

# Modélisation de trajectoires cible/caméra : requêtes spatio-temporelles dans le cadre de la vidéosurveillance

Dana Codreanu, André Peninou, Florence Sedes

\*Université de Toulouse – Université Paul Sabatier – IRIT UMR 5505  
118 Route de Narbonne, 31062 Toulouse Cedex 9 - France,  
{codreanu, peninou, sedes}@irit.fr

## 1 Introduction

Le nombre de caméras de vidéosurveillance installées dans le monde augmente chaque jour. En France, le système de la RATP déployé sur Paris comprend 9000 caméras fixes et 19000 mobiles. Lors de faits particuliers (e.g., agressions, vols), les opérateurs de vidéo surveillance se basent sur les indications spatiales et temporelles de la victime et sur leur connaissance de la localisation des caméras pour sélectionner les contenus intéressants pour l'enquête. Deux grands problèmes peuvent alors survenir : (1) le temps de réponse est long (jusqu'à plusieurs jours de traitement) et (2) un risque important de perte de résultats à cause d'une mauvaise connaissance du terrain (appel à des opérateurs extérieurs). Le but de notre recherche est de définir des outils d'assistance aux opérateurs qui puissent, à partir d'une trajectoire donnée, sélectionner de façon automatique les caméras pertinentes par rapport à la requête.

## 2 Contexte de la vidéosurveillance et problématique

Dans un système de vidéosurveillance le contenu est acquis par des caméras fixes et mobiles installées dans des contextes différents (à l'extérieur ou à l'intérieur des bus, des stations de métro, dans les rues, etc.).

Lorsqu'une victime d'une agression porte plainte, on lui demande de décrire les éléments qui pourront aider à trouver les segments vidéos pertinents. Les éléments d'une telle description sont : la *Localisation* (e.g., Paris), la *Date* et l'*Heure* (e.g., 10 octobre 2013 entre 10h et 11h), l'*Événement* recherché (e.g., vol de sac), et la *Trajectoire* de la victime avant et après l'événement (e.g., Rue de Rivoli, de la sortie du musée du Louvre jusqu'à l'entrée de la station de métro Chatelet). Les opérateurs, en se basant sur les aspects spatiaux et temporels de la requête et sur leur propre connaissance de la localisation des caméras, sélectionnent les caméras qui auront référencé r des informations concernant la trajectoire de la personne pour ensuite analyser les enregistrements vidéo.

Le but est de définir un(des) nouveau(x) opérateur(s) prenant comme entrée une trajectoire cible définie comme un ensemble d'unités  $u_k$  (segments cartographiques de la rue de Rivoli)

Modélisation de trajectoires cible/caméra

et un intervalle temporaire et qui calcule la liste d'extraits vidéo pertinents pour la requête :

$$hasSeen : u_1, u_2, \dots, u_n, [t_1, t_2] \Rightarrow \begin{cases} c_1 : t_{start}^1 > t_{end}^1, u_k (1 \leq k \leq n) \\ \dots \\ c_m : t_{start}^m > t_{end}^m, u_k (1 \leq k \leq n) \end{cases}$$

ce qui équivaut à dire que la caméra  $c_i$  a potentiellement "filmé" le segment  $u_k$  de la trajectoire de la requête entre  $t_{start}^i$  et  $t_{end}^i$ , avec  $t_1 \leq t_{start}^i \leq t_2$  et  $t_1 \leq t_{end}^i \leq t_2$ .

### 3 Modélisation des prises de vue de caméras et des trajectoires de déplacement

Afin de pouvoir effectuer les requêtes nécessaires pour assister les opérateurs, nous proposons un modèle intégrant d'une part les réseaux routier et de transport et d'autre part : (1) les caméras fixes et leur position sur le réseau routier, (2) les changements des champs de vue des caméras fixes dans le temps et (3) les caméras mobiles et leur attachement à un objet mobile.

