# Harnessing Location-Context for Content-based Services in Vehicular Systems

Vasudha Ramnath, Joo-Hwee Lim, Jean-Pierre Chevallet[1], Daqing Zhang
Institute for Infocomm Research
[1]Image Processing & Application Laboratory
21 Heng Mui Keng Terrace, Singapore 119613
Email: {vasu,joohwee,viscjp,daqing}@i2r.a-star.edu.sg

*Abstract*— **By incorporating personalised content as context, vehicular telematic services can be made more user-centric, as opposed to vehicle-centric. We develop an ontology based model to integrate both context and content in one unifying framework to facilitate the creation of such services. Scenes captured spontaneously by the user can be a useful index to his or her intent and interests. Our empirical study shows that location-context leads to a higher precision in recognizing these captured scenes. A "tourist information" prototype based on scene recognition is built, to illustrate one such location-enhanced, user-centric service.**

Keywords : context aware, telematics, content analysis, location-based ontology, framework

## I. INTRODUCTION

Advances in automotive telematics has made possible the development of a large number of services such as navigation assistance, remote monitoring and preventive maintenance, traffic control. All these services are primarily directed towards the key requirements of safety and convenience of the occupants of the automotive unit, and enabling the upkeep of the vehicle itself in a roadworthy state. Sensing technologies play an important role here, in monitoring the state of the treads on the tyres, pressure and temperature within the engine etc. However, other information, such as user-specific road-preferences, inputs on dynamic traffic conditions etc. can help to provide the road-user with a more satisfying experience.

*Context* is defined as any information about the concepts, entities and relationships in a domain [1]. In a vehicular system, the state of the vehicle, the user's preferences, the state of the traffic system and the knowledge of the navigable paths within the system are some examples of useful context. Context-aware or situation-aware systems are those that use context information to give the user a more personalised, and hence more satisfying, service.

Work on incorporating higher-level context, such as that pertaining to a Person or Activity into automotive telematics has been reported in [2]. The use of an OSGi-based [3] standard platform within a vehicle, that allows dynamic provisioning of services, and provides a context infrastructure to support context aggregation and reasoning, has also been reported [4].

The use of context enables a user-centric approach to service provisioning, by capturing information about the user's intent and interests. Purely GPS-based navigation services, for e.g., tend to be car-centric, and may fail to capture the user's intent. To give an example of the latter, "If in the vicinity of Little India, and traffic flow is smooth, make detour to pick up fresh cottage cheese". In our work, we enhance the context-infrastructure, which supports rule-based inferencing and semantic queries, to incorporate content (specifically images) as context sources. Image analysis generates important context.

Route guidance systems typically combine positioning information with digital maps and knowledge of landmarks. *Landmarks* are important in wayfinding, as they help to organize space in terms of reference points in the real-world [5], and have an implicit notion of location. Cognitive landmarks [6] are generic landmarks that have cultural or historic significance. It may also include landmarks that have personal significance to a user, such as a house where the user once lived, a friend's dwelling etc. Personal landmarks are formed by semantic associations in a person's memory, that are not captured in an aerial view of the world. They form part of a person's context.

Images of landmarks are one type of content that may serve as a context source. The user may capture landmarks of interest using a wireless-enabled camera, or a camera-phone. A "landmark-aware" wayfinding service could use this content, in combination with position information, to create a *personalized overlay* on a navigation map, or provide personalized routing instructions.

However, to achieve this, the captured content needs to be analysed, semantically understood, and annotated, for incorporation into a *Context Knowledge Base* (CKB) accessible to any of the context-aware vehicular services.

Analysing and automatically annotating the content, in the form of images, is the focus of some related research efforts. In MMM, a system for Mobile Metadata Management [7], the emphasis is on empowering the end-user to create metadata at the time of content capture. Hence, they have integrated user-feedback into the content-gathering and annotation cycle. However, they do not focus on the process of content analysis, or the use of context therein. Also, it appears that "variants" in the images, such as faces, and complex context such as that pertaining to Activity, is "guessed" using regularity trends, unlike our model, which leverages a variety of context sources within the infrastructure for UserContext, ActivityContext

etc. Infoscope [8] was one of the early 'information augmentation' devices reported. The 'Augmentation in the City" application provides enhanced information on a building in a scene captured by the user's camera. Position information, from a GPS was used to allow the image matching algorithm to solve the scene recognition problem in less time. However, the Infoscope architecture does not have a model for the use of generic context entities. In IDeixis [13], the user is equipped with a set-up similar to ours to achieve a point-by-photograph medium. Content-based image retrieval is performed by leveraging similarity matching techniques such as Color Histogram and Fourier Transform to find images in the database that match the user's input. The world-wide-web serves as the image database. The system then extracts keywords (or annotations) from the web-pages found, and uses those keywords to perform a text-based web-search to retrieve additional web pages (and information). However, it appears that their current system does not integrate contextual cues into the image-search.

