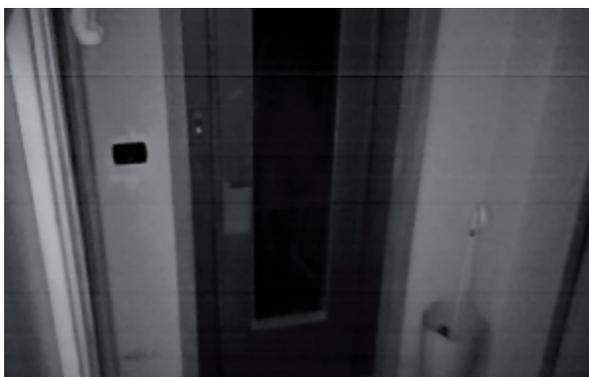


Figura 2: La telecamera in modalità night. Una persona, in un piano diverso da quello sorvegliato, accende le luci della scala e chiama l'ascensore. La differenza di luminosità complessiva della scena genera un falso allarme.

siva molto più alta di un soggetto che veste di grigio. Potrebbe quindi accadere che l'algoritmo non genera un allarme (Falso Negativo) quando il soggetto è più piccolo di quello ipotizzato o quando il suo colore è poco diverso dallo sfondo. Il problema più grave è quello relativo alla presenza di molti Falsi Positivi. Se si verifica un cambio di luminosità cambiano di conseguenza e contemporaneamente tutti i pixel, anche se non è presente alcun soggetto nella scena. Il fatto che tutti i pixel cambino valore, sebbene non di tantissimo, è tale che la differenza sia grande, cioè da superare la soglia di allarme. Un esempio è mostrato in **Figura 2**.

È INTELLIGENTE?

È evidente che un algoritmo di motion detection effettua un'analisi dell'immagine - e quindi può considerarsi sul piano teorico un algoritmo di video analisi - ma la semplicità delle elaborazioni effettuate lo colloca al di fuori della categoria dei sistemi di video analisi intelli-

gente. La sua natura fa sì che l'impiego consigliato non sia quello relativo alla generazione di allarmi, in quanto se ne otterrebbe una scarsa affidabilità in relazione alla notifica di molti falsi positivi, ma piuttosto di comandare la registrazione: basandosi sull'evidenza che è inutile registrare se non ci sono variazioni nella scena, il motion detector avvia la registrazione quando rileva una variazione e la interrompe quando si ritorna in una soluzione di stabilità. In questo caso l'effetto della generazione erronea di una variazione (causata ad esempio da un cambio di luminosità) ha un effetto trascurabile, quello di aver avviato inutilmente una registrazione per pochi secondi. L'eliminazione di questi effetti negativi richiede quindi un algoritmo che sia capace effettivamente di riconoscere ciò che è presente nella scena, quali persone, oggetti o animali, e di analizzarne il movimento, annullando quindi ogni problematica legata a variazioni di luminosità; in tal modo si potranno addirittura discernere situazioni complesse come una porta che si apre per il vento senza che vi siano persone presenti o simili, come richiede un'applicazione professionale di mercato.⁽¹⁾

⁽¹⁾ Come il prodotto A.I. Intrusion al link <http://www.aitech-solutions.eu/ai-intrusion>