Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej Nr 36

XXIII Seminarium

ZASTOSOWANIE KOMPUTERÓW W NAUCE I TECHNICE' 2013
Oddział Gdański PTETiS
Referat nr 27

METODY ŚLEDZENIA OBIEKTÓW W ROZPROSZONYCH SYSTEMACH MONITORINGU WIDEO

Karol LISOWSKI, Andrzej CZYŻEWSKI

Katedra Systemów Multimedialnych, Wydział Elektroniki, Telekomunikacji i Informatyki, Politechnika Gdańska
tel: +48 58 347 16 36, +48 58 347 13 01 e-mail: {lisowski, andcz}@sound.eti.pg.gda.pl

Streszczenie: Systemy monitoringu wideo stały się powszechną częścią zarówno przestrzeni publicznej jak również miejsc o ograniczonym dostępie. Nadzór obszaru o dużej powierzchni wymaga rozmieszczenia wielu kamer. Skuteczna analiza przez człowieka dużej liczby obrazów wideo jest praktycznie niemożliwa. Dlatego rozwijane są metody służące do automatycznego przetwarzania wideo ukierunkowanego na analizę kontekstową. W przypadku niepokrywających się pól widzenia kamer znaczenia nabiera również reidentyfikacja obiektów w różnych kamerach. Ten referat koncentruje się na przeglądzie metod śledzenia obiektów pomiędzy kamerami. Docelowo automatyczna analiza ma ułatwić śledzenie sytuacji na dużym obszarze poprzez wskazanie strumieni wideo skojarzonych z pewnymi istotnymi zdarzeniami.

Słowa kluczowe: monitoring wizyjny, reidentyfikacja, śledzenie obiektów.

1. WPROWADZENIE

1.1. Rozwój systemów monitoringu i automatyczna analiza kontekstowa

Systemy monitoringu wideo stały się nieodłączną częścią przestrzeni publicznej oraz zamkniętych obszarów wymagających specjalnego nadzoru. Cel, który przyświeca instalacji kamer to głównie podniesienie bezpieczeństwa osób i mienia w danym miejscu. Potrzeba bezpieczeństwa spowodowała gwałtowny wzrost liczby kamer monitoringu. Szczególnym przykładem jest tutaj Wielka Brytania, w której liczba kamer monitoringu w 2012 roku przekroczyła 8 milionów. Warto zwrócić również uwagę na zdolności poznawcze człowieka, który potrafi utrzymywać wysoki stopień koncentracji jedynie przez około 20 minut [1]. Podzielność uwagi przy obserwacji obrazów z wielu kamer jednocześnie również stanowi tutaj problem [2]. Tak gwałtowny rozwój systemów monitoringu połączony z ograniczoną percepcją człowieka spowodował wzrost znaczenia automatycznej analizy wideo w systemach monitoringu wideo.

Automatyczna analiza video składa się z kilku etapów tj.:

- odejmowanie tła i detekcja obiektów,
- śledzenie obiektów w obrębie jednej kamery,
- wykrywanie zdarzeń,
- śledzenie obiektów pomiędzy kamerami.

Na każdym kroku przetwarzania do strumieni wideo przypisywane są kolejne metadane związane z co raz szerszym kontekstem sytuacji obserwowanej przez kamery. Kolejne etapy przetwarzania mogą być potraktowane jako warstwy (lub inaczej poziomy) analizy wideo (rys. 1).

Istotną fazą przetwarzania jest powiązanie ze sobą ujęć z różnych kamer zawierających ten sam obiekt. Pozwala to określić tzw. ścieżkę ruchu danego obiektu. Ma to szczególne znaczenie przede wszystkim w rozbudowanych wielokamerowych systemach monitoringu, gdzie rejestrowany jest obraz z wielu kamer przez całą dobę i w razie ewentualnej potrzeby odnalezienia danego materiału wideo należałoby przejrzeć wiele godzin nagrań.

Dodatkowo w ostatnim etapie przetwarzania można zbudować model zachowania obiektów, który jest pomocny przy reidentyfikacji obiektów.