## II. CONTENT AND CONTEXT

### A. Synergy Scenarios

Context is important to content processing, be it an application for automatic image annotation such as in a photo album, or query-by-example, as in iDeixis and SnapToTell Our current prototype *uses* the Concept of Space or Location for image understanding and querying. In IDeixis, image analysis has been used to *derive* location context. An application that uses additional context such as Person and Activity to recognize faces using feature vector matching, for instance, is doable. Activity and Location context would be used to select the possible persons who could appear in a certain image.

The orientation of the user is another type of context that can be inferred from image analysis. This can be used to provide navigational cues. For e.g., panoramic views (360 deg) of locations could be captured and stored in a database. The user is asked to provide an image taken from his current orientation. Image-matching techniques can then be used to find the orientation of the user.

Location, orientation and intent contexts can be used to extract the relevant images of landmarks from an application database and juxtaposed on a GIS navigational map to guide the user to a destination (intent).

Context can also be extracted from content-trails such as a series of images taken during an event.

### B. Representation

There are several efforts underway to represent context information ontologically. For instance, spatial ontologies such as OpenCyc [9] and RCC [10] provide ways to express spatial relations and for qualitative reasoning based on spatial coordinates. In [12], the authors provide an ontology to integrate geometric, set-theoretic, graph-based and semantic models of location. However, there are no standard, universally agreed ontologies for many concepts. The SOUPA Ontology [11] is an effort to create a Standard for Ubiquitous

and Pervasive Applications. This is somewhat similar to the notion of Upper Level Context Ontology (ULCO) that is adopted by the Semantic Space. The ULCO is easily extendible, and in our current implementation, we extend the ContextEntity → ComputingEntity → Application hierarchy to create one entity specific to the SnapToTell application, where <SnapToTell, usesContext, Location> , and in its vocabulary, the term "zone" is mapped to Location → OutdoorLocation → Region in the Semantic Space using the owl:equivalentProperty attribute.

The Web Ontology Language [18] or OWL provides a formal and expressive means to represent the instances, concepts, relations and axioms in any domain. The integration of any upcoming standard context ontologies into the Semantic Space can be done easily, facilitated by the Jena [17] ontology engine, which handles semantic web constructs expressed in OWL.

The context related to the content can be expressed using the Dublin Core [23] and FOAF [22] vocabularies. Image capture context such as aperture, shutter-speed etc. as well as Dublin Core attributes and FOAF attributes can be extracted from the Extended Image File Format or EXIF [24] headers that are now part of every JPEG image.

## III. THE PERVASIVE MEDIA FRAMEWORK

In our work, we incorporate a content analysis framework, which we call "Pervasive Media Framework", within an existing context infrastructure (Fig. 1) thereby allowing content and context to inform each other. We integrate our content analysis algorithms with the Semantic Space [19], a pervasive computing infrastructure that supports context representation, semantic querying, and reasoning within a so-called smart space. An instance of a Semantic Space would be created within the vehicle. Context sources that are represented within a smart space include low-level sensors such as motion and pressure sensors, and devices, such as the In-Vehicle GPS receiver, Web Service agents [21], and Upper-level Context, that includes UserProfile and Activity context. Networking between different semantic space instances is the topic of other research.

The context-aware Content Analysis enGine, or CAG[1], sits at the core of our media framework.

The context-aware CAG plays two important roles –

- *Context Consumer:* The CAG *uses* available context while analysing the content, to specifically "detect" certain entities. In the context of the Activity "Visit to a tourist spot" undertaken by the User "Donald", the CAG "looks" for specific persons, e.g Donald and his nephews, based on User and Activity context, and Tourist Sites such as the Merlion, Underwater World etc, based on Location context.

