

3.2 Parco – Visione Artificiale per la sicurezza di persone in luoghi pubblici

R. Cucchiara, A. Prati

ImageLab – Università di Modena e Reggio Emilia

Sappiamo bene come la qualità della vita in una città non significhi solo benessere economico, ma anche sicurezza sociale e fruibilità delle numerose opportunità che la città offre: parchi, giardini per l'infanzia, piazze e luoghi pubblici a disposizione del cittadino. Purtroppo però, sia per l'aumentata criminalità comune, sia per l'insicurezza dovuta al diffondersi del terrorismo nelle città di tutto il mondo, i problemi di ordine pubblico sono molto sentiti dalla popolazione. Il rapporto del Sole 24Ore del 2004 riporta che le città Emiliane ne sono particolarmente sensibili. Malgrado il tenore di vita sia estremamente elevato nella nostra regione, città come Reggio Emilia, Bologna e Modena si sono trovate rispettivamente al 66°, 75° e 98° posto su 103 per quel che riguarda la percezione della gravità attribuita ai problemi di criminalità e ordine pubblico.

In risposta a queste esigenze Reggio Emilia, come tante altre città si sta adeguando con una fitta rete di telecamere collegate alla polizia. Gli attuali sistemi di sorveglianza sono normalmente composti da telecamere fisse o brandeggiabili, piazzate per avere un buon campo di vista (in inglese Field of View, FoV) su zone sensibili potenzialmente pericolose, collegate attraverso una rete dedicata alla centrale di polizia. Là le immagini possono essere visualizzate su monitor o memorizzate se necessario, in ottemperanza alle leggi del Decalogo della Videosorveglianza sulla privacy (D.lgs. n.196/2003). Tali sistemi spesso funzionano da deterrente ma non possono essere efficaci in tempo reale come prevenzione al crimine, dato che la contemporanea attenta osservazione di tutti le centinaia i flussi video 24 ore su 24, non e' praticabile. E' ormai assodato come solo attraverso strumenti di analisi automatici o semiautomatici si potranno ottenere un contributo preventivi e non solo un supporto alla indagine a posteriori. Ciò rende indispensabile la ricerca in Visione Artificiale e Pattern Recognition. La seconda disciplina si occupa di tecniche di riconoscimento e classificazione di forme, e di dati misurati, spesso mediante tecniche probabilistiche e basate sull'incertezza. La prima si occupa di tutte le metodologie per permettere al calcolatore di emulare comportamenti percettivi e visivi tipici dell'essere umano, quale l'individuazione di oggetti in movimento nella scena, l'estrazione di informazioni visuali (colori, forme), la ricostruzione dello spazio tridimensionale a partire da immagini.

Le tecniche sviluppate in questo ambito permettono di estrarre automaticamente dati visuali e di inferire una conoscenza su quanto osservato come ad esempio la presenza di persone e i loro spostamenti. Esse permettono di fornire un ausilio al riconoscimento di potenziali situazioni di pericolo o focalizzare l'attenzione degli operatori nel centro di controllo su situazioni fuori dalla norma. Pensiamo alla presenza di persone a terra, o persone in fuga sospetta, persone stazionarie in zone normalmente di passaggio, ecc. Questo diretto alla popolazione. Invece i video catturati da zone della città e direttamente disponibili ad esempio via Web potrebbero migliorare la sensazione di sicurezza e migliorare la fruibilità di spazi pubblici. Queste possibile impiego di Webcam e Telecamere della città si scontra però con le giuste richieste di privacy dei cittadini ripresi.

Fatte queste premesse, il dimostratore "Parco" si è posto alcuni obiettivi:

- ✓ realizzazione automatica di viste allargate (note come panoramiche o "mosaici") dei parchi cittadini e la loro sua diffusione tramite Internet ai cittadini
- ✓ realizzazione di un sistema di video-sorveglianza automatica per l'individuazione e l'inseguimento ("tracking) di persone nella scena.

Il primo obiettivo ha sfruttato telecamere dome brandeggiabili come quella installata ad-hoc nel Parco del Popolo. Sistemi di creazione di mosaici da video sono già disponibili sul mercato. Nel caso specifico però sono state impiegate innovative tecniche di visione artificiale per poter ricostruire la scena in ogni condizione ed in tempo-reale, pur non conoscendo il movimento della telecamera, capace di muoversi a 360° anche manovrata manualmente da agenti dell'ordine. Le tecniche sviluppate sono basate sul calcolo automatico delle corrispondenze tra due immagini consecutive ottenute mentre la telecamera è in movimento e sulla rettificazione delle immagini con un modello affine. Figura 3.2.1 mostra un esempio.



Figura 3.2.1. Mosaico creato automaticamente

Questa panoramica del parco potrà essere messa a disposizione dei cittadini tramite una pagina Internet ad-hoc, ma solo una volta che le informazioni sensibili sulle persone saranno state rimosse. Il sistema e' direttamente impiegabile qualora la telecamera adotti un livello di zoom come in figura in cui l'identità dell'individuo non può essere svelata. Il sistema comunque può adattarsi a zoom diversi: in alcuni casi dovrà essere possibile oscurare l'identità delle persone.