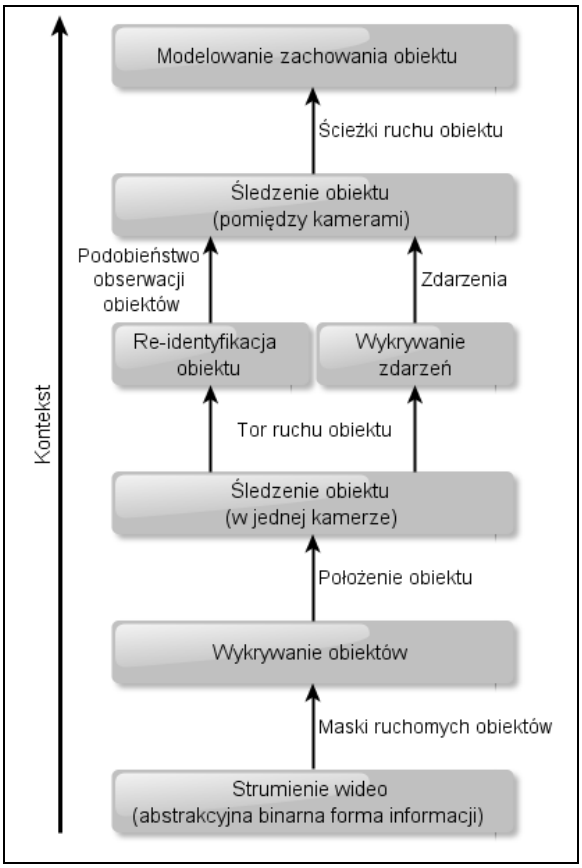
1.2. Wykrywanie obiektów

Pierwszym krokiem analizy wideo jest odróżnienie statycznego tła od obiektów, które zmieniają swoje położenie w kadrze kamery. Najprostsza metoda polega na uznaniu pierwszej ramki za tło a następnie wykrywanie różnic względem niej. Jednak jest to niepraktyczne przy założeniu, że tło może się stopniowo zmieniać. Dodatkowo nie można przyjąć, że pierwsza ramka obrazu wideo nie zawiera ruchomych obiektów. Znacznie skuteczniejsze są metody wykorzystujące modelowanie tła. Takie algorytmy mają zdolność adaptacji m.in. do zmiennych warunków oświetlenia. Znane metody detekcji obiektów to na przykład GMM (ang. *Gaussian Mixture Model*) [3] oraz Codebook [4]. Wynikiem działania takiego algorytmu są sylwetki (maski) obiektów.

1.2. Śledzenie obiektów w polu widzenia jednej kamery

Kiedy odróżnianie tła od ruchomych obiektów jest możliwe następnym problemem jest uznanie danej sylwetki jako ten sam obiekt w kolejnych ramkach obrazu. Najpowszechniej stosowanymi metodami śledzenia obiektów jest śledzenie kształtu. W systemie czasu rzeczywistego jakim jest system monitoringu wideo istotny jest kompromis pomiędzy skutecznością użytej metody śledzenia a kosztem

obliczeniowym zastosowanego algorytmu. Zadowalający stosunek skuteczności do złożoności obliczeniowej na analizie statystycznej związanej z minimalizacją błędu pomiędzy przewidywanym i faktycznym położeniem obiektu w kadrze kamery można osiągnąć przy użyciu metod adaptacyjnych. Przykładami takich algorytmów mogą być filtr Kalmana lub filtr cząsteczkowy [5].



Rys. 1. Etapy kontekstowego przetwarzania danych wideo z zaznaczeniem wyników analizy na kolejnych warstwach

1.3. Wykrywanie zdarzeń

W systemach monitoringu wideo możliwe jest wykrywanie zdarzeń obserwowanych w kamerze. Mogą to być proste zdarzenia takie jak wejście w zabroniony obszar, przekroczenie zdefiniowanej bariery lub ruch pod prąd. Jednak możliwe jest również wykrywanie bardziej złożonych zdarzeń na przykład porzucenie bagażu albo różnorakie zachowania tłumu (panika, stłoczenie przy drzwiach itp.)

2. ŚLEDZENIE OBIEKTÓW POMIĘDZY KAMERAMI

W rozproszonych systemach monitoringu wideo przy dużej liczbie kamer nie możliwe jest aby poruszający się obiekt zawsze był widoczny chociaż w jednej z kamer. Najczęściej obserwowaną sytuacją jest całkowite zniknięcie obiektu z pola widzenia kamery i pojawienie się po pewnym okresie w innej kamerze. Podstawowym zagadnieniem jest tutaj powiązanie ze sobą obserwacji z dwóch kamer i podjęcie decyzji, że to jest ten sam obiekt. Tworzenie takich powiązań pozwala otrzymać ścieżki ruchu obiektu na nadzorowanym obszarze. W takim przypadku operator systemu ma dostęp do kilku przydatnych funkcjonalności, które dostarcza automatyczna analiza wideo, takich jak możliwość odnalezienia fragmentów materiału wideo z różnych kamer powiązanych z daną obserwacją obiektu lub wykrytym zdarzeniem. Ponadto, może zaznaczyć obiekt, który będzie automatycznie śledzony. Operator może zająć się innymi zadaniami ale w każdej chwili operator może wrócić do obserwowania tego obiektu.