  The application extracts these desired contexts from the CKB using RDQL [16], which supports querying over semantic models. A sample RDQL query based on the triple

---

[1]In Fig. 1, the CAG is represented by the Image Matching and Annotation Engine
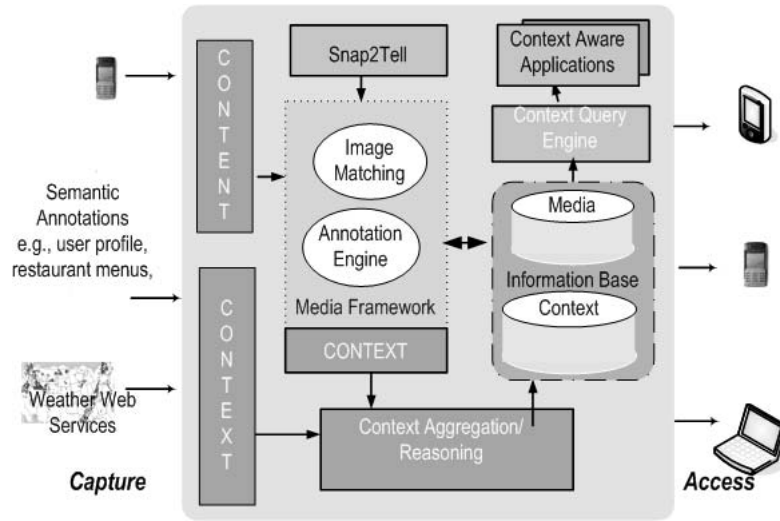
Fig. 1. Context and Content Infrastructure

pattern `<subject,predicate,object>` is shown in Listing 1.

Listing 1. Sample RDQL Query sent to the CKB

```
SELECT ?shop, maxDetour, ?currentLoc
WHERE(?shop, <rdf:type>, <FreshProduce>)
  , (?shop, <shopLocation>, ?l1)
  , (vehicle, <rdf:type>, ?currentLoc)
AND
  routeCalc(l1, currentLoc) < maxDetour;
```

- *Context Producer:* The CAG *generates* context by analysing the content, and identifying, with higher confidence-levels, the following contextual information: Donald and his nephews atop the Merlion, Huey with a Shark at the Underwater World, and so on. Content analysis generates additional context attributes for existing semantic descriptors.

Our context infrastructure uses OWL to represent and express context for modelling and sharing. The CAG, being one of the agents in the architecture, also uses the same format for the context it creates. Listing 2 shows the context markup composed when the CAG locates the Merlion landmark in an image captured by the vehicle user Tom.

Listing 2. CAG-created context instance

```
<User rdf:about="#Tom"> <locatedNear>
   <rdf:about="#Merlion/">
</User>
```

The Semantic Space uses the Jena2 generic rule engine to perform forward-chaining reasoning over the CKB. This permits developers to write first-order logic rules for a particular application based on its needs. As an example, Listing 3 shows a rule that directs the application to seek the face of a certain person in the image.

Listing 3. Rule using Activity and Location Context

```
type(?user, User), type(?activity, Picnic)
  , imageCaptureLocation(?image, ?loc1)
```
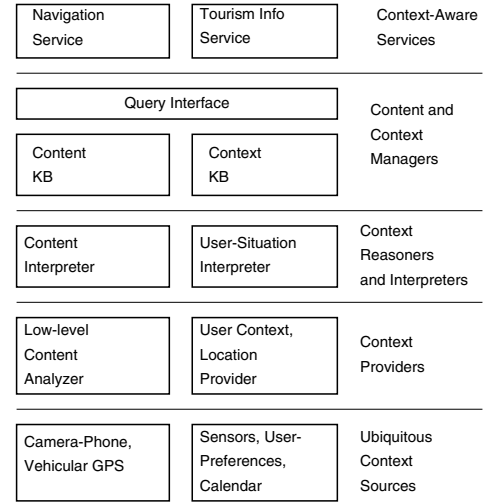


Fig. 2. Context-Aware Service Architecture

```
, location(?activity, ?loc1)
, startActivity(?user, ?time1)
, endActivity(?user, ?time2)
, imageCaptureTime(?image, ?time3)
, greaterThan(time3, ?time1)
, lessThan(time3, ?time2)
==> situation(?user, AtPicnic)
==> Check for ?user presence in ?image
```

The architecture of our framework is shown in Fig 2. Using this framework, we have built a tourist information application, SnapToTell that delivers, in real-time, annotations describing the importance of a landmark identified by the user. We find that the performance of the content-analysis engine is *significantly enhanced* by contextual cues from the infrastructure, and a synergistic relationship between content and context is observed. In section V, we also provide empirical results to show the power of position information, readily obtained from the GPS receiver on a vehicle, in enhancing the accuracy of

the image-matching algorithm.

