
PhotoMap: annotations spatio-temporelles automatiques de photos personnelles pour les utilisateurs nomades

Windson Viana, José Bringel, Jérôme Gensel, Marlène Villanova Oliver, Hervé Martin

Laboratoire d'Informatique de Grenoble, équipe STEAMER.

681, rue de la Passerelle, 38402 Saint Martin d'Hères

{Windson.Viana-de-Carvalho,Jose.De-Ribamar-Martins-Bringel-Filho,

Jerome.Gensel, Marlene.Villanova-Oliver, Herve.Martin } @imag.fr

RÉSUMÉ. L'augmentation croissante de la quantité de photos numériques personnelles rend les processus d'organisation, de recherche et de visualisation de plus en plus difficiles. Les outils d'annotation manuels existants demandent trop de temps disponible des utilisateurs pour effectuer l'annotation. L'annotation automatique de photos à l'aide de dispositifs mobile est donc la solution émergente pour faciliter la gestion d'images personnelles. Cet article propose un système, appelé PhotoMap, pour l'annotation automatique et pour la visualisation spatio-temporelles de collections de photos prises par un utilisateur nomade. PhotoMap utilise des nouvelles technologies du Web Sémantique pour inférer des informations sur les photos et améliorer les processus de visualisation et de consultation futures. Cet article présente une démonstration d'utilisation de l'application PhotoMap lors d'une promenade touristique réalisée à Grenoble.

ABSTRACT. The increasing number of personal digital photos makes their management, retrieval and visualization a difficult task. Existing tools performing manual annotation are time consuming for the users. To annotate automatically these photos using mobile devices is the emerging solution to facilitate the management of personal photos. In this context, this paper proposes a mobile and location based system, called PhotoMap, for the automatic annotation and the spatial-temporal visualization of photo collections. PhotoMap uses the new technologies of the Semantic Web in order to infer information about the taken photos and to improve both the visualization and retrieval. We also present a demonstration of the PhotoMap application during a tourist tour in the city of Grenoble.

MOTS-CLÉS : Annotation sémantique, ontologies, capture de contexte, dispositifs mobiles, services localisés et Web 2.0 spatial.

KEYWORDS: Semantic Annotation, ontologies, context sensing, mobile devices, location based services, and spatial Web 2.0

1. Introduction

La banalisation des appareils photos numériques et la baisse du coût du stockage de données provoquent la croissance quasi exponentielle de la quantité de photos numériques sur les ordinateurs personnels (Monaghan *et al.*, 2006). Si le coût de la prise et du stockage des photos numériques ne pose plus de problèmes majeurs, la grande quantité de photos rend les processus d'organisation, de recherche et de visualisation plus ardu. Une des solutions existantes pour faciliter la gestion d'images personnelles est l'utilisation d'annotations. Le processus d'annotation consiste en l'ajout de métadonnées sur les photos. Ces métadonnées décrivent le contexte de prise de la photo (la localisation, le climat, les personnes présentes) et/ou les informations sur le contenu de la photo (les personnes ou les objets qui apparaissent sur la photo) (Sarvars *et al.*, 2004). Les annotations permettent le développement d'outils de visualisation plus agréables, comme la visualisation spatiale d'une collection de photos. Les annotations permettent l'emploi de méthodes de recherche plus puissantes, capables de répondre à des requêtes complexes, par exemple : « quelles sont les photos que j'ai pris à Grenoble lors des trois derniers hivers ? ». Divers produits commerciaux et de travaux de recherche proposent des outils d'aide à la saisie manuelle d'annotation. Par exemple, Panoramio¹ et Flickr² permet l'association d'une information spatiale (coordonnée géographique) et l'ajout de « tags » (mots clés) à chaque photo. Hormis certains utilisateurs (tels que les professionnels), la majorité des personnes ne sont pas prêtes consacrer du temps à indexer et annoter correctement chacune de leurs photos (Nars, 2005). De plus, il peut exister un écart considérable entre l'instant de la prise de la photo et le moment de l'annotation. Cet intervalle de temps peut compliquer la tâche de l'utilisateur qui doit alors se rappeler des informations associées à photo (Où la photo a été prise ? Qui était avec moi ce jour là ?).

Bientôt, l'utilisation de nouveaux dispositifs mobiles changera radicalement l'annotation de photos. Ces dispositifs disposeront d'appareils photos numériques intégrés de haute résolution et remplaceront, petit à petit, les appareils photos numériques traditionnels. L'ajout progressive de capteurs (GPS, lecteurs RFID, compas) mettre à disposition des informations sur le contexte de prise d'une photo (coordonnées GPS, température) (Yamaba *et al.*, 2005). L'interprétation et l'inférence sur le contexte de la prise génèrent des informations qui peuvent être utilisées pour annoter automatiquement une photo (Ames *et al.*, 2007). L'autre avantage de l'utilisation d'un dispositif mobile est de permettre à l'utilisateur de compléter, au moment où il a pris la photo, l'annotation générée.