W celu realizacji śledzenia pomiędzy kamerami należy wykorzystać możliwie dużo przesłanek, które może dostarczyć nam system monitoringu. Wśród nich mogą być cechy fizyczne obiektu związane z jego wyglądem i cechami fizycznymi (tzw. deskryptory cech fizycznych), ale również zależności czasowo-przestrzenne występujące pomiędzy kamerami. Dodatkową przesłanką może być też sposób zachowania obiektów na obserwowanym obszarze.

2.1. Deskryptory cech fizycznych

W celu porównania dwóch obserwacji pomocny może być wygląd widzianych obiektów. Takie rozróżnienie może nastąpić przy użyciu deskryptorów wyglądu obiektu. Deskryptory obrazu można podzielić ze względu na opisywaną przez nie cechy obrazu na następujące typy:

- Deskryptory koloru [6]:
  - Histogram koloru [7];
  - DCD (ang. *Dominant Color Descriptor*) [8];
  - CLD (ang. *Color Layout Descriptor*) [9];
- Deskryptory tekstury [11]:
  - TBD (ang. *Texture Browsing Descriptor*) [12];
  - EHD (ang. *Edge Histogram Descriptor*) [13];
  - SIFT (ang. *Scale Invariant Feature Transform*) [14];
- Deskryptory kształtu:
  - Momenty geometryczne obrazu [10];
  - Momenty centralne obrazu [10];
  - Szkieletyzacja kształtu (np. String Codes, Tree Grammar) [15];
  - PGH (ang. Pairwise Geometric Histogram) [16].

Pod każdym z typów deskryptorów podano kilka przykładów deskryptorów danego typu. Standardem służącym do opisu danych multimedialnych jest MPEG-7, w którym zdefiniowana jest większość wymienionych powyżej deskryptorów. Ważną cechą deskryptorów cech fizycznych stosowanych przy re-identyfikacji obiektów jest ich odporność na zmiany oświetlenia i na różnice w ustawieniach parametrów akwizycji obrazu w kamerze (przede wszystkim chodzi o balans bieli).

2.2. Topologia sieci kamer w systemie monitoringu

Możliwość i czas przejścia pomiędzy daną parą kamer jest uzależniony od odległości w przestrzeni między kamerami i od fizycznych przeszkód znajdujących się pomiędzy polami widzenia tej pary kamer. Może zaistnieć również przypadek, że przejście pomiędzy dwiema kamerami nie jest możliwe bez pojawienia się w polu widzenia trzeciej. Do opisu tych zależności służy graf topologii. Dokładniej jest to digraf ważony, w którym wierzchołki odpowiadają kamerom, krawędzie determinują możliwość przejścia pomiędzy daną parą kamer w określonym kierunku, zaś wagi na krawędziach opisują czas takiego przejścia.

MOST WIEDZY Pobrano z mostwiedzy.pl

2.3. Model zachowania obiektów na obserwowanym obszarze

Ze względu na pewną niepewność związaną z trajektorią ruchu obiektu w miejscach nieobjętych polem widzenia systemu monitoringu należy zastosować metody predykcji jego ruchu. Takie algorytmy muszą bazować na modelu zachowania obiektów opisującym, które krawędzie w topologii sieci kamer są używane częściej niż inne oraz definiującym najprawdopodobniejsze kolejne ruchy danego obiektu. Różne struktury danych mogą zostać użyte to opisu zachowania obiektów, a także w różny sposób można je wykorzystać. Przykłady modeli zachowania obiektów i ich zastosowania zostaną opisane w dalszej części referatu.

2.4. Metody śledzenia obiektów pomiędzy kamerami

Ponieważ dane wejściowe dla algorytmów śledzenia pomiędzy kamerami są podobne, w literaturze możemy odnaleźć podejścia do tego problemu wykorzystujące w różny sposób jedną albo kilka wyżej wspomnianych przesłanek ale w różnoraki sposób. Część podejść koncentruje się na poprawie skuteczności reidentyfikacji obiektów w oparciu o deskryptory cech fizycznych. Powstają również pomysły użycia różnych metod modelowania zachowania obiektów obserwowanych przez system monitoringu wideo. Istnieją również publikacje, w których opisane są metody łączenia informacji z wielu przesłanek w celu podjęcia decyzji podczas śledzenia obiektu pomiędzy kamerami.

2.4.1 Metody wyznaczania i porównywania deskryptorów cech fizycznych

Najczęściej stosowanym deskryptorem wyglądu jest histogram kolorów [7]. Znane są również rozwiązania, w których w celu łatwiejszego porównywania histogramów kolorów są one aproksymowane rozkładem Gaussa. Ponadto, w literaturze można odnaleźć podejścia ukierunkowane na większą skuteczność decyzji o identyczności dwóch obiektów. Dotyczą one śledzenia osób i polegają one na podziale sylwetki na części (przykładowo: głowa, tors i nogi) i wyznaczeniu oddzielnego histogramu dla każdej z tych części [17].