Dans notre approche nous proposons un modèle à quatre couches indépendantes : (1) Réseau routier, (2) Réseau de transport, (3) Objets, (4) Caméras.

La couche *Réseau Routier* est basée sur l'approche de modélisation sous forme de graphe consacrée dans la littérature (Liu et al., 2012). Le réseau routier est un graphe  $G = (N, A)$ , dans lequel les noeuds sont les croisements des rues et les arcs sont les rues. Le *Réseau de Transport* est aussi modélisé sous forme de graphe. A ce niveau, les noeuds sont constitués par des stations de bus qui sont situées sur des segments de rues. Des séquences ordonnées de noeuds de transport forment des lignes (e.g., des lignes de bus). La couche *Objets* modélise les positions des objets fixes et mobiles par rapport aux couches sous-jacentes. Chaque *Objet Mobile* transmet périodiquement sa position en fonction de différentes stratégies (e.g., chaque fois que l'objet change de segment) que nous ne traitons pas dans ce papier. Nous supposons que nous avons au moins une mise à jour de la position par segment de rue. L'ensemble de points avec les timestamps associés forment leur trajectoire (Gutting et al., 2006).

Au-dessus de toutes ces couches, nous modélisons les *Caméras* de Vidéosurveillance. Un schéma simplifié de cette couche est illustré dans la Figure 1. Cette couche est composée des caméras fixes et mobiles. Les caméras fixes ont une position 2D fixée au moment de l'installation. Les caméras mobiles sont associées à un objet mobile (e.g., bus) et leur trajectoire est la même que celle de l'objet.

La nouvelle génération de caméras a des capteurs GPS incorporés et même des boussoles. Les technologies développées autour de ces caméras rendent possible l'extraction automatique des caractéristiques de prise de vue, par exemple : l'orientation, le zoom, la distance focale, etc. En se basant sur ces éléments, il est possible de modéliser le champ de vue et de tracer ces modifications dans le temps (Arslan Ay et al., 2010).

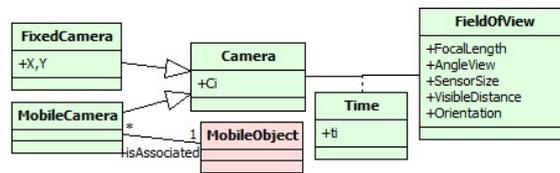


FIG. 1 – Modélisation UML de la couche Caméras

## 4 Exemples de requêtes de recherche de caméras sur une trajectoire

Pour les tests que nous avons faits nous avons utilisé le dataset Oldenburg 1 du générateur de Brinkhoff<sup>1</sup>. Nous avons modifié le générateur pour utiliser des bases de données Oracle Spatial en entrée et en sortie. Cela a généré les tables NODE(ID, POSITION : SDO\_GEOMETRY), EDGE(ID, SEGMENT : SDO\_GEOMETRY) pour le réseau routier et MO(IDOBJ, POSITION : SDO\_GEOMETRY, TIMESTAMP) où les trajectoires des objets mobiles sont stockées.

Nous avons rajouté les tables concernant les caméras :

- FOV(IDFOV, FOCALDISTANCE, VISIBLEDISTANCE, ANGLEVIEW, SENSORSIZE, ORIENTATION, TIMESTAMP) : les caractéristiques du champ de vue qui peuvent changer dans le temps ;
- FC(IDFC, POSITION : SDO\_GEOMETRY, IDFOV) : les caméras fixes avec leur position et le champ de vue associé. Les caractéristiques du champ de vue sont dynamiques et l'historique de ces changements est stocké dans FOV(IDFOV,..., TIMESTAMP) ; Pour calculer la géométrie du champ de vue d'une caméra donnée à un instant donné il faut joindre ces deux tables en utilisant IDFOV et vérifier l'intervalle de temps en fonction du TIMESTAMP ;
- MC(IDMC, IDOBJ) : l'association des caméras mobiles aux objets mobiles ;