Dans ce contexte, cet article propose un système pour l'annotation automatique et pour la visualisation spatio-temporelle de collections de photos prises par un utilisateur nomade. Notre proposition, appelée PhotoMap, est un système Web de

¹ <http://www.panoramio.com/>

² <http://www.flickr.com/>

gestion de collection de photos prises à l'aide de dispositifs mobiles. L'objectif de PhotoMap est d'offrir des outils pour l'organisation, la navigation, la consultation et le partage de photos. PhotoMap représente une évolution d'autres approches d'annotation automatique pour les utilisateurs nomades, en introduisant un nouveau concept d'annotation de photos : l'annotation spatio-temporelle d'un trajet de photos. L'application mobile du PhotoMap capture automatiquement les informations sur le contexte de prise de chaque photo (la localisation, la date et l'heure, et les dispositifs Bluetooth environnantes) et sur le parcours que l'utilisateur a accompli pour les prendre. L'application Web du PhotoMap utilise l'annotation sur la collection pour dériver des informations spatio-temporelles (la saison, la température) et sociale (les amis présents) sur chaque photo de la collection. PhotoMap utilise ces informations dérivées afin d'offrir une interface de consultation et visualisation des collections de photos.

L'article est organisé de la façon suivante. La section 2 présente les approches d'annotation sémantique existantes. Nous donnons une vision générale de PhotoMap, de l'ontologie OWL créée pour annoter les collections de photos et des mécanismes de capture et d'inférence utilisées dans la section 3. Une illustration d'une utilisation du système est présentée dans la section 4. Les autres approches d'annotation à l'aide de dispositifs mobiles sont présentées dans la section 5. Enfin, dans la section 6, la conclusion et les perspectives de ce travail sont exposées.

2. Annotation sémantique de photos

Afin d'organiser leurs collections de photos personnelles, la majorité des utilisateurs modifie le nom des images pour aider les recherches futures. Ils rangent leurs photos en fichiers qui symbolisent des périodes (par exemple, été 2006), des événements (un voyage, une conférence) et des endroits différents (photos de Saint Etienne) (Matellanes *et al.*, 2006). Néanmoins, ce type d'organisation rudimentaire n'évite pas le problème de la recherche lorsque le nombre de photos augmente. Il est toujours difficile de trouver une photo spécifique lorsqu'on veut la montrer ou la partager avec une autre personne. L'annotation peut faciliter la tâche de gestion des photos personnelles. L'annotation d'une image constitue l'outil principal pour associer de la sémantique à une image. L'ajout de métadonnées sur une photo enrichir sa description et permet la construction d'outils de consultation et de visualisation plus performants. Les mécanismes d'annotation de photos existants sont classés selon le sujet décrit dans l'annotation. Ainsi, on trouve l'annotation sur le contenu et l'annotation sur le contexte de prise d'une photo. Mais, l'annotation peut être différenciée aussi selon la manière dont elle est produite (manuellement, automatiquement ou semi automatiquement).

L'annotation sur le contenu consiste à décrire ce qu'il y a dans une image, par exemple, les objets et les personnes qui apparaissent sur une photo (un chien et Pierre), les relations entre les objets de la scène (le chien de Marie est à gauche de

Pierre), ou l'activité illustré par la photo (une promenade au parc) (Sarvas *et al.*, 2004). Il existe des outils simples d'annotation, comme Flickr et ACDS³, permettant l'ajout de descriptions textuelles sur le contenu. D'autres outils, comme (Lux *et al.*, 2004) et (Hollink *et al.*, 2004), proposent une description plus complexe à l'aide d'un graphe de concepts et des ontologies spatiales en RDF⁴. Ces outils permettent des annotations sur l'image, sur ses régions, et sur les relations spatiales entre eux (« la région A est à droite de la région B »). Néanmoins, ces outils d'annotation de contenu demande aux utilisateurs de consacrer du temps à annoter leurs photos. Cependant, lors du moment de la description, il est difficile pour l'utilisateur d'identifier quels seront les bénéfices futurs apportés par l'annotation (Ames *et al.*, 2007). Il existe des outils qui recommandent des annotations sur le contenu en exploitant la correspondance entre les aspects bas niveau de l'image (texture, couleurs, segments) et ceux d'images préalablement annotées (Wang *et al.*, 2006). Néanmoins, ces outils produisent encore un écart sémantique entre les annotations suggérées et les annotations attendues par les utilisateurs. Dans cette première phase de conception de PhotoMap, nous ne nous intéressons pas à la recommandation d'annotations sur le contenu l'image. Notre approche se concentre sur la génération d'annotations d'une photo en utilisant uniquement les inférences sur le contexte de la prise de photo. Nous pensons en effet, comme (Naaman *et al.*, 2004) et (Nars, 2005), que les annotations suggérées avec ces mécanismes présentent des résultats plus utiles pour la gestion de photos personnelles.