Proponowaną i często stosowaną miarą pomiaru różnicy pomiędzy dwoma histogramami kolorów jest odległość Bhattacharyya. Dla jednowymiarowego rozkładu tę odległość opisuje następujący wzór [18]:

$$d_b(p_1, p_2) = \frac{1}{4}\frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2}\ln\left(\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2}\right) \qquad (1)$$

gdzie: d_b – odległość Bhattacharyya, p_1, p_2 – dwa rozkłady prawdopodobieństwa, σ_1, σ_2 – odchylenia standardowe tych rozkładów, μ_1, μ_2 – wartości średnie tych rozkładów

Pomimo swojej prostoty i łatwości wykorzystania w algorytmach identyfikacji histogram kolorów nie jest odpowiednim deskryptorem cech fizycznych w przypadku zmiennego oświetlenia i różnych paramentów technicznych kamer użytych w systemie monitoringu. Rozwiązanie tego problemu może być dokonane na dwa sposoby:

- wyznaczenie transformacji fotometrycznej pomiędzy parami kamer, która będzie niwelować różnice w parametrach urządzeń rejestrujących obraz i w zmianach w oświetleniu (np. pomiędzy naturalnym i sztucznym oświetleniem) [19][20];
- poszukiwanie deskryptorów odpornych na niekorzystne z punktu widzenia re-identyfikacji zmiany w wyglądzie obiektu obserwowanego przez różne kamery [21]

Badania wskazują, że wysoką odporność na niestałość wyglądu tego samego obiektu mają deskryptory SURF [22] i SIFT [14]. Opierają się one na analizie punktów charakterystycznych w sylwetce obiektu i rozróżniają również teksturę obiektu. Dzięki tym właściwościom łatwiej pozwalają rozróżnić obiekty o podobnym histogramie kolorów, lecz zupełnie różnym wyglądzie [21].

Ciekawym podejściem jest również algorytm kompensacji zmian kolorów CCN (ang. *comprehensive color normalization*) zaproponowany przez G. D. Finlaysona [23] i in. Ta metoda jest w stanie zniwelować zmiany koloru wynikające z różnego kąta padania światła na obserwowany obiekt albo będące następstwem zmiany koloru oświetlenia. Jednak algorytm CCN nie działa poprawnie w przypadku, kiedy kąt padania światła zmienia się jednocześnie ze zmianą jego koloru.

2.4.2 Metody odkrywania topologii sieci kamer

Kolejnym elementem niezbędnym do re-identyfikacji obiektów, a co za tym idzie również do powiązania ze sobą dwóch różnych obserwacji obiektu, jest użycie topologii sieci kamer. Bardzo często w metodach śledzenia osób pomiędzy kamerami zależności czasowo-przestrzenne muszą być dane *a-priori*. Jednak istnieją również metody pozwalające na odkrywanie nieznanej wcześniej topologii na podstawie analizy statystycznej wyników przetwarzania wielu godzin materiału wideo.

R. Farrell i in. [24] zaproponowali tworzenie statystycznego modelu sylwetki osoby ze wszystkich obserwacji. Następnie należało odnaleźć wśród wszystkich obserwacji obiektów takie, które posiadały pewne cechy dystynktywne pozwalające odróżnić je w skuteczny sposób od sylwetek podobnych do modelowej. Mówiąc inaczej, osoby ubrane w nietypowy sposób stanowią źródło danych niezbędnych do odkrycia topologii danej sieci.

Również K.Tieu i in. [25] przedstawili metodę odkrywania zależności czasowo-przestrzennych w oparciu o analizę statystyczną. Autorzy założyli niezmienność średniego czasu przejścia pomiędzy parami kamer oraz to, że zmiana w wyglądzie obiektów przy przejściu między tymi samymi kamerami również jest stabilna. Określanie powiązania jest oparte na pomiarze entropii w metadanych pochodzących z automatycznej analizy wideo.

Sposób uzyskiwania topologii sieci wykorzystujący algorytm CCN zaproponowali C. Niu i E. Grimson [26]. Dodatkowo poza grafem topologii wynikiem działania tego algorytmu jest model probabilistyczny uwzględniający częstość przejść obiektów pomiędzy daną parą kamer, jak również wiążący ze sobą konkretne części kadrów kamer w pary. Te połączenia wyznaczane są w oparciu o korelację krzyżową sygnałów opisujących częstość pojawiania się i znikania obiektów w konkretnej części kadru dla danej kamery.

