A New Context-Aware Computing Method for Urban Safety

Hyeon-Woo Kang and Hang-Bong Kang(✉)

Department of Digital Media, Catholic University of Korea, Bucheon, Gyonggi-Do, Korea
znxlwm@gmail.com, hbkang@catholic.ac.kr

Abstract. Recently, various research efforts have been made to analyze urban environments. Particularly, predicting urban safety from by means of visual perception is very important for most people. In this paper, we propose a context-aware urban safety prediction method by measuring the contexts of urban environments through visual information. In our context-aware evaluation, we define and extract positive and negative visual associations with urban safety. Then, we add these associations to a computational model of urban safety. Our experimental results show better performance than previous approaches.

Keywords: Urban safety · Context · Visual perception · Context-aware computing

1 Introduction

Recently, considerable research efforts have been devoted to analyzing urban environments [1] - [6]. In particular, research on the connection between criminal disorder and urban environments suggests that criminal behavior can be explained in terms of neighborhood disorder; this is known as the 'broken window theory'. That is, if we ignore apparently trivial disorders in the community, that trivial disorder will spread to the entire community [7], [8]. So, it is important to explore those disorders in predicting criminal behavior from urban environment data.

In previous work [2] - [6], many researchers attempted to explore street safety, wealth, uniqueness and interesting spots based on visual perception of urban environments. Salesses et al. [2] and Naik et al. [3] proposed predicting street safety using Support Vector Regression (SVR) from global features such as Histogram of Gradients (HOG), GIST and DeCAF. Ordonez et al. [4] explored perceptual characteristics of urban environments in terms of wealth, uniqueness, and safety. They also proposed computational models to jointly predict urban perceptual characteristics. Arietta et al. [5] showed predictive relationships between visual elements and city attributes such as crime rates, theft rates and danger perception. Khosla et al. [6] proposed an approach to look beyond the immediately visible urban scene using visual elements for predicting the distance to interesting places or crime rates for those places. This method of gathering predictive perceptual characteristics from visual data is similar to the tasks needed to measure the aesthetic quality of images [9] or their memorability [10].

V. Murino et al. (Eds.): ICIAP 2015 Workshops, LNCS 9281, pp. 298–305, 2015.
DOI: 10.1007/978-3-319-23222-5_37

(a) (b)

Fig. 1. (a) safe place (safety score : 6.08), (b) unsafe place (safety score : 3.03)

However, for accurate safety analysis, these approaches still need to explore appropriate semantic information in predicting their perceptual characteristics. For example, Fig. 1(a) and 1(b) from the Place Pulse 1.0 dataset [2] show a safe and an unsafe place, respectively, measured using the street-score method [3]. In this approach, Fig. 2a is same place as Fig. 1a, except that there is a car. Even though the generic car in Fig. 2a may be replaced by a police-car or a crashed car such as Fig. 2b and Fig. 2c, the safety scores in Fig. 2 are still similar to Fig. 1a. This is because the semantic information regarding cars was not reflected well in predicting street-scores. Thus, if we predict public safety for a given location, it is necessary to extract much more information from visual perception than its basic semantics. In other words, context awareness is required to analyze visual perceptions regarding urban environments.

(a) (b) (c)

Fig. 2. (a) Normal car (safety score : 2.87), (b) Police car (safety score : 3.02), (c) Crashed car (safety score : 2.88)

In this paper, we proposed a new approach to incorporate contextual information into predicting urban safety using visual perception. We extract contextual information as a combination of spatial and temporal contexts. Then, we measure the effects of context in visual perception as an index of positives and negatives. Finally, we compute urban safety using our contextual information.

2 Context Modeling

To evaluate safety scores for given places by using urban visual perception, it is desirable to adapt a safety metric to the changing context information on the given place [11]. In this Section, we will discuss the context information gained from visual perception that may affect the safety metric.

(a) Pedestrian (b) Police car (c) Crashed car (d) Night

Fig. 3. Examples of spatial and temporal context

2.1 Context Information

We define the context information of visual perception V on urban environments as the combination of spatial and temporal contexts in visual perception. It is as follows:

$$V = S + T \tag{1}$$

The spatial context S includes various objects and their relationships. In view of urban safety, variables such as people, buildings, cars, streets, surrounding environments and activities at a given place are important elements to model spatial context S.

The temporal context T refers to the time slot and relationships within it. In view of urban safety, the time slot in a given place is strongly related to crime rates. The same place may feel safe during daylight, but unsafe during the night. Therefore, these elements should be included in the temporal context T. Fig. 3 shows some examples of spatial and temporal context.

Based on this spatial and temporal context information, we have to change the urban safety metric accordingly. From the point of view of the safety metric, some context information will increase urban safety values, but other kinds may decrease it. Thus, it is necessary to extract positive and negative contextual objects for urban safety from both spatial and temporal contexts.

2.2 Positive Contextual Objects in Urban Safety

We define positive contextual objects in urban safety as police car, policeman, a crowd of people, a clean street with numerous flowers and natural light etc. If those objects or environments exist at any given spot, they make people feel safe. Therefore, it is necessary to detect those positive contextual objects when measuring urban safety by visual perception.

