# Technical Report

Department of Computer Science
and Engineering
University of Minnesota
4-192 Keller Hall
200 Union Street SE
Minneapolis, MN 55455-0159 USA

TR 14-022

## You Can’t Smoke Here: Towards Support for Space Usage Rules in Location-aware Technologies

Pavel Samsonov, Xun Tang, Johannes Schoening, Werner Kuhn, Brent Hecht

September 22, 2014

# You Can't Smoke Here: Towards Support for Space Usage Rules in Location-aware Technologies

Pavel Samsonov*, Xun Tang§, Johannes Schöning*, Werner Kuhn†, Brent Hecht‡§,

*Hasselt University - tUL - iMinds; §Department of Computer Science and Engineering, University of Minnesota, †Department of Geography, University of California, Santa Barbara; ‡GroupLens Research, University of Minnesota

{pavel.samsonov, johannes.schoening}@uhasselt.be, {bhecht,xuntang}@cs.umn.edu, kuhn@geog.ucsb.edu

**ABSTRACT**

Recent work has identified the lack of *space usage rule* (SUR) data – e.g. "no smoking", "no campfires" – as an important limitation of online/mobile maps that presents risks to user safety and the environment. In order to address this limitation, a large-scale means of mapping SURs must be developed. In this paper, we introduce and motivate the problem of mapping space usage rules and take the first steps towards identifying solutions. We show how computer vision can be employed to identify SUR indicators in the environment (e.g. "No Smoking" signs) with reasonable accuracy and describe techniques that can assign each rule to the appropriate geographic feature. We also discuss how our methods can be applied to large repositories of spatially-referenced images (e.g. Google Street View) to generate global-scale datasets of SURs.

**INTRODUCTION**

In 2013, a hunter started an illegal campfire that grew out of control and ended up damaging California's famous Yosemite National Park [4]. By violating a *space usage rule* (SUR) – a restriction against campfires – this hunter caused severe environmental and property damage and was a serious hazard to public safety.

SURs are not limited to constraints on campfires. Most of us encounter space usage rules frequently as we go about our day. From "no smoking" to "no fishing" to "no swimming", these rules maintain public health, enforce important laws, and protect fragile ecosystems. More generally, space usage rules are a critical mechanism through which governments and other stakeholders (e.g. landowners) manage our interaction with our environment.

However, despite their importance and ubiquity, space usage rules are absent from location-aware technologies. Schöning et al. [5] recently reported that while traditional paper maps frequently inform map readers of the space usage rules in the depicted area, no popular mobile or online map does the same. This omission is more than just a missing feature. As people become more and more dependent on their mobile devices as guides to unfamiliar spaces, the lack of support for SURs threatens to undermine the awareness of SURs and reduce their benefits.

The potential of space usage rules for location-aware technologies extends well beyond improvements to online and mobile maps: space usage rules can also enable an

**Figure 1: An example of a "no-sign" showing a *space usage rule* (SUR), specifically "no dogs allowed".**

*entirely new class of context-aware applications*. For instance, it is easy to imagine a space usage rule-based app that tells smokers if it is legal to light a cigarette in their current location and, similarly, an app that tells hunters where it is okay to start a campfire.

One can also easily imagine straightforward algorithms that provide routing instructions to help dog owners avoid "no dogs allowed" areas when walking their dogs and those that generate vacation recommendations for specific areas that allow activities of interest (e.g. climbing, fishing, diving, swimming, etc.). Along the same lines, SUR-based technologies could also help people negotiate complex issues with spatial components, such as laws that regulate where one can bring a concealed weapon in several U.S. states and laws that restrict the flight area of personal drones.

However, before online/mobile maps can match the feature set of paper maps when it comes to space usage rules and before novel SUR-based technologies can be developed, a critical problem must be solved: *space usage rules must be mapped*. As we will show below, outside of a very small number of OpenStreetMap (OSM) tags, no dataset of SURs currently exists.

In this paper, we introduce the first techniques for the widespread mapping of space usage rules. Doing this mapping accurately and on a global scale is a challenging task that will require a variety of approaches. The goal of this paper is to demonstrate that the time is right to begin addressing the SUR mapping problem by demonstrating the feasibility of one family of approaches: those informed by computer vision. The objective of our computer vision-based technique is to mine publicly available geotagged photos for "no-signs" such as those shown in Figure 1 ("no dogs allowed") and Figure 2 ("no biking", "no campfires", etc.) and map the corresponding space usage rules to the appropriate spatial regions. While we draw on existing work for several components of our computer vision approach, we are the first to combine them, especially for the detection of "no-signs".