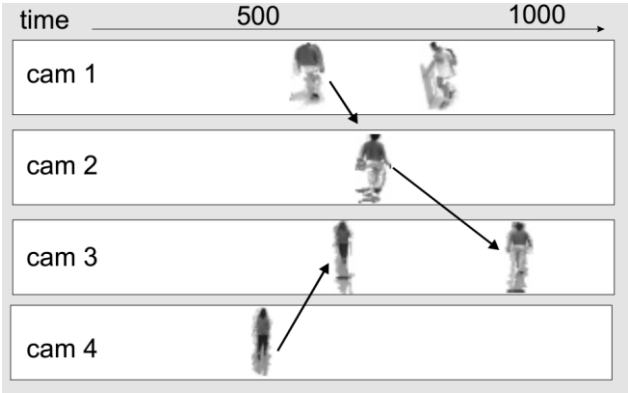
2.4.3 Modelowanie zachowania osób

Istnieją również inne metody modelowania zachowania obiektów. Podobnie jak metody odkrywania topologii modelowanie zachowania wymaga analizy dużej ilości danych

MOST WIEDZY Pobrano z mostwiedzy.pl

wideo. W tym przypadku celem przetwarzania jest uzyskanie statystyki ścieżek ruchu obiektów i częstości przejść pomiędzy poszczególnymi parami kamer. Wiedza zawarta w modelu zachowania opisuje, które ścieżki są częściej, a które rzadziej uczęszczane. Ponadto, model pozwala przewidywać w jaki sposób będzie poruszał się pewien obiekt w oparciu w wiedzę zebraną w przeszłości.

V. Kettnaker i R. Zabih [27] zastosowali modele Markowa pierwszego rzędu do przewidywania najbliższego ruchu obiektu. Modele Markowa wyższych rzędów były użyte do predykcji kilku kolejnych ruchów obiektu pomiędzy kamerami. Do śledzenia obiektów pomiędzy kamerami zostało użyte prawdopodobieństwo warunkowe, które zostało użyte jako przesłanka dotycząca zachowania obiektów.



Rys. 2. Przykładowy wynik powiązania ze sobą obserwacji kilku obiektów [27]

Filtry cząsteczkowe poza śledzeniem obiektu w kadrze pojedynczej kamery mogą posłużyć do śledzenia pomiędzy kamerami. Oczywiście musi być wcześniej znana statystyka ruchu obiektów na obserwowanym obszarze, dzięki której możliwe jest zamodelowanie powiązań pomiędzy kamerami. Takie rozwiązanie problemu modelowania ruchu obiektów także zostało zaprezentowane w ogólnodostępnej literaturze [17].

Kolejnym podejściem pozwalającym zbudować model zachowania obiektów w systemie monitoringu bezpośrednio ze zbioru ścieżek pochodzącego z przeszłości jest zastosowanie idei grafów przepływów [28]. Ponieważ grafy przepływów są oparte na teorii zbiorów przybliżonych, możliwe jest łatwe budowanie reguł dotyczących przyszłych ruchów obiektu i wykorzystanie ich jako przesłanki w metodach śledzenia pomiędzy kamerami [28]. Przykładowo, każda krawędź w grafie przepływów może być opisana za pomocą parametru zwanego pewnością (ang. *cartainty*), który estymuje prawdopodobieństwo względne tego, gdzie pojawi się obiekt, który właśnie zniknął w danej kamerze.

2.4.4 Metody łączące kilka przesłanek

Powstały również rozwiązania dokonujące fuzji kilku rodzajów przesłanek i dokonujące re-identyfikacji danego obiektu pomiędzy kamerami. W tym referacie zostaną przedstawione jedynie wybrane. Przykładem mogą tutaj być prace V. Leunga i in. [29], w których wykorzystano histogram kolorów do opisu wyglądu obiektu, znaną *a-priori* topologię sieci kamer oraz model probabilistyczny określający częstość przejść pomiędzy poszczególnymi parami kamer. Jako środowisko testowe wykorzystano kamery znajdujące się na peronach metra. Dzięki temu zależności czasowe pomiędzy kamerami były łatwe do określenia. Ponadto, nie pojawiały się drastyczne różnice w oświetleniu ze względu na brak światła słonecznego na peronach pod ziemią.

W literaturze została również opisana metoda śledzenia obiektów pomiędzy kamerami zaproponowana przez O. Javeda i in. [20] Również używała ona histogramu kolorów do porównywania wyglądu obiektów. Miarą podobieństwa wyglądu obiektów była wspomniana wcześniej odległość Bhattacharyya. Do kojarzenia ze sobą wszystkich przesłanek służy sieć bayesowska. Przedstawiony algorytm ma dwie fazy. Pierwszą, w której na bazie dużej ilości danych wideo następuje nauka topologii sieci kamer i statystycznego modelu zachowania obiektów. Drugą fazą jest wykorzystanie uzyskanych zależności czasowo-przestrzennych i modelu zachowania przy śledzeniu obiektów między kamerami.