Comme requête nous avons considéré des unités  $u_k$  in EDGE.SEGMENT et l'intervalle de temps  $[t_1, t_2]$ . Par la suite nous allons présenter deux exemples de requêtes qui permettent de retrouver les caméras dans la base de données :

Q1 : Sélectionner les caméras fixes dont le champ de vue a intersecté un segment donné :

```

SELECT C.IDFC FROM < EDGES E, FC C, FOV F>
WHERE {
SDO_RELATE(E.SEGMENT, geometry(F, C.POSITION), 'mask=anyinteract')='TRUE'
AND C.IDFOV=F.IDFOV AND E.ID=idSegmentRequete
AND ((F.TIMESTAMP<= t2 AND F.TIMESTAMP >= t1) OR
(F.TIMESTAMP<t1 AND NOT EXISTS (SELECT F1.TIMESTAMP FROM FOV F1
WHERE F1.TIMESTAMP>F.TIMESTAMP AND F1.TIMESTAMP<=t1
AND F1.IDFOV=F.IDFOV))) }
  
```

1. <http://iapg.jade-hs.de/personen/brinkhoff/generator/>

## Modélisation de trajectoires cible/caméra

La requête sélectionne les caméras dont la géométrie du champ de vue (calculée à partir des caractéristiques stockées dans la table FOV et de la position de la caméra par la fonction *geometry*(F, C.POSITION)) intersecte la géométrie du segment de rue donné (SDO\_RELATE est un opérateur spatial qui vérifie si deux géométries ont n'importe quel type d'intersection (paramètre anyinteract)) dans l'intervalle recherché.

Q2 : Sélectionner les caméras mobiles dont la trajectoire a intersecté un segment donné :

```
SELECT MC.IDMC FROM < EDGES E, MO, MC>
WHERE { (SDO_RELATE(E.SEGMENT, MO.POSITION), 'mask=anyinteract')
        ='TRUE'
AND t1 <= MO.TIMESTAMP AND MO.TIMESTAMP <= t2
AND MO.IDOBJ=MC.IDOBJ AND E.ID=idSegmentRequete }
```

La deuxième requête sélectionne les caméras mobiles associées aux objets mobiles dont la trajectoire a intersecté le segment donné entre  $t_1$  et  $t_2$ .

## 5 Conclusion et perspectives

Afin d'aider les opérateurs, nous proposons une approche de modélisation spatiale et temporelle des caméras de vidéosurveillance dans le cadre d'un réseau routier, modélisation basée sur des informations spatiales et des trajectoires des objets mobiles. A partir de ce modèle on peut sélectionner la liste des caméras qui ont potentiellement "filmé" la trajectoire d'une personne. Nous envisageons d'étendre le modèle pour prendre en compte le réseau de caméras à l'intérieur de stations de métro ou de train. Un autre aspect concerne l'ordonnement des résultats en fonction de la distance des caméras par rapport aux segments de la requête.

## Références

- Arslan Ay, S., R. Zimmermann, et K. SeonHo (2010). Effective map-matching on the most simplified road network. *Multimedia Systems* 16, 105–125.
- Gutting, R., V. Almeida, et Z. Ding (2006). Modeling and querying moving objects in networks. *VLDB Journal* 15, 165–190.
- Liu, K., Y. Li, F. He, J. Xu, et Z. Ding (2012). Effective map-matching on the most simplified road network. *In Proceedings of the 20th International Conference on Advances in Geographic Information Systems*, 609–612.

## Summary

Our work concerns the assistance to the operators of videosurveillance in the search for particular videos (e.g., attack of persons, lost object) by the automatic identification of a set of cameras likely to have filmed a required scene. This paper presents a multi-layer modeling (cameras fixed and mobile, road network, public transport network) whose characteristic consists in rather including calculation data of the geometry of the cameras viewpoints on the network layer than to store this geometry as such.