

Alessia Saggese^(*)

Capire la traiettoria degli oggetti in movimento: gli algoritmi di tracking

Il rilevamento di oggetti in movimento è una tappa fondamentale per realizzare un buon algoritmo di video analisi. Ma è evidente che da solo non basta: è possibile verificare se una persona ha un comportamento anomalo limitandosi a rilevarne solo la presenza? O ancora, è possibile contare il numero di persone che entrano o escono da un negozio senza conoscere le traiettorie seguite? Scopriamolo insieme.

Il rilevamento di comportamenti anomali o il conteggio di persone sono esempi semplici che lasciano intravedere l'esigenza di introdurre ulteriore "intelligenza" negli algoritmi di video analisi, finalizzata all'estrazione della dinamica degli oggetti presenti nella scena. In altre parole, per analizzare un video abbiamo bisogno di estrarre in qualche modo la traiettoria degli oggetti, siano essi persone o veicoli, nonché alcuni dati ad essa associati come la velocità istantanea: gli algoritmi che assolvono a tale compito sono detti "di tracking".

A CHE SERVE

L'importanza del tracking a questo punto dovrebbe essere così chiara da immaginare che ogni algoritmo di video analisi ne abbia necessità. Non ne siete ancora convinti? E allora immaginiamo di essere vicino ad uno sportello bancomat. Un comportamento normale potrebbe essere quello di una persona che è ferma (in attesa di effettuare il prelievo), poi si incammina verso lo sportello, si ferma di nuovo (per prelevare) ed infine si allontana camminando. Cosa ben diversa è invece il caso di una persona che si avvicina allo sportello, si ferma e poi inizia a correre. Cosa ci fa venire in mente?

^(*) Project Manager di A.I. Tech <http://www.aitech-solutions.eu>

Usando traiettorie e velocità stiamo concludendo che è altamente probabile che il comportamento della persona sia anomalo e che ci si trovi di fronte a un ladro. Senza esserne pienamente consapevoli, abbiamo appena usato anche noi un algoritmo di tracking.

COME FUNZIONA

Il tracking è un problema tutt'altro che semplice. L'idea generale è quella riassunta nella **Figura 1**: sulla sinistra vediamo gli oggetti con la relativa traiettoria, rilevati dall'algoritmo di tracking al fotogramma precedente. Sulla destra vediamo invece tutto ciò che è in movimento al fotogramma corrente e che è stato rilevato dall'algoritmo di "moving object detection", trattato nel precedente numero di a&s Italy e approfondito sul web⁽¹⁾. L'obiettivo di un algoritmo di tracking a questo punto sarà effettuare la migliore associazione tra gli oggetti inseguiti fino al fotogramma precedente e tutto ciò che è stato rilevato al fotogramma corrente. Più in generale, il tracking ha il compito di associare le posizioni degli oggetti che si corrispondono in frame successivi, ottenendo così fotogramma per fotogramma la traiettoria da essi disegnata. Tale associazione può essere basata su differenti proprietà. La cosa più semplice è quella di valutare la **posizione**: un oggetto viene associato a quello che nel frame precedente occupava la posizione più vicina. Meccanismi più sofisticati prevedono l'introduzione della **predizione del movimento** degli oggetti grazie alla valutazione della loro storia passata: in altre parole, se un oggetto sta camminando in una certa direzione con una certa velocità, è molto probabile che al fotogramma