Une partie des informations qui aident les personnes à se rappeler d'une photo sont les aspects contextuels du moment de sa prise (« quand », « où », « avec qui »). **L'annotation sur le contexte** d'une photo consiste à la décrire à l'aide de ces aspects contextuels. La date et l'heure de prise d'une photo sont les informations sur le contexte les plus faciles à capturer par les appareils photos numériques traditionnels. L'outil proposé par (Nars, 2005) et Adobe PhotoShop Album⁵ utilisent ces informations pour établir un ordre chronologique sur les photos. Elles génèrent des regroupements temporels de ces images afin de rendre plus facile la consultation.

Une autre ressource d'information importante sur le contexte d'une photo numérique est l'archive EXIF (*EXchangeable Image File*). Lors de la prise d'une photo, la plupart des appareils photos numériques ajoute à l'en-tête de l'image une description au format EXIF. Cette description possède des métadonnées sur la photo (nom, taille, la date), sur l'appareil de photo (modèle, résolution maximale) et sur les réglages utilisés pour prendre la photo (l'intensité du flash, l'ouverture du diaphragme). L'archive EXIF permet aussi l'ajout au descriptif de l'image des coordonnées GPS. Ces informations spatiales sont ajoutées : (i) **automatiquement**, au moment de la prise de photo à l'aide d'une connexion directe à un récepteur GPS ; (ii) **manuellement**, des outils peuvent synchroniser un trajet stocké par un récepteur GPS avec l'instant de prise de chaque photo.

³ <http://www.acdsee.com/>

⁴ <http://www.w3c.org>

⁵ <http://www.adobe.com/>

La possibilité d'accéder à la fois aux données spatiales et temporelles sur une photo est exploitée par les approches PhotoCompas (Naaman *et al.*, 2004) et WMMX (Toyama *et al.*, 2003). PhotoCompas utilise des Services Web pour dériver des informations sur l'espace et le temps (telles que la saison, la température et la quantité de lumière). PhotoCompas transforme les coordonnées GPS en une hiérarchie de localités (telle que Pays > Département > Ville) et offre une interface de requête spatiale en utilisant cette hiérarchie. Tous les attributs de l'annotation générée (tels que l'altitude, la saison) peuvent être utilisés afin d'effectuer des recherches sur les collections d'images. WMMX est un projet de recherche de Microsoft dont l'objectif est de construire un outil de visualisation des images à l'aide d'un SIG (Système d'Information Géographique). WMMX utilise une base de données spatiale pour indexer des photos personnelles géolocalisées. WMMX offre une interface de visualisation avec des cartes fournies par MapPoint⁶. Capturer des informations sur le contexte de la prise de la photo est la principale difficulté des solutions d'annotation sur le contexte comme PhotoCompas et WMMX. A l'heure actuelle, il existe peu d'appareils photos numériques disposant de GPS intégré. De plus, l'information temporelle (date et heure) de la prise d'une photo n'est pas toujours fiable, étant donné que la plupart des personnes ne règlent pas correctement l'horloge de leurs appareils photos. Notre proposition, PhotoMap, utilise les dispositifs mobiles comme source d'information sur le contexte afin d'augmenter la quantité et la qualité de cette information. Ces dispositifs sont dotés de nouveaux capteurs (GPS, Bluetooth) pour l'acquisition de données sur le contexte. De plus, le support des plates-formes de programmation (telles que Symbian, J2ME) rend possible le développement d'applications pour ces dispositifs. Ceci permet la mise en œuvre d'outils d'annotation plus puissants (Ames *et al.*, 2007). L'autre avantage est que l'information sur l'heure et la date du jour, maintenue par ces dispositifs, est en général plus fiable, étant donné que la plupart de personnes utilisent leurs dispositifs mobiles comme montre ou agenda.

3. PhotoMap

PhotoMap est un système Web et mobile pour l'annotation et la visualisation des photos prises par des utilisateurs nomades. L'objectif de PhotoMap est de fonctionner à la fois comme un système d'annotation automatique et de consultation spatio-temporelle de collections de photos. La Figure 1 montre l'architecture générale de PhotoMap. Le client mobile du PhotoMap est un MIDlet J2ME⁷ qui permet l'utilisateur mobile de prendre des photos, d'ajouter des commentaires et de regrouper ses photos en collections au moment de la prise. L'application mobile capture automatiquement les informations sur le contexte de prise de chaque photo (la localisation, la date et l'heure et les dispositifs Bluetooth environnants). La capture est mise en œuvre en interrogeant des capteurs disponibles sur le dispositif

⁶ <http://mappoint.msn.com/>

⁷ www.java.com

(un GPS intégré ou connecté via Bluetooth). Le client mobile stocke les annotations, les photos et le parcours que l'utilisateur a accompli pour les prendre. Ces annotations et ces photos sont envoyées au serveur Web. Les annotations sont décrites à l'aide d'une représentation sémantique. Le serveur de PhotoMap utilise l'annotation sur la collection et accède des services Web distants pour dériver des informations sur le contexte de prise de la photo (telles que la saison, le climat).