To compute the positive contextual effects, we re-scored the images in which positive objects were introduced. Using these data, the positive contextual effects are computed as follows:

$$P_k = (\sum_i p_{i,k} - po_{i,k}) / n_k \tag{2}$$

where n_k is the number of total images including object class k, $p_{i,k}$ is the score of the modified image, and $po_{i,k}$ is the score of the original image.

2.3 Negative Contextual Objects in Urban Safety

As negative contextual objects in urban safety, we define crashed cars, streets covered with graffiti, dark night, etc. If these objects or surrounding environments feature in a given place, people feel less safe there. Thus, these negative contextual objects should also be considered in predicting urban safety.

To compute the negative contextual effects, we also re-scored the images in which positive objects are placed appropriately. The negative contextual effects are computed as follows:

$$N_k = \left(\sum_i no_{i,k} - n_{i,k} \right) / n_k \tag{3}$$

where n_k is the number of total images including object class k, $n_{i,k}$ is the score from the modified image, and $no_{i,k}$ is the score of the original image.

2.4 Contextual Effects in Urban Safety

Using the positive and negative contexts, we can compute the visual contextual effects for safety as follows:

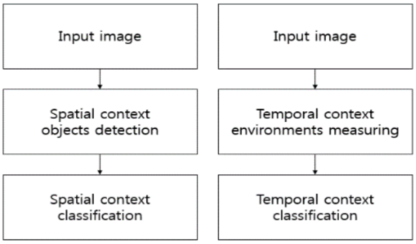
$$V = P_{k1} - N_{k2} \tag{4}$$

where k_1 and k_2 are the object class for the positive and negative contexts, respectively.

3 Context Aware Computing for Safety Prediction

3.1 Context Extraction

Fig. 4 shows our context extraction methods. Spatial and temporal context extraction from the image is performed separately and independently. In the spatial context object detection step, we extract meaningful objects from visual perception using Felzenszwalb et al's approach [20]. This approach represents objects using mixtures of deformable part models. These models are trained using a discriminative method that requires only bounding boxes for the objects in an image. This approach uses strong low-level features based on histograms of oriented gradients (HOG) [12], [13] as image representation. HOG used to create an HOG feature pyramid. The HOG feature pyramid is composed of the HOG features of each scale of a standard image pyramid. The HOG pyramid is decomposed into deformable parts and spatial relations among the deformable parts.



(a) (b)

Fig. 4. Context extraction methods: (a) Spatial context extraction method. (b) Temporal context extraction method

Learning was represented as latent SVM (discriminative learning with latent variables). Latent SVM is the general class of energy-based models. This approach used part positions and hard negatives as latent variables. This approach is a detector with good performance.

For spatial context classification, the extracted objects are classified into positive, neutral and negative contexts using the Random Forest algorithm. For example, the detected cars are categorized as contributing to a positive context (e.g., police car), a neutral context (e.g., normal car), or a negative context (e.g., crashed car).

Similarly, temporal context is also computed in two stages. In the temporal context environment measurement, a given time is determined by extracting four attributes such as daytime, dawn/evening, night and dawn from visual perception. To detect attributes related with time, we extract HOG, self-similarity features (SSIM), GIST and geometric context color histograms [17]. Using these features and image attributes regarding time at a given place, we train Support Vector Regression (SVR). Following this, we determine the time of the visual perception using SVR.

In the temporal context classification, temporal context is classified into positive and negative categories. The temporal context of noon or daytime is classified as a positive context, but dawn or evening and night is classified as a negative context.

3.2 Context Aware Urban Safety Metric

Based on context information including positive and negative context, we propose a new urban safety metric. Our urban safety metric $U_{(s,t)}$ is computed as follows:

$$U_{(s,t)} = street\_score + V = street\_score + P_k - N_k$$

$$= street\_score + \left(\sum_{k1} p_{i,k1} - po_{i,k1}\right)/n_{k1} - \left(\sum_{k2} no_{j,k2} - n_{j,k2}\right)/n_{k2} \qquad (5)$$

where s and t refer to spatial and temporal context, respectively.

4 Experimental Results

To predict urban safety from visual data using context-awareness, we computed the positive and negative contexts of that place. This contextual information is reflected in the urban safety score of a given place. Fig. 5 shows an overview of our approach. Using visual perception values, we detect spatial and temporal context information that affects urban safety. Then, we classify those contexts into positive and negative context using SVM and SVR which are trained by Place Pulse DB [2]. After that, we compute the safety score for a given place by incorporating positive and negative context information. In this Section, we will discuss some experimental results for context-aware urban safety.

For detecting contexts from visual data, we constructed a DB consisting of several categories of objects such as cars, pedestrians, different time of the day, etc. In the case of cars, the car detection average precision score from PASCAL 2006 dataset is 0.64. To detect the temporal context, we extract HOG, SSIM, GIST and geometric context color histograms. Then, we learned visual attributes such as daytime, evening, night and dawn using SVR.