Questo problema e' cruciale sia per le telecamere brandeggiabili che fisse. Avendo a disposizione un sistema capace di individuare nella scena ("segmentare") le persone, farne il tracking nel tempo ed estrarne l'aspetto visuale, il problema può essere risolto. Il secondo obiettivo del dimostratore e' proprio quello di dimostrare come le più avanzate ricerche in video-sorveglianza automatica possono ottenere traccia delle persone anche in ambienti complessi come un parco cittadino.

Queste tecniche costituiscono lo stato dell'arte a livello mondiale, e non erano mai state sperimentate prima di ora in ambienti fuori da laboratori o da situazioni ben strutturate. Infatti oltre al problema di modellazione teorica, e' necessario impiegare tecniche sofisticate per distinguere oggetti dallo sfondo, eliminare ombre, distinguere persone che si sovrappongono da un dato punto di vista o che sono parzialmente nascoste da oggetti della scena come alberi panchine o giochi dei bambini.

L'individuazione e segmentazione delle persone in movimento viene fatto sulla base della loro differenza rispetto allo sfondo. Il loro inseguimento nel tempo (con il calcolo della traiettoria) viene effettuato grazie a modelli probabilistici che tengono conto dell'apparenza della persona e che si sono dimostrati molto robusti alle occlusioni (molto frequenti in ambienti complessi come un parco). Le informazioni estratte dalle singole telecamere vengono sfruttate per riconoscere che la persona individuata da una telecamera e' la stessa vista da un altro punto da un'altra telecamera.

Un esempio di risultato dalle telecamere di Parco del Popolo è riportato nella Figura 3.2.1

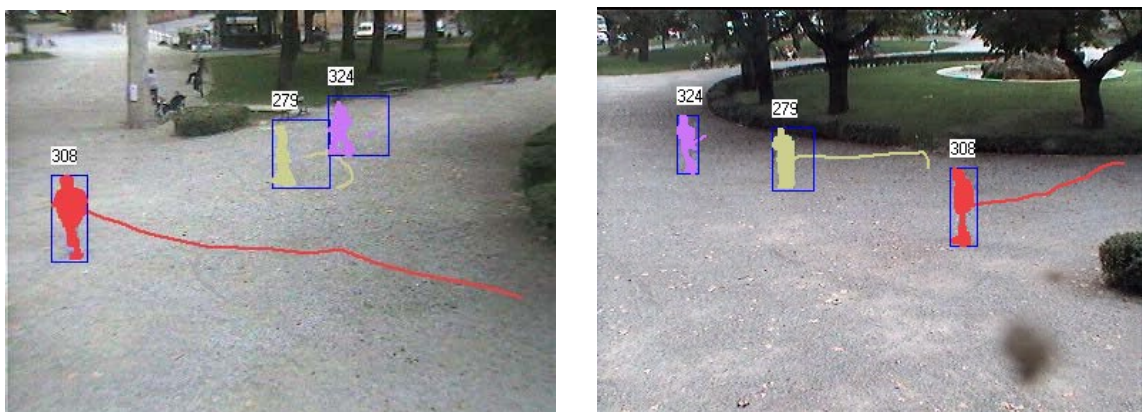


Figura 3.2.1. Esempio di individuazione ed inseguimento di persone

Avendo queste informazioni, è anche possibile oscurarle. Le immagini come quelle in figure permettono di individuare la scena, possono essere messe disponibili in Web per informazioni turistiche e per la sicurezza ma nel contempo sono rispettose della privacy dell'individuo.

Queste informazioni possono essere sfruttate per monitorare zone pericolose e di interesse: un tipico esempio di applicazione è il conteggio di persone che attraversano un'area o l'analisi delle traiettorie associate (un esempio in Figura 3.2.2).

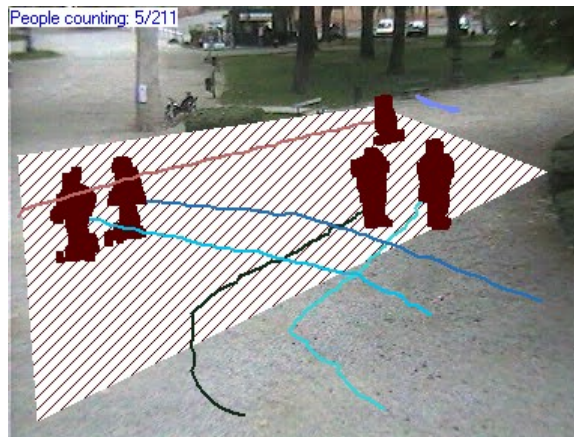


Figura 3.2.2. Esempio di conteggio in area definita

La ricerca sviluppata in questo ambito ha prodotto ottimi risultati scientifici, di eccellenza internazionale (come evidenziato nel capitolo 5); ha anche mostrato la fattibilità di nuove generazioni di sistemi di sorveglianza che crediamo possano essere alla base dell'uso futuro, più efficace ed accettabile per i cittadini, di telecamere e della video sorveglianza.