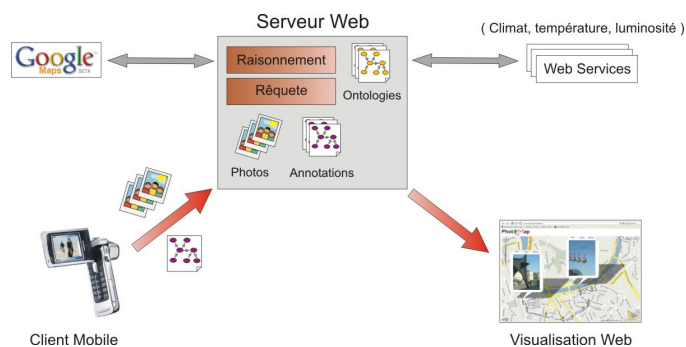


Figure 1. L'architecture générale de PhotoMap

PhotoMap utilise les technologies du Web Sémantique afin d'inférer des informations sur le contexte social des photos et identifie automatiquement les amis de l'utilisateur présents lors de la prise de chaque photo. Les annotations résultantes de ces inférences sont exploitées par l'outil de visualisation de PhotoMap. Le site Web PhotoMap permet à l'utilisateur, à l'aide de services Web 2.0 spatiaux, de visualiser ses photos et le trajet parcouru pour les prendre, de voir les endroits où il a pris ces photos et de consulter l'annotation riche qui a été générée.

3.1. L'ontologie *ContextPhoto*

Différentes structures de représentation sont employées pour décrire des annotations d'une image (telles que texte, EXIF, MPEG7). Cependant, le MPEG7 est un standard davantage adapté à la description du contenu, et le format EXIF ne supporte pas les attributs désirés pour notre système de consultation. Ainsi, notre approche propose l'utilisation d'une représentation sémantique de l'annotation basée sur des ontologies. Les ontologies expriment, de manière formelle et explicite, les concepts et relations d'un domaine. Nous avons modélisé une ontologie, *ContextPhoto*, pour l'annotation en utilisant le standard OWL-DL⁸. Ce standard de représentation présente l'expressivité de la logique de description et reste encore décidable. Ce

⁸ www.w3c.org

choix permet l'utilisation de mécanismes d'inférences existants dans le but d'enrichir l'annotation. L'ontologie est aussi utilisée pour améliorer les processus futurs de requête et de visualisation de photos. La **Figure 2** montre les principaux concepts et relations de notre ontologie. La principale proposition de l'ontologie ContextPhoto est l'association d'un trajet spatial et d'une durée temporelle à une collection de photos. (Naaman *et al.*, 2004) et (Matellanes *et al.*, 2006) montre, en utilisant des tests d'utilisabilité, que le regroupement de photos en événements (vacances, randonnée) est la manière la plus utilisée par les personnes pour se souvenir d'une photo. Le concept *EventCollection* de ContextPhoto évoque l'idée d'une collection de photos associées à un événement. Un intervalle de temps (concept *Interval*) et un ensemble ordonné de points d'un trajet (*TrackPoint*) sont associés à *EventCollection* afin d'incorporer des aspects temporels et spatiaux à la collection. *EventCollection* possède une propriété (*hasPhotos*) qui décrit les photos (*Photo*) qui appartient à la collection. Le concept *Photo* contient les propriétés de base d'une image (telles que son nom, sa taille). Chaque photo peut avoir différents types d'annotation : les annotations textuelles (*AnnotationTag*) et les annotations contextuelles (*Shot Context*). Ce choix de modélisation permet à l'ontologie d'exprimer des annotations manuelles décrites par les utilisateurs et des annotations générées automatiquement.