One major benefit of computer vision-based approaches is that they can be used on large datasets of geographically-referenced imagery, particularly those available to Google and Microsoft in their Streetview and Streetside corpora. While we show below that using publicly available image datasets allows us to increase the number of SURs available in OpenStreetMap by a significant amount, our methods below have been developed with an eye towards these much larger corpora and the exponential increase in SURs that would result if similar methods were applied to them.

To summarize, this note makes the following contributions:

1. We introduce and motivate the *space usage rule mapping problem*, discussing the need for SURs in existing technologies (e.g. mobile maps) and highlighting technologies that could be enabled with a large-scale dataset of space usage rules.
2. We report the results of a small survey of SURs, finding that rules can be complex and can rely on indicators in the environment (e.g. "no-signs").
3. We show that the only dataset of usage rules – SUR-like "tags" in OSM – is very limited in size, scope, and geographic extent (e.g. there are only 7 places tagged with "no-fishing" in all of North America).
4. We introduce a computer vision-based technique that can map space usage rules by automatically detecting "no-signs" (i.e. indicators of usage rules) in geotagged photos. We also identify straightforward approaches to assign the corresponding usage rules to the correct spatial feature in OpenStreetMap.

## SURVEY OF SPACE USAGE RULES

In order to inform the design of our large-scale space usage rule mapping techniques, we conducted a survey of SURs. Because Schöning et al. [5] found many space usage rules on maps of public parks, we surveyed the websites of 25 well-known parks in urban locations and 25 well-known parks in rural locations using Wikipedia's lists of parks articles (e.g. "List of national parks of the United States"). 4-5 parks from each populated continent were selected for both types.

Overall, we found that there was an average of 7 SURs listed per urban park website and 6.9 per rural park website. Over half (56%) of these rules applied not to the entire park, but to specific places in the park, and many of the rules were restricted to *types of areas* (e.g. paths, grass areas), adding complexity to SUR mapping efforts. In addition, in certain cases, the areas in which rules applied were not fully specified on the website. For instance, the website for New York City's Central Park lists several specific places where dogs must be leashed and adds that the rule also applies in "other areas where signs requiring dogs to be leashed are posted." Examining this "where posted" phenomenon in more detail, we found it to be somewhat common. For instance, in the U.S. state of Minnesota, a business owner can ban people from carrying guns into her/his business by posting a sign.

The complexity and non-specific nature of official usage rules led us to our computer vision-based "no-sign" detection approach as our first SUR mapping effort. This approach is robust against the "where posted" phenomenon, and can capture structured rules (e.g. park- or city-wide rules) as well in many cases. In the discussion section, we highlight other possible SUR mapping approaches – e.g. crowdsourcing using a custom data entry interface – and how they can complement the techniques described here.

## USE OF USAGE RULE TAGS IN OSM

As noted above, the only existing dataset of space usage rules of any size is embedded in OpenStreetMap. These rules are encoded by OSM contributors through the use of "tags" that are applied to spatial features. In order to understand the coverage of these tags, we examined the tags on all spatial features in the entire global OSM dataset.

In our analysis, we focused on three SUR tags in particular: no-dogs, no-smoking, and no-fishing. These tags were selected as they were the top-used common SUR tags in our OSM dataset. The results of our mining of OSM for our three tags of interest can be seen in Table 1.

The most important takeaway for Table 1 is that very few SURs are encoded in OSM. By far the most common of the three tags is no-smoking, but there are only 13,976 spatial features *total* in all of OSM that have this tag. This represents less than 0.0006% of all features in OSM. The situation is even sparser for the other tags; only 57 features in the entire world have been tagged with no-fishing.

Table 1 reveals that existing approaches based on crowdsourced volunteered geographic information have failed thus far to generate a dataset of SURs of a useful size. This means that more sophisticated approaches like our computer vision-based technique are needed to develop

| Region | no-dogs | no-fishing | no-smoking |
|---|---|---|---|
| Europe | 137/1738/338 | 0/5/44 | 6858/37/3996 |
| Asia | 1/0/0 | 0/0/0 | 586/1/153 |
| North America | 5/472/11 | 0/4/3 | 1131/0/411 |
| South America | 0/4/2 | 0/0/0 | 214/0/43 |
| Central America | 0/0/0 | 1/0/0 | 68/0/58 |
| Africa | 0/1/0 | 0/0/0 | 107/0/90 |
| Australia | 5/20/1 | 0/0/0 | 171/0/52 |
| **Overall** | **148/2235/352** | **1/9/47** | **9135/38/4803** |

**Table 1: OSM SUR tag distribution for points/lines/polygons.**

the global or semi-global SUR dataset necessary to support SUR-based applications. It is important to note that more targeted crowdsourcing efforts may yield better results, something that we touch on in the discussion section.