3. PRZYKŁADOWA METODA MODELOWANIA ZACHOWANIA

Przykładowym wspomnianym powyżej sposobem modelowania zachowania jest modelowanie za pomocą grafu przepływów opartego na teorii zbiorów przybliżonych [28]. Na bazie zbioru ścieżek z przeszłości *P* budowany jest graf przepływów.

$$P = \{p : p = \{c_{x0}, c_{y1}, \ldots, o_k\}\} \tag{2}$$

gdzie: P – zbiór ścieżek, p – ścieżka ruchu, c_{xi}, c_{yi} – identyfikatory kamer w i-tym kroku, $i \in \{1,2,...,k\}$ – numer kroku w ścieżce ruchu, o_k – identyfikator wyjścia z obserwowanego obszaru *w k-tym kroku*

Do wyznaczenia parametru pewności (ang. *certainty*) służą poniższe wzory:

$$\sigma(c_{xi}, c_{yj}) = \frac{\varphi(c_{xi}, c_{yj})}{\varphi(G)} \tag{3}$$

gdzie: $\sigma(c_{xi}, c_{yj})$ – częstość występowania w zbiorze ścieżek, które zawierają przejście z kamery c_x w i-tym kroku do kamery c_y w j-tym kroku; $\varphi(c_{xi}, c_{yj})$ – liczba ścieżek, w których występuje przejście z kamery c_x w i-tym kroku do kamery c_y w j-tym kroku; $\varphi(G)$ – liczba wszystkich ścieżek w zbiorze P

$$\sigma(c_{xi}) = \frac{\varphi(c_{xi})}{\varphi(P)} \tag{4}$$

gdzie: $\sigma(c_{xi})$ – częstość występowania w zbiorze ścieżek, które zawierają pojawienie się obiektu w kamerze c_x w i-tym kroku; $\varphi(c_{xi})$ – liczba ścieżek, w których występuje pojawienie się obiektu w kamerze c_x w i-tym kroku; $\varphi(P)$ – liczba wszystkich ścieżek w zbiorze P

$$cer(c_{xi}, c_{yj}) = \frac{\sigma(c_{xi}, c_{yj})}{\sigma(c_{xi})} \tag{5}$$

gdzie: $cer(c_{xi}, c_{yj})$ – miara pewności tego, że obiekt, który pojawił się w kamerze c_{xi} w i-tym kroku następnie pojawi się w kamerze c_{yj} w j-tym kroku $\sigma(c_{xi}, c_{yj})$ – częstość występowania w zbiorze ścieżek, które zawierają przejście z kamery c_x w i-tym kroku do kamery c_y w j-tym kroku;

MOST WIEDZY Pobrano z mostwiedzy.pl

4. WYNIKI UŻYCIA GRAFU PRZEPŁYWÓW

Do określenia miar pewności został użyty zbiór 925 ścieżek pochodzących z testowego zestawu czterech kamer (kamery: A,B,C,D). Chcąc przewidzieć dalsze kroki w ścieżce obiektu można potraktować miarę pewności *certainty*, jako estymatę prawdopodobieństwa warunkowego, że obiekt widziany w jednej kamerze będzie widziany w następnej albo opuści obserwowany obszar. Te prawdopodobieństwa są obliczane za pomocą następującego wzoru:

$$cer[c_{xn},\ldots,c_{zm}]=\prod_{i=xn}^{i=zm} cer(c_i,c_{i+1}) \qquad (6)$$

gdzie: $cer[c_{xn},\ldots,c_{zm}]$ – miara pewności tego, że obiekt, który pojawił się w kamerze c_x w n-tym kroku pokona ścieżkę, która zakończy się w c_z w m-tym kroku

Tabela 1. przedstawia możliwe dalsze kroki w ścieżce obiektu, który w swoim pierwszym kroku pojawił się w kamerze A. Jak widać obiekty pojawiające się w tej kamerze prawie zawsze (w ponad 90% przypadków) wychodziły z obserwowanego obszaru.

Tablica 1. Prawdopodobieństwa dalszych kroków w ścieżce ruchu obiektu, który wszedł na obserwowany obszar w kamerze A [28]

Dalsza ścieżka	Miara pewności
[A1, B2]	0.0287
[A1, B2, A3]	0.0045
[A1, B2, A3, Out4]	0.0045
[A1, B2, Out3]	0.0242
[A1, D2]	0.0656
[A1, D2, Out3]	0.0656
[A1, Out2]	0.9057