## IV. SnapToTell: A location-aware picture-based information-access application

As described in the preceding paragraph, SnapToTell allows a user to query information about a landmark or object, based on a real image captured by a mobile device such as a camera phone. When the user seeks information about a certain monument or scene, she simply "snaps" a picture of the same with her mobile phone. The In-Vehicle GPS receiver must declare itself to be a context source and register with the Semantic Space server instance in the vehicle. The client application on the mobile phone acquires the GPS coordinates by registering a query for the same with the Semantic Space server. Location coordinates obtained from the context server are tagged to the image by the SnapToTell client to create an MMS message, which is then sent to the SnapToTell server.

In our current prototype, location context plays an important role in refining the directory search, since the image-based representation of the scenes and objects is organised according to the hierarchical classification of zones (broad-scope) and locations (narrow-scope).

Using Singapore as a test bed, we divide the map of Singapore into zones. A *zone* includes several *locations*, each of which may in turn contain a number of *scenes*. A scene is characterized by images taken from different viewpoints, distances, and possibly lighting conditions.

390 images were collected for the image-match set, which were found to pertain to 78 scenes, and these 78 scenes were further organized into 15 locations.

## V. Empirical Study on the Usefulness of Context

In our empirical study, we have investigated the extent to which location context reduces the image search space and improves performance. As our data set is limited, we have adopted the leave-one-out methodology for evaluation. In all tests, each image of the test collection is considered as a query and is removed from the collection. This image is tested for color histogram similarity matching [20] against the rest of the collection: we compute histogram similarity between this query image and all other images, and we sort the similarities in descending order. A similarity score of 100 is considered a perfect match. Fig. 3 plots the average precisions of correctly matched images over histograms of different dimensions. *Priming provided by location context* ("zone-based" and "location-based") *clearly improves the precision of color-based image retrieval* when compared to a matching query with all database images ("match-all"). In particular, more precise context information ("location-based") is better than broader contextual cues ("zone-based"). More details of the empirical study and image analysis may be found in [14].

## VI. Conclusion

We have shown how an ontology-based model can help to integrate both context and content in one unifying framework. Reasoning enabled by such a model makes it possible to
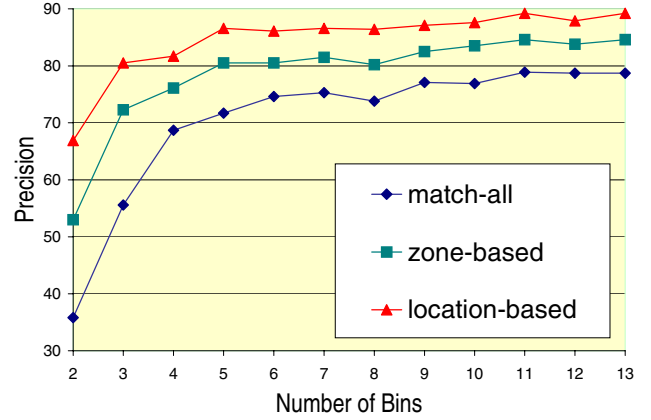


Fig. 3.   Impact of Location Context in Scene Matching

deliver a rich set of features in telematics systems. Incorporating user-specific content such as personal landmarks into the context environment, and combining with other sources of context such as user activity schedules, can help the person make efficient use of time and resources to provide a feeling of satisfaction.

Specifically, we have provided empirical data to show how location context improves the accuracy of content-analysis. Analysed content is an important source of context for user-centric telematics services.

In trials of SnapToTell , we found that the accuracy of the image-matching algorithm (color histogram) was greatly affected by the type of camera device used. In future, we intend to incorporate the characteristics of the user's camera into inference rules to improve the image-matching. In addition, weather context, such as cloud cover, dust or haze measure etc. will be used to perform adjustments on the color-histogram. Indeed, we also plan to make the image-analysis algorithm an item of "context", such that the application need not be restricted to using only the color histogram, but could choose a series of 'filtering' mechanisms, based on other context.