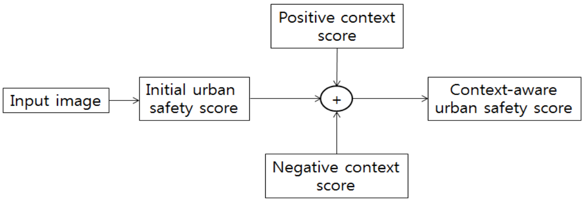


Fig. 5. An overview of our approach

Table 1 and Table 2 show weight values of each piece of contextual information relating to spatial and temporal context. To compute the weight values for spatial and temporal contextual information, we perform 3,000 crowd-sourced scoring experiments on police-car, crashed car, pedestrian and time slot. After that, we compute the weight values for police-car, crashed car and pedestrian. We also run similar experiments for time slot data. Fig. 6 shows our experimental results on the Place Pulse 1.0 dataset [2] and compares them with a previous approach. Our approach dealt with and reflected context information more accurately.

Table 1. Spatial context weights

Spatial context objects	Weights
Police-car	2.01
Crashed car	-1.76
Pedestrian	0.6

Table 2. Temporal context weights

Temporal context objects	Weights
Dawn	-1.20
Night	-2.47

5 Conclusions

In this paper, we proposed a context-aware urban safety computing method based on visual perception. We detect context information related to urban safety and extract positive and negative contexts to compute safety values accurately. By incorporating both positive and negative contexts, our urban safety computing method shows better results compared to previous approaches.

In the future, we will extend our work into a pervasive computing environment for a robust safety prediction system in which various sensor data are captured and handled. In addition, a more sophisticated high level context extraction method should be developed.

Acknowledegements. This work was supported by the Basic Science Research Program through the National Research Foundation (NRF) funded by the Ministry of Education, Science and Technology in Korea (2010-0024641).

References

1. Lynch, K.: The image of the city, vol. 1. MIT press (1960)
2. Salesses, P., Schechtner, K., Hidalgo, C.A.: The collaborative image of the city: mapping the inequality of urban perception. PloS one **8**(7) (2013)
3. Naik, N., Philipoom, J., Raskar, R., Hidalgo, C.: Streetscore-predicting the perceived safety of one million streetscapes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 793–799 (2014)
4. Ordonez, V., Berg, T.L.: Learning high-level judgments of urban perception. In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds.) ECCV 2014, Part VI. LNCS, vol. 8694, pp. 494–510. Springer, Heidelberg (2014)
5. Arietta, S., Efros, A., Ramamoorthi, R., Agrawala, M.: City forensics: Using visual elements to predict non-visual city attributes. IEEE Transactions on Visualization and Computer Graphics **20**, 2624–2633 (2014)
6. Khosla, A., An, B., Lim, J.J., Torralba, A.: Looking beyond the visible scene. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2014)
7. Wilson, J.Q., Kelling, G.L.: Broken windows. Atlantic monthly **249**(3), 29–38 (1982)
8. Keizer, K., Lindenberg, S., Steg, L.: The spreading of disorder. Science **322**(5908), 1681–1685 (2008)
9. Marchesotti, L., Perronnin, F., Larlus, D., Csurka, G.: Assessing the aesthetic quality of photographs using generic image descriptors. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 1784–1791 (2011)

10. Isola, P., Xiao, J., Parikh, D., Torralba, A., Oliva, A.: What makes a photograph memorable? In: 2011 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), pp. 145–152 (2014)
11. Bettini, C., Brdiczka, O., Henricksen, K., Indulska, J., Nicklas, D., Ranganathan, A., Riboni, D.: A survey of context modelling and reasoning techniques. Pervasive and Mobile Computing **6**(2), 161–180 (2010)
12. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. IEEE Computer Vision and Pattern Recognition **1**, 886–893 (2005)
13. Vondrick, C., Khosla, A., Malisiewicz, T., Torralba, A.: HOGgles: visualizing object detection features. In: ICCV 2013, pp. 1–8 (2013)
14. Stefano, L.D., Mattoccia, S., Tombari, F.: An efficient algorithm for exhausrive template matching based on normalized cross correlation. In: Image Analysis and Recognition, pp. 322–327 (2004)
15. Chang, C.C., Lin, C.J.: Libsvm: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) **2**(3) (2011)
16. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiely (2000)
17. Laffont, P.Y., Ren, Z., Tao, X., Qian, C., Hays J.: Transient Attributes for High-Level Understanding and Editing of Outdoor Scenes. ACM Transactions on Graphics – Proceedings of ACM SIGGRAPH 2014 **33**(4) (2014)
18. Lu, C., Lin, D., Jia, J., Tang, C.: Two-class weather classification. In: CVPR (2014)
19. Roser, M., Moosmann, F.: Classification of weather situations on single color images, intelligent vehicles symposium, pp. 798–803 (2008)
20. Felzenszwalb, P., McAllester, D., Fowlkes, C.: Discriminatively and trained, multiscale, deformable part model. In: CVPR (2008)