successivo continui a camminare nella stessa direzione e mantenendo una velocità costante. Un'altra proprietà spesso utilizzata è il **colore**: se due persone indossano rispettivamente una t-shirt rossa con jeans blue, e una tuta nera, l'analisi del colore sicuramente aiuta. Pensiamo ad esempio ad una situazione in cui i relativi oggetti sono molto prossimi e seguono la stessa traiettoria: i precedenti criteri basati sulla distanza e sulla storia del movimento non risulterebbero sufficientemente caratterizzanti. Il colore è in questo caso l'unica informazione su cui basarsi per fare una corretta associazione. Sulla base di queste considerazioni, molti dei moderni algoritmi di tracking sfruttano la combinazione tra l'informazione basata sul colore e quella basata sulla posizione al fine di aumentare l'affidabilità del risultato. È evidente che queste, sebbene siano le più comuni, sono solo alcune delle proprietà che un algoritmo di tracking può impiegare: ne sono esempio l'**altezza** di una persona espressa in pixel, o magari l'altezza reale, espressa cioè in centimetri, estratta attraverso tecniche di "Inverse Perspective Mapping" che, grazie ad una preliminare calibrazione della telecamera, sono in grado di valutare le dimensioni reali di una persona a partire dalla loro posizione e la dimensione espressa in pixel.

PROBLEMI COMUNI

A questo punto ci si potrebbe chiedere perché mai tanti scienziati siano impegnati sul problema del tracking, visto che la soluzione parrebbe tanto semplice. C'è da dire che questa non vuole essere una descrizione esaustiva delle problematiche associate al tracking (un inte-

⁽¹⁾ <http://www.aitech-solutions.eu/category/pils/>

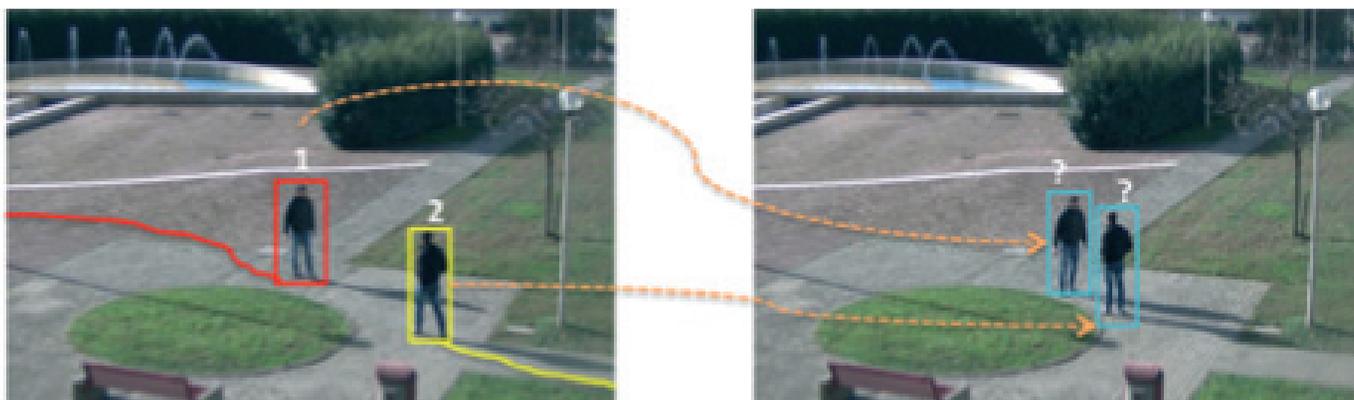


Figura 1: Un algoritmo di tracking all'opera: sulla sinistra gli oggetti rilevati al frame precedente e la relativa traiettoria; sulla destra gli oggetti in movimento rilevati dall'algoritmo di "moving object detection", ancora privi di identità. L'algoritmo di tracking effettua l'associazione di tali oggetti nei frame successivi, come indicato dalle frecce.

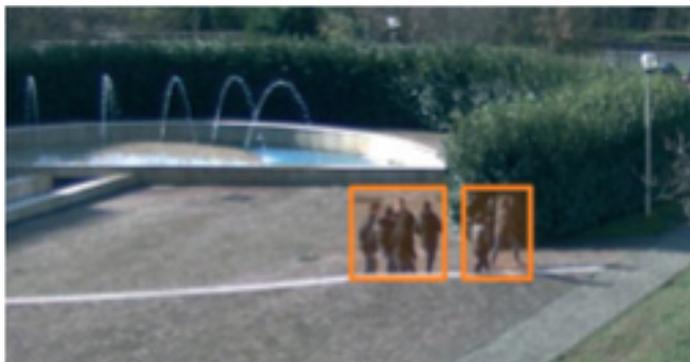


Figura 2: Il problema delle occlusioni: due persone che camminano nella stessa direzione sulla destra; sulla sinistra invece si vedono cinque persone ferme. Un buon algoritmo di tracking deve essere capace, anche in questi casi, di inseguire singolarmente le persone che si muovono nella scena.