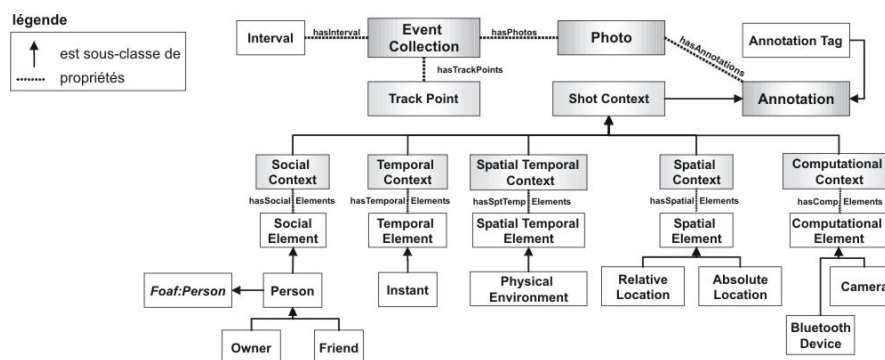


Figure 2. L'ontologie ContextPhoto

Les annotations contextuelles sont de cinq types : spatiales (*Spatial Context*), temporelles (*Temporal Context*), sociales (*Social Context*), informationnelles (*Computational Context*) et spatio-temporelles (*Spatial-Temporal Context*). Ces concepts intègrent les principaux éléments de description du contexte de prise d'une photo (« où », « qui était présent », « quand », « quoi »). Un utilisateur peut se souvenir d'une photo à l'aide d'une adresse (« la rue du pont »), d'une description de l'endroit (« le parc Mistral ») ou d'un point spécifique (« sommet du mont Aguille »). ContextPhoto incorpore ces différents niveaux sémantiques de description d'une localisation en utilisant le concept *SpatialContext*. Celui-ci peut être défini de façon géométrique à l'aide du concept *Absolute Location*. Ce concept importe l'ontologie

NeoGeo⁹, une représentation en OWL de GML (*Geographic Markup Language*). Une localisation peut être aussi décrite de façon relative (« à droite de » et « à gauche de ») et comme un adresse postal à l'aide du concept *Relative Location*.

Le temps est un autre aspect important pour l'organisation et la recherche des photos. Cependant, la date n'est pas forcément l'attribut le plus utilisé lorsqu'une personne souhaite chercher une photo de façon temporelle (Naaman *et al.*, 2004). Les attributs tels que le moment de la journée, le mois, l'année se montrent plus pertinents pour la recherche. Le concept *Temporal Context* permet d'associer à une photo un instant. Celui-ci possède des propriétés qui décrivent séparément toutes les facettes temporelles d'une date et d'une heure (telles que le mois, le jour de la semaine, le moment de la journée). Le concept *SpatialTemporalContext* contient les éléments du contexte de prise qui dépendent du temps et de l'espace pour être calculés. Dans cette première version de ContextPhoto, un seul élément spatio-temporel a été défini : l'environnement physique. Il contient des propriétés pour décrire la saison, la température et le moment de la journée en relation au coucher du soleil. Ces attributs sont utilisés à la fois pour enrichir la description d'une image et pour faire des requêtes (« donne-moi toutes les photos prises lors de l'été 2006 avant le coucher du soleil »). Le contexte informationnel d'une photo (*Computational Context*) décrit les dispositifs numériques présents lors de la prise de photo (l'appareil photo, les dispositifs Bluetooth environnants). Le concept *Camera* décrit les caractéristiques de l'appareil photo et les réglages de l'appareil utilisés pour prendre la photo. Ce concept intègre les attributs du format EXIF. Le concept *Bluetooth Device* contient les adresses Bluetooth des dispositifs environnants. Ce concept joue un rôle important lors de la dérivation du contexte social d'une photo.

Une des caractéristiques innovantes de l'ontologie ContextPhoto est la capacité de décrire le contexte social d'une photo. ContextPhoto fait référence à la proposition de (Monaghan *et al.*, 2006) d'utiliser l'adresse Bluetooth du dispositif personnel de l'utilisateur comme l'identificateur de son profil FOAF¹⁰. L'ontologie FOAF permet la description d'une personne (nom, email, image personnelle) et de son réseau social (les connaissances de la personne). Le concept *Person* de ContextPhoto élargit le concept *foaf:Person* en ajoutant une propriété *hasBluetoothDevice*. D'autres classes sont définies à partir de *Person* : *Owner* et *Friend*. Le concept *Owner* décrit le propriétaire de la photo (le photographe). Le concept *Friend* représente les amis du propriétaire qui étaient présents au moment de la prise de la photo. Ce concept est défini à l'aide d'une condition nécessaire et suffisante de restriction sur la classe *Person*. Cette restriction est décrite en langue naturelle comme : « les instances de *Friend* sont toutes les instances de *Person* qui appartiennent à la propriété *foaf:knows* de *Owner* et qui contiennent un *BluetoothDevice* présent dans le contexte de la photo ». Cette restriction est utilisée par le raisonneur lors du processus d'inférence pour définir quels amis de *Owner* étaient présents au moment de la prise de la photo.