## MINING USAGE RULES FROM GEOTAGGED PHOTOS

The goal of our computer-vision based approach is to identify "no-signs" in geotagged photos like those in Figures 1 and 2 and assign the corresponding SURs to the correct spatial features in OSM. In this section, we cover the first portion of this task – identifying "no-signs" – and the following section discusses our approaches to the SUR assignment problem.

To develop our computer vision approach, we used a dataset of geotagged Flickr images. We focused specifically on the three SURs above, which means we concentrated on finding "no dog", "no fishing", and "no smoking" signs in Flickr images. We developed datasets of Flickr images for these three tasks by downloading all geotagged Flickr images that had the terms "no dog", "no fishing", or "no smoking" in their titles, descriptions, and/or tags. In this way, we acquired 29,981 images for no-fishing, 28,921 images for no-dogs, and 17,268 images for no-smoking.

Once we finalized the collection of our datasets, we began developing our novel computer vision approach to "no-sign" detection. This approach occurs in two stages: (1) general sign detection and (2) the application of a sparse coding-based "no-sign" filter. It is important to point out that our work is not the first to address the more general "sign detection" problem. Several approaches based on visual salience detection (e.g. [2]) to identify speed limit signs have been proposed. However, these approaches assume that the sign is the most salient feature in an image. We found that this was not at all true for most of the photos in our three datasets and, due to this assumption violation, it was necessary to develop our own approach.

### *Stage 1: General Sign Detection Algorithm*

The aim of the first stage of our "no-sign" detection approach is to extract candidate signs from the original images, as the keyword search alone does not tell us if a "no-sign" is *actually* in the photo. Indeed, examining 3000 randomly-selected images in our three datasets, we found the "keyword-only" baseline precision to be less than 5%.

Object detection is a popular topic in the computer vision community and we rely on prominent object detection techniques for this stage of the approach. Specifically, we apply the Viola-Jones object detection framework [6] with Local Binary Pattern (LBP) [3] image features. In our implementation, 159 "no-signs" cropped from the overall images are used as positive training data, while 1,060 "sign-free" random images are used as negative training data.

| Method | "No dogs" | "No fishing" | "No smoking" |
|---|---|---|---|
| Stage 1 | .425 | .364 | .548 |
| Both Stages | .840 | .829 | .915 |

**Table 2. Detection precision for all three types of signs using the first stage of the algorithm only and both stages (general sign detection and filtering with sparse coding)**

Using this dataset of 1,219 images, a 25-stage cascade classifier was trained. Figure 2 shows some results from our sign detection and demonstrates that this algorithm has relatively good robustness when handling image illumination variation and image tilting. However, the classification performance of this algorithm was not sufficiently high. Table 2 shows the precision of this stage of our approach. While precision for all types of signs was significantly higher than the 5% baseline, it was around or below 50% in every case. As such, in order to increase the performance of our approach, we added a second stage of processing, which is described below. Images in which our general sign detection algorithm finds "no-signs" get passed on to the second stage.

### *Stage 2: Filtering with Sparse Coding*

The second stage of our computer vision approach leverages sparse coding [7]. Using the sparse coding model, we developed a non-learning algorithm for detecting different "no-signs" through which most false positives among candidate signs are pruned, while most true positives are retained.

The sparse coding model tries to sparsely represent input objects. In the case that an input object is similar to a small number of bases, the representation residual is low, otherwise, the residual is high. The set of bases in a sparse coding model, therefore, should contain instances able to cover various kinds of the objects under consideration (true positive signs in our task). The bases in our approach contain edges of standard cropped signs having varying appearances as well as varying angles. The goal was to have our model be robust to different sign designs as well as color and background variation. In total, our basis set consists of eighty bases.

Following the application of sparse coding, we found that precision significantly increased across all three types of "no-signs" (second row, Table 2). In total, after filtering out the relatively small number of false positives, our algorithm found 431 "no dogs" signs, 100 "no fishing" signs, and 638 "no smoking" signs. These photos were then passed on to the SUR-to-spatial feature assignment techniques described below. Although the total number of photos found was not enormous, if this type of approach were applied to a dataset like Google Street View (with some improvements, as described below), the number of extracted SURs would


**Figure 2: Sample results of the general detection algorithm.**

increase dramatically. In our work, we heavily biased precision over recall in order to ensure the entire pipeline – from photo to OSM tag – was effective. Future work will seek to increase recall, which was not possible to assess given the large numbers of photos involved (over 70,000).