The end-user trials were not conducted from a moving vehicle. We feel that a gyro-stabilised camera may be needed, which is quite expensive currently. This is an area for future investigation.

## References

[1] Thomas P. Moran and Paul Dourish (eds), "Context-Aware Computing : HCI Journal Special Issue", Vol 16 2001, Numbers 2-4.
[2] Vidales P. and Stajano F, "The sentient car : context-aware automotive telematics", *Proceedings of the First European Workshop on Location Based Service, LBS-2002*, London Sept 2002.
[3] "Open Service Gateway Initiative", http://www.osgi.org
[4] Daqing Zhang, Xiao Hang Wang and Kai Hackbarth, "OSGi Based Service Infrastructure for Context Aware Automotive Telematics", *Vehicular Technology Conference* , Italy, 2003.

[5] C. Brenner and B. Elias, "Extracting Landmarks for Car Navigation Systems Using existing GIS Databases and Laser Scanning", *ISPRS Archives, Vol XXXIV, Part 3/W8*, Munich, Sept. 2003.

[6] Sorrows and Hirtle, "The nature of landmarks for real and electronic spaces", *Freksa and Mark (eds). Spatial Information Theory: Cognitive and Computational Foundations of Geographic Information Science, Springer Verlag, pp 51-64.*

[7] Risto Sarvas et al, "Metadata creation system for mobile images", *MobiSys*, Boston, USA, 2004.

[8] I. Haritaoglu, "Infoscope: Link from real world to digital information space", *Ubicomp 2001: Ubiquitous Computing: Third International Conference* , Atlanta, Georgia, USA.

[9] D.B. Lenat and R. V. Guha, "Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project", Addison-Wesley, February 1990.

[10] D. A. Randell, Z. Cui, and A. Cohn, "A spatial logic based on regions and connection. In B. Nebel, C. Rich and W. Swartout, editors, KR'92. *Principles of Knowledge Representation and Reasoning: Proceedings of the Third International Conference*, pages 165-176. Morgan Kaufmann, 1992.

[11] Harry Chen, Filip Perich, Tim Finin and Anupam Joshi, "SOUPA: Standard Ontology for Ubiquitous and Pervasive Applications", *Proceedings of the First Annual International Conference on Mobile and Ubiquitous Systems*, Boston, USA, 2004.

[12] Thibaud Flury, Gilles Privat, Fano Ramparany, "OWL-based location ontology for context-aware services", *Artificial Intelligence in Mobile Systems*, Nottingham, UK, 2004.

[13] Konrad Tollmar, T. Yeh and T. Darrell, "Ideixis : Searching the web with mobile images for location-based information", *Mobile HCI 2004 : 6th International Symposium.*

[14] J.H. Lim, Chevallet, J.-P., and Merah, S.N.,
"SnapToTell: Ubiquitous Information Access from Camera", *Mobile HCI International Workshop on Mobile and Ubiquitous Information Access*, September 2004

[15] Chakravarthy, S., Krishnaprasad, V., Anwar, E., and Kim, S.-K. "Composite Events for Active Databases: Semantics, Contexts and Detection." VLDB, 1994.

[16] L. Miller, A. Seaborne and A. Reggiori, "Three Implementations of SquishQL, a simple RDF Query Language", *Proc. 1st Int'l Semantic Web Conference* (ISWC 2002), LNCS 2342, Springer Verlag, 2002, pp 423-435.

[17] J.J. Carroll et al., *Jena: Implementing the Semantic Web Recommendations*, tech report HPL-2003-146, Hewlett Packard Laboratories, Bristol.

[18] "Web Ontology Language (OWL), W3C Recommendation", `http://www.w3.org/TR/owl-ref/`

[19] Xiaohang et al, "Semantic Space: An Infrastructure for Smart Spaces", *IEEE Pervasive Computing, July-Sept 2004.*

[20] M.J. Swain and D.N. Ballard. Color indexing. *Intl. J. C omputer Vision*, 7(1): 11-32, 1991.

[21] "A Web Services Primer", `http://webservices.xml.com /pub/a/ws/2000/04/04/webservices/`

[22] "FOAF Project", `http://www.foaf-project.org`

[23] "Dublin Core", `http://purl.org/dc/elements/1.1`

[24] "EXIF Overview", `http://www.w3c.org/2003/12/exif/`