⁹ <http://mapbureau.com/neogeo/neogeo.owl>

¹⁰ <http://www.foaf-project.org/>

3.2. Processus d'inférence sur l'annotation

Le client mobile de PhotoMap envoie au serveur des instances de ContextPhoto contenant des données capturées par le dispositif. Chaque collection possède une annotation qui contient un identifiant de l'utilisateur (*foaf: mbox*), les points du trajet parcouru, la date et l'heure de chaque photo, les coordonnées de GPS associées à chaque photo et les adresses *Bluetooth* des dispositifs environnants. Le processus d'inférence et d'interprétation d'annotation est mis en œuvre après le transfert de ces annotations. L'objectif principal de ce processus est d'enrichir les informations décrites dans les instances de l'ontologie ContextPhoto. L'exécution de cette étape est essentielle pour permettre la création de requêtes complexes et de haut niveau sur les informations décrites dans les annotations. A la fin de ce processus, l'annotation OWL résultante est composée par les données capturées et par les informations inférées. PhotoMap utilise l'approche proposée par (Naaman *et al.*, 2004) qui consiste accéder des Services Web distants afin d'enrichir les informations spatiales et spatio-temporelles de l'annotation de chaque photo. Cette approche réduit le coût de développement et bénéficie des avantages de la technologie des Services Web (réutilisation, standardisation, distribution). Dans un premier temps, PhotoMap interroge le Service Web *AddressFinder*¹¹ afin de transformer les coordonnées GPS de chaque photo en des adresses géographiques. Le serveur de PhotoMap lie la propriété *gml:pos* des instances de la classe *Absolute Location* de chaque photo. Cette propriété décrit les coordonnées GPS au format WSG84 (*World Geodetic System*). Les coordonnées sont envoyées au Service Web et le résultat de la requête est enregistré dans l'ontologie. PhotoMap crée une instance du concept *Address* et l'insère dans l'annotation spatiale de la photo.

Une deuxième étape de dérivation consiste à séparer les composants temporels de la date et l'heure. Ainsi, le serveur calcule les valeurs des propriétés tels que *hasDayofWeek*, *hasMonth* et *hasYear*. Une fois calculées les attributs spatiaux et temporels, le serveur de PhotoMap génère des informations sur l'environnement physique de chaque photo. PhotoMap dérive la température, la saison, la luminosité et le climat en utilisant les coordonnées GPS et la date/heure annotées par le client mobile. Le serveur utilise le service *Weather Underground*¹² pour renseigner sur le climat et la température. Pour découvrir l'heure de coucher du soleil, PhotoMap utilise le service *Sunrise and Sunset Times*¹³. La saison est calculée dans le serveur en utilisant la date et la coordonnée GPS. La dernière étape de dérivation a pour objectif de découvrir les personnes présentes au moment de la prise de la photo (le contexte social). Le serveur de PhotoMap utilise l'identificateur de l'utilisateur pour récupérer le profil de l'utilisateur dans les registres des profils de PhotoMap. Les propriétés «*foaf:Knows*» et «*rdf:seeAlso*» du profil FOAF de l'utilisateur sont exploitées afin de retrouver des profils de ces amis qui sont distribués sur

¹¹ <http://ashburnarcweb.esri.com/>

¹² www.weatherunderground.com/

¹³ <http://www.earthtools.org/>

l'internet. Après la capture des profils et l'insertion de ces données dans l'annotation, le processus de classification du raisonnement est appelé dans le but d'établir quels amis de l'utilisateur étaient présents au moment de chaque photo.

3.3. Visualisation sur le Web

PhotoMap offre une application pour la visualisation et la consultation de collections de photos. Nous avons choisi d'implémenter la visualisation spatiale des collections de photos à l'aide de cartes. Des études d'utilisabilité montrent que la visualisation spatiale de photos en utilisant des cartes présente des avantages d'interactivité comparativement à la visualisation textuelle hiérarchique d'une localisation (Naaman *et al.*, 2004). De plus, la structure hiérarchique d'une localisation ne permet pas de représenter de manière intuitive un parcours.

Après la sélection d'une collection, PhotoMap affiche les photos et le trajet parcouru pour les prendre. L'utilisateur peut consulter l'annotation riche générée. Pour afficher des cartes, le serveur PhotoMap utilise Google Maps pour visualiser de manière spatiale ces photos et ces parcours. Google Maps est largement utilisé, étant donnée qu'il présente des mécanismes d'interactivité standardisés (pan, zoom, vision mixte) et facilite son intégration lors de l'implémentation de nouveaux services Web 2.0 comme PhotoMap. Lors du processus de visualisation spatio-temporelle d'une collection de photos, le serveur PhotoMap exécute la lecture de l'instance de l'ontologie ContextPhoto associée à cette collection en utilisant l'API Jena. PhotoMap ajoute sur la carte une marque pour chaque photo de la collection avec leurs annotations. PhotoMap lit toutes les coordonnées GPS du parcours de la collection afin de tracer une poly ligne qui représente le trajet.

4. Illustration de l'utilisation du PhotoMap

Cette section décrit une démonstration d'utilisation de l'application PhotoMap à partir d'un scénario qui consiste en une promenade touristique réalisée à Grenoble (France) par un groupe de touristes (Pierre, Joseph et Marie). Pierre possède un dispositif mobile K750i, connecté par Bluetooth à un GPS. Les autres deux touristes portent des mobiles dotés de Bluetooth activé. L'application mobile de PhotoMap est installée dans le dispositif de Pierre. Au début de la promenade, Pierre active le client mobile de PhotoMap afin de prendre des photos et d'enregistrer automatiquement son trajet. La **Figure 3a** présente Pierre lors du moment de la prise d'une photo de Joseph. Le client mobile capture les coordonnées GPS pendant tout le parcours. Dans cette démonstration, PhotoMap a capturé 109 coordonnées GPS et Pierre a pris 13 photos de points touristiques de Grenoble. La durée totale de la promenade est d'environ 36 minutes. À la fin de la promenade, Pierre utilise le client PhotoMap dans le but de transférer la collection de photos et les annotations générées (les instances de l'ontologie ContextPhoto) au serveur PhotoMap. Après ce

transfert, le serveur de PhotoMap exécute les processus d'interprétation et d'inférence des annotations. Ainsi, Pierre peut exécuter le module de visualisation spatio-temporelle du serveur PhotoMap afin de voir le trajet et les photos prises.