## ASSIGNING RULES TO SPATIAL FEATURES

The final step of our computer vision-based approach is to assign the SUR in a geotagged photo of a "no-sign" to the correct spatial feature in OpenStreetMap (e.g. park region, building). In other words, the goal in this step is to take, for instance, a "no dogs" sign outside of a playground and tag the playground feature in OSM with the no-dogs tag.

To evaluate methods for accomplishing this task, we first downloaded all available OSM data in a 250m buffer around all 1,169 photo geotags, excluding the small percentage of cases (less than 10%) where there was very little OSM data (less than 2KB) in this buffer zone. Next, we searched for pre-existing no-fishing, no-dogs, and no-smoking tags in each buffer zone around a photo geotag and found 0, 1, and 51 tags, respectively. Because of the limited number of no-dogs and no-fishing tags, we focused on no-smoking tags for the remainder of our sign-to-region study.

We used the 51 no-smoking photo/tag combinations as ground truth for testing a variety of straightforward approaches to assigning a "no-sign" SUR photo to the corresponding spatial feature. It is important to note that the lack of a bigger ground truth data demonstrates there is very little overlap between geotagged photos of "no-signs" and tagged features in OSM. In fact, we found that just with our preliminary dataset of "no-signs" photos mined using the algorithm described above, we can boost the number of features with SUR tags in OSM by 15.0% for no-dogs, by 171.9% for no-fishing, and by 4.2% for no-smoking.

We evaluated four basic algorithms for assigning the no-smoking tag to the correct spatial feature (Table 3). The most straightforward of the algorithms – simply choosing the nearest OSM feature – has the highest accuracy (97.6%). This suggests that assigning point-based SUR indicators to spatial features may be straightforward.

## DISCUSSION & CONCLUSION

In this paper, we have discussed the need for location-aware systems (e.g. mobile maps) to incorporate space usage rules like "no smoking" and "no campfires", surveyed some of the new location-aware technologies that would be enabled with SURs, and focused on the key problem of mapping SURs. We also demonstrated that computer vision approaches – combined with a straightforward technique to assign SUR indicators found in photos to spatial features – can help us address this problem. However, as noted above, a number of other techniques for developing SUR datasets likely exist, and our immediate future work involves investigating some of these techniques.

| Method | Accuracy |
|---|---|
| Tag the closest OSM feature | 97.6% |
| Tag the closest OSM node | 87.8% |
| Tag the closest OSM polygon | 12.2% |
| Tag the closest polygon belonging to the categories [Restaurants, Fast Food, Cafe, Pub, Bar] (which often have no-smoking tags), else select closest node belonging to these categories | 85.4% |

**Table 3. The accuracy of various methods for assigning SURs in photos to the appropriate OSM feature.**

One promising approach involves using crowdworkers to capture SURs from official websites. We are developing an "If This Then That"[1] (IFTTT)-like interface that simplifies the encoding of complex SURs, e.g. Alaska State Parks' "Discharge of firearms within ½ mile of any developed park facility is prohibited". Major challenges include developing web mining techniques to identify SUR web pages (and to possibly automatically encode simple rules).

Finally, it is important to reiterate that while we demonstrated that computer vision approaches can be used find SUR indicators (i.e. "no-signs") in geotagged Flickr images, the most significant value of this type of approach is in its application to large corpora of spatially-referenced imagery (e.g. Google Street View). In order for this to occur, several additional challenges must be addressed. In particular, our techniques must be tested on additional types of "no-signs", recall may need to improve, and methods for distinguishing between different types of "no-signs" must be developed. Early work on this latter problem suggests that similar approaches to those above can be effective. We were able to achieve 65.1% accuracy classifying images of four types of "no-signs" (the three considered above plus "no swimming") using sparse coding techniques.

## REFERENCES

[1] Donoho, D.L. and Elad, M. 2003. Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ1 minimization. *PNAS*. 100, 5 (Mar. 2003), 2197–2202.

[2] Itti, L., Koch, C. and Niebur, E. 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE TPAMI*. 20, 11 (1998), 1254–1259.

[3] Ojala, T., et al. 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*. IEEE TPAMI*. 24, 7 (Jul. 2002), 971–987.

[4] Rogers, P. 2013. Rim Fire was started by hunter's illegal campfire. *San Jose Mercury News*.

[5] Schöning, J., Hecht, B. and Kuhn, W. 2014. Informing Online and Mobile Map Design with the Collective Wisdom of Cartographers. *DIS '14*.

[6] Viola, P. and Jones, M.J. 2004. Robust Real-Time Face Detection. *IJCV*. 57, 2 (May 2004), 137–154.

[7] Wright, J., et al. 2010. Sparse Representation for Computer Vision and Pattern Recognition. *Proceedings of the IEEE*. 98, 6 (Jun. 2010), 1031–1044.

[1] http://www.ifttt.com/