ro numero di a&s Italy non basterebbe!), ma piuttosto una presentazione dei principi di funzionamento e delle finalità. Per questo motivo, l'unico problema che nominiamo (tra i tantissimi!) è quello legato alla presenza di **occlusioni**: la telecamera appiattisce l'immagine su un piano a due dimensioni, per cui può capitare che per periodi di tempo più o meno lunghi un oggetto in movimento sia parzialmente o completamente nascosto da altri oggetti presenti all'interno della scena. Pensiamo, ad esempio, ad un muretto o ad un palo che potrebbe impedire la vista di una persona, anche all'occhio umano, per qualche fotogramma, oppure a due persone che

camminano insieme all'interno della scena in modo tale che una delle due sia sempre nascosta agli occhi della telecamera. Ovviamente un algoritmo di tracking deve continuare ad inseguire tali oggetti anche quando non si vedono, rendendo tale compito ancora più complesso. Insomma, agli algoritmi di tracking stiamo chiedendo di fare una cosa che all'occhio umano non è consentita: devono inseguire anche quando non vedono!

Per meglio chiarire il concetto, un esempio è mostrato nella **Figura 2**⁽²⁾.

CONCLUSIONI

Tanto complesso quanto potente. Questo è il messaggio conclusivo di questo articolo. Abbiamo visto come gli algoritmi di tracking debbano spesso risolvere problemi tutt'altro che banali anche per l'occhio umano (il problema delle occlusioni, ad esempio), ma risultano fondamentali per consentire il corretto funzionamento della maggior parte degli algoritmi di video analisi oggi presenti sul mercato⁽³⁾. Ogni buon algoritmo di video analisi che si rispetti è basato su un algoritmo di tracking, e non tutti gli algoritmi di tracking funzionano allo stesso modo. Per cui, quando scegli un prodotto di video analisi, scegli sempre quello dotato del migliore algoritmo di tracking presente sul mercato.

⁽²⁾ Ecco un video con maggiori dettagli sul funzionamento dell'algoritmo di tracking di A.I. Tech in caso di occlusioni, algoritmo che si è classificato nelle prime posizioni in una competizione mondiale. <https://www.youtube.com/watch?v=FTCMFeEGW78%20>

⁽³⁾ Per capire meglio le potenzialità degli algoritmi di tracking, esplorate alcuni prodotti di A.I. Tech come AI-PEOPLE (<http://www.aitech-solutions.eu/ai-people/>), che effettua il conteggio di persone che attraversano una regione in una data direzione, usando il tracking per comprendere la direzione del movimento delle persone. Oppure AI-INTRUSION (<http://www.aitech-solutions.eu/ai-intrusion/>), che verifica la presenza di una persona che "si intrufola" in un'area proibita. Se il sistema rivelasse solo la presenza di una persona all'interno dell'area, ad esempio utilizzando i più conosciuti algoritmi di "motion detection", un elevato numero di falsi allarmi sarebbero inviati al sistema (addirittura, un allarme diverso per ciascun fotogramma, poiché il sistema non ha alcuna informazione sulla "storia" della scena, non avendo informazioni su cosa sia successo nei fotogrammi precedenti, su quanto tempo una persona abbia trascorso nella scena e così via). Grazie al tracking, il sistema è capace di identificare se l'allarme sia stato invece generato da una sola persona e quindi da una sola intrusione, generando così un singolo allarme.



Figura 3: L'algoritmo di tracking di A.I.Tech all'opera in due differenti scenari, uno outdoor e uno indoor: le traiettorie, ciascuna con un colore diverso, sono disegnate sovrapposte all'immagine.