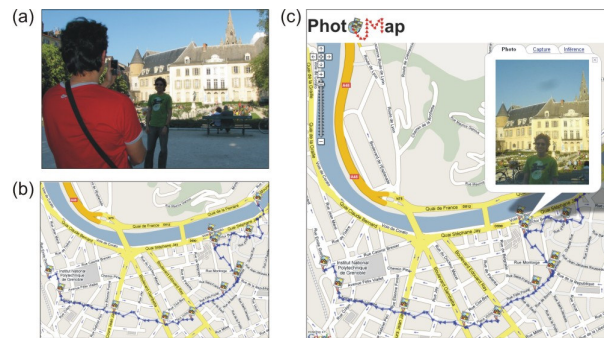


Figure 3. L'utilisation du client mobile et du serveur Web de PhotoMap.

La **Figure 3b** illustre une visualisation du parcours effectué par les trois touristes. La **Figure 3c** présente une photo de Joseph localisée sur le trajet parcouru. La visualisation de parcours à l'aide d'un zoom montre différents espacements entre les points du trajet. Une concentration de points correspond aux moments où le groupe de touristes s'est arrêté (capturer une photo, attente d'un feu piéton). La **Figure 4** présente la visualisation d'une photo de Marie placée sur la carte de Grenoble en mode satellitaire. L'onglet « Capture » décrit les données brutes de l'annotation: les coordonnées GPS, la date et l'heure, et les trois adresses physiques des dispositifs *Bluetooth* voisins. L'onglet « Inférence » présente les informations dérivées par le serveur PhotoMap: le lieu de la photo (Jardin de Ville, Grenoble - FR), la température (23°) et la saison (printemps). L'onglet « Inférence » montre aussi le nom et la photo de deux amis qui étaient présents au moment de la prise de la photo. Il est important de souligner que le dispositif a trouvé trois adresses *Bluetooth* et l'application a inféré correctement que seulement deux parmi eux étaient des dispositifs amis du propriétaire.

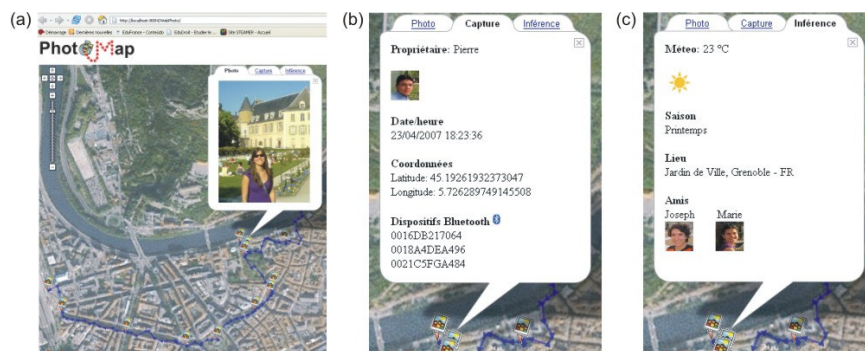


Figure 4. Visualisation des annotations inférées sur le contexte

5. Travaux voisins

Quelques travaux de recherche proposent l'annotation de photos à l'aide de dispositifs mobiles. Par exemple, le système **MMM Image Gallery** (Sarvas *et al.*, 2004) propose une application d'annotation semi-automatique pour les appareils Nokia 3650. Lors d'une prise de photo, l'application capture l'ID de la cellule GSM, l'identité de l'utilisateur et l'instant de prise de la photo. L'image est envoyée à un serveur avec son annotation. Le serveur exécute un processus de manipulation sur l'image pour en déduire des informations sur son contenu (si la photo a été prise à l'extérieur ou à l'intérieur). L'annotation produite est soumise à l'utilisateur pour une validation. Néanmoins, l'outil ne présente pas de mécanismes de visualisation et de consultation de photos. Outre, MMM Image Gallery présente des problèmes de précision sur la localisation occasionnés par l'utilisation de l'ID de la cellule GSM. **Zonetag** (Amens *et al.*, 2007) est un projet de recherche Yahoo pour l'annotation des images à l'aide de dispositifs mobiles. Zonetag permet aux utilisateurs de prendre des photos et de les publier sur Flickr. Zonetag utilise des informations sur l'utilisateur et sa localisation afin de recommander des mots clés d'annotations qui sont renvoyés au dispositif. L'utilisateur peut utiliser les fonctionnalités de Flickr pour visualiser ses images. Notre proposition, PhotoMap, offre automatiquement plus d'information sur le contexte de prise d'une photo se comparé à ZoneTag.