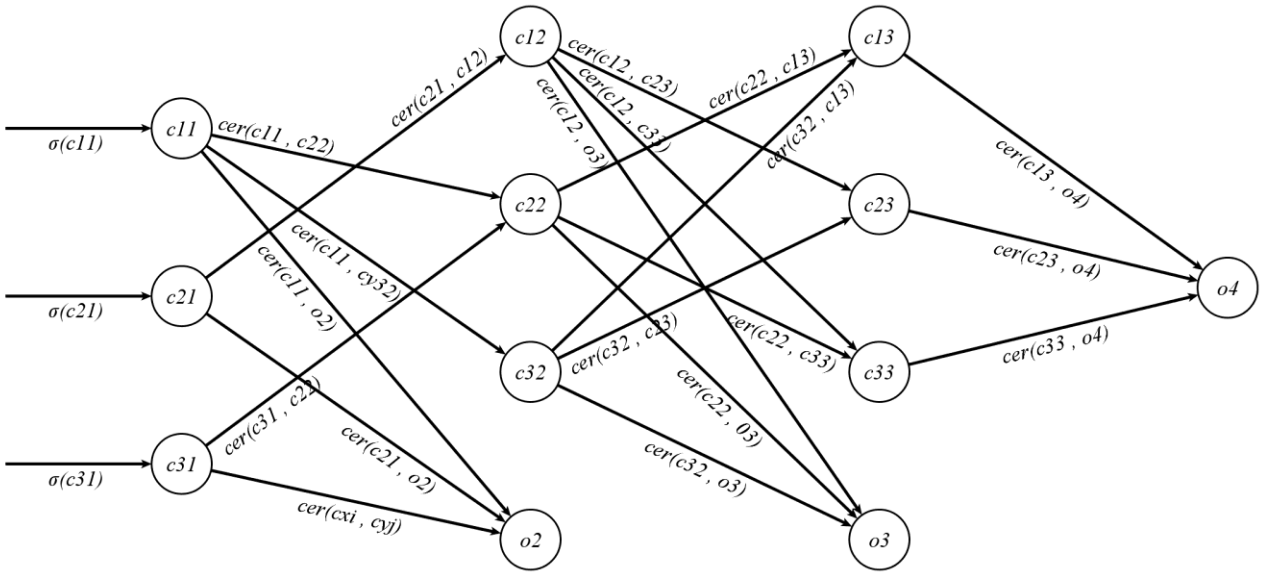
5. WNIOSKI

Śledzenie obiektów jest podstawową funkcjonalnością systemu monitoringu video, która zapewnia szereg możliwości odciążających jego percepcje operatorowi takiego systemu. Kiedy obrazy wideo są obserwowane przez operatora na bieżąco system monitoringu wideo wspomaga percepcję człowieka. Dodatkowo tagowanie materiału (przykładowo etykietą odnoszącą się do pewnego wykrytego zdarzenia) udostępnia możliwość łatwiejszego i szybszego przeglądania dużych zbiorów materiału wideo.

Chcąc powiązać ze sobą różne obserwacje obiektów należy wykorzystać wiele przesłanek tj. wygląd, zależności czasowo-przestrzenne, a także sposób zachowania. Kiedy wszystkie te przesłanki otrzymujemy w sposób automatyczny musimy wziąć pod uwagę, że błędy wprowadzane na kolejnych etapach analizy wideo będą miały wpływ na błędne działanie śledzenia obiektów między kamerami (nastąpi tzw. propagacja błędu). Można z tego wywnioskować, że im szerszy jest kontekst wyników analizy wideo tym większych błędów trzeba się spodziewać.

Pomimo to, rozwijanie algorytmów związanych ze śledzeniem obiektów w systemie monitoringu jest nadal intensywnie badanym zagadnieniem. Ma to związek z szybkim rozwojem systemów monitoringu instalowanych głównie w miejscach publicznych.

Ponadto, w pełni automatyczna analiza ograniczy potrzebę ciągłego bezpośredniego dostępu do strumieni wideo przez operatora systemu monitoringu. Implikuje to zwiększenie bezpieczeństwa i prywatności osób znajdujących się w nadzorowanym obszarze.



Rys. 3. Przykładowy graf przepływów w systemie monitoringu wizyjnego dla trzech kamer

4. BIBLIOGRAFIA

1. Ferenbok J., Clement A.: Hidden Changes: from CCTV to 'Smart' video surveillance", Eyes Everywhere: The Global Growth of Camera Surveillance, 2011, s. 235–253
2. Lisowski K., Czyżewski A.: Cartographic Representation of Route Reconstruction Results in Video Surveillance System, Multimedia and Internet Systems: Theory and Practice Advances in Intelligent Systems and Computing Volume 183, 2013, s. 35-44
3. Nascimento J.C., Marques J.S.: Performance evaluation of object detection algorithms for video surveillance, IEEE Transactions on Multimedia, vol.8, no.4, 2006, s.761-774
4. Ferenbok J., Clement A.: Hidden Changes: from CCTV to 'Smart' video surveillance, Eyes Everywhere: The Global Growth of Camera Surveillance, 2011, s. 235–253
5. Chilgunde A., Kumar P., Ranganath S., Weimin H.: Multi-Camera Target Tracking in Blind Regions of