(**Monaghan *et al.*, 2006**) proposent une démarche pour dériver des informations sur le contexte social d'une photo. Le travail suggère la possibilité de créer une application mobile qui capture les adresses Bluetooth des dispositifs environnants. L'utilisation de requêtes sur un registre universel de profils FOAF¹⁴ permet d'indiquer les personnes présentes au moment de la prise de photo. PhotoMap adopte la même proposition pour dériver le contexte social d'une photo. Cependant, PhotoMap utilise une approche différente pour dériver l'information sur le contexte social. PhotoMap n'utilise pas un registre universel de profils FOAF. PhotoMap parcourt les profils des utilisateurs et utilise les liens RDF afin de récupérer les profils des amis de l'utilisateur. PhotoMap, différemment des autres approches, joint l'annotation sur le contexte spatial et temporel avec l'annotation sociale.

6. Conclusion et travaux futurs

PhotoMap aide à diminuer le processus coûteux d'organisation de photos. Le processus de capture et d'inférence sur le contexte de prise d'une photo offre de nouvelles façons de visualiser et de consulter des collections de photos. Les résultats initiaux d'utilisations de PhotoMap sont très prometteurs. A court terme, les fonctionnalités de requêtes temporelles et attributaires seront ajoutées à l'interface de visualisation. L'ontologie ContextPhoto sera couplée à des ontologies de raisonnement spatial et temporel afin de créer un outil puissant de requête sur les

¹⁴ <http://www.foaf-project.org/>

photos. A moyen terme, un modèle de collaboration, de respect à la vie privée et de partage de photos et trajets sera intégré à PhotoMap afin de transformer le système en un Wiki mobile pour les touristes itinérants.

7. Bibliographie

- Ahern S., Davis M., Eckles D., King S., Naaman M., Spasojevic M., «ZoneTag: Designing Context-Aware Mobile Media Capture to Increase Participation». *Actes du Workshop on Pervasive Image Capture and Sharing (PICS 2006- Ubicomp 2006)*, USA, 2006.
- Ames M., Naaman M., «Why We Tag: Motivations for Annotation in Mobile and Online Media». *Actes de la Conférence on Human Factors in computing systems (CHI 2007)*, San Jose, CA, USA, 2007.
- Hollink L., Nguyen G., Schreiber G., Wielemaker J., Wielinga B., Worring M., «Adding Spatial Semantics to Image Annotations». *Actes du 4ème International Workshop on Knowledge Markup and Semantic Annotation*, 2004.
- Lux M., Klieber W., Granitzer M., «Caliph & Emir: Semantics in Multimedia Retrieval and Annotation», *Actes de la 19ème International CODATA Conference 2004: The Information Society: New Horizons for Science*, Berlin, Allemagne, 2004.
- Matellanes A., Evans A., Erdal B., «Creating an application for automatic annotations of images and video», *Actes du 1ème International Workshop on Semantic Web Annotations for Multimedia (SWAMM)*, Edinburgh, Scotland, 2006.
- Monaghan F., O'Sullivan D., «Automating Photo Annotation using Services and Ontologies», *Actes de la 7ème International Conference on Mobile Data Management (Mdm'06)*, Washington, DC, USA, Mai 2006. IEEE Computer Society, p. 79-82.
- Naaman M., Harada S., Wang Q., Garcia-Molina H., Paepcke A. «Context data in geo-referenced digital photo collections». *Actes de la 12ème ACM international Conference on Multimedia (MULTIMEDIA '04)*, New York, NY, USA, 2004, ACM, p.196-203.
- Nars E., «Visualisation multi-échelles pour photos personnelles», *Actes de la 17ème conférence francophone sur l'Interaction Homme-Machine (IHM 2005)*, Toulouse, Septembre 2005, ACM Press, International Conference Proceedings Series, p. 259-262.
- Sarvas R., Herrarte E., Wilhelm A., Davis M., «Metadata creation system for mobile images», *Actes de la 2ème International Conference on Mobile Systems, Applications, and Services (MobiSys '04)*, Boston, MA, USA, 2004, ACM, p 36-48.
- Toyama K., Logan R., Roseway A., «Geographic location tags on digital images». *Actes de la 11ème ACM international Conference on Multimedia (MULTIMEDIA '03)*, Berkeley, CA, USA, Novembre 2003, ACM Press, p. 156-166.
- Yamaba, T., Takagi, A., Nakajima, T., Citron: A context information acquisition framework for personal devices, *Actes de la 11ème International Conference on Embedded and real-Time Computing Systems and Applications*. 2005.
- Wang L., Khan L., «Automatic image annotation and retrieval using weighted feature selection», *Journal of Multimedia Tools and Applications*, 2006, ACM, p.55-71.