MOST WIEDZY Pobrano z mostwiedzy.pl

Cameras with Non-overlapping Fields of View, Proc. British Machine Vision Conference, 2004
6. Cieplinski L.: 2001. MPEG-7 Color Descriptors and Their Applications, Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns (CAIP '01), Springer-Verlag, London, UK, 2001, s.11-20
7. Novak C.L., Shafer S.A.: Anatomy of a color histogram, Computer Vision and Pattern Recognition, 1992. Proceedings CVPR '92., 1992 IEEE Conference on Computer Society, 1992, s. 599-605
8. Shao H., Wu Y., Cui W., Zhang J.: Image Retrieval Based on MPEG-7 Dominant Color Descriptor, 2008. ICYCS 2008. The 9th International Conference for Young Computer Scientists, 2008, s.753-757
9. Kasutani E., Yamada A.: The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval, 2001. Proceedings. 2001 International Conference on Image Processing, 2001, s. 674-677
10. Hu M.-H.: Visual pattern recognition by moment invariants, IRE Transactions on Information Theory, 1962, s. 179-187
11. He D.-C., Wang L.: Texture Unit, Texture Spectrum, And Texture Analysis, IEEE Transactions on Geoscience and Remote Sensing, 1990, s.509-512
12. Lee K.-L., Chen L.-H.: An efficient computation method for the texture browsing descriptor of MPEG-7, Image and Vision Computing, 2005, s. 479-489, ISSN 0262-8856
13. Park D. K., Jeon Y. S., Won C. S.: Efficient use of local edge histogram descriptor. In Proceedings of the 2000 ACM workshops on Multimedia, 2000, s. 51-54
14. Tao Y., Skubic M., Han T., Xia Y.: Performance Evaluation of SIFT-Based Descriptors for Object Recognition, IMECS, 2010, s. 17–20
15. Teboul O., Kokkinos I., Simon L., Koutsourakis P., Paragios N.: Shape grammar parsing via Reinforcement Learning, 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011 s. 2273-2280
16. Huet B., Hancock E. R.: Structurally Gated Pairwise Geometric Histograms for Shape Indexing, 2000
17. Honggab K., Romberg J., Wolf W.: Multi-camera tracking on a graph using Markov chain Monte Carlo, 2009. ICDSC 2009. Third ACM/IEEE International Conference on Distributed Smart Cameras, 2009, s.1-8
18. Kailath T.: The Divergence and Bhattacharyya Distance Measures in Signal Selection, IEEE Transactions on Communication Technology, Vol. 15, No. 1, Luty 1967
19. Lev-Tov A., Moses Y.: Path recovery of a disappearing target in a large network of cameras, Proceedings of the Fourth ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC '10), 2001, s. 57-64.
20. Javed O., Shafique K., Shah M.: Appearance modeling for tracking in multiple non-overlapping cameras, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005., s. 26-33
21. Ellwart D., Czyzewski A.: Visual objects description for their re-identification in multi-camera systems, in Video Surveillance System, Multimedia and Internet Systems: Theory and Practice Advances in Intelligent Systems and Computing Volume 183, 2013
22. Bay H., Tuytelaars T., Gool L.V.: SURF : Speeded Up Robust Features. Computer Vi-sion and Image Understanding, 2008
23. Finlayson G. D., Schiele B., Crowley J. L.: Comprehensive Colour Image Normalization, Fifth European Conference on Computer Vision; 1998
24. Farrell R., Davis L.S.: Decentralized discovery of camera network topology, Second ACM/IEEE International Conference o Distributed Smart Cameras, 2008. ICDSC 2008, s.1-10
25. Tieu K., Grimson D. G.: Inference of non-overlapping camera network topology by measuring statistical dependence," ICCV 2005. Tenth IEEE International Conference on Computer Vision, 2005, s.1842-1849
26. Niu C., Grimson E.: Recovering Non-overlapping Network Topology Using Far-field Vehicle Tracking Data, ICPR 2006. 18th International Conference on Pattern Recognition, 2006, s. 944-949
27. Kettnaker V., Zabih R.: Bayesian multi-camera surveillance, 1999. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999
28. Czyżewski A., Lisowski K.: Employing flowgraphs for forward route reconstruction in video surveillance system, Journal of Intelligent Information Systems, 2013
29. Leung V., Orwell J., Velastin S.A.: Performance evaluation of re-acquisition methods for public transport surveillance, 2008. ICARCV 2008. 10th International Conference on Control, Automation, Robotics and Vision, 2008, s. 705-712
30. Nascimento J.C., Marques J.S.: Performance evaluation of object detection algorithms for video surveillance, IEEE Transactions on Multimedia, 2006, s.761,774

MOST WIEDZY Pobrano z mostwiedzy.pl