

Research Article

Entropy-Based and Weighted Selective SIFT Clustering as an Energy Aware Framework for Supervised Visual Recognition of Man-Made Structures

Ayman El Mobacher, Nicholas Mitri, and Mariette Awad

Electrical and Computer Engineering Department, The American University of Beirut, P.O. Box 11-0236, Riad El Solh, Beirut 1107 2020, Lebanon

Correspondence should be addressed to Mariette Awad; mariette.awad@aub.edu.lb

Received 14 July 2013; Accepted 30 September 2013

Academic Editor: Yudong Zhang

Copyright © 2013 Ayman El Mobacher et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Using local invariant features has been proven by published literature to be powerful for image processing and pattern recognition tasks. However, in energy aware environments, these invariant features would not scale easily because of their computational requirements. Motivated to find an efficient building recognition algorithm based on scale invariant feature transform (SIFT) keypoints, we present in this paper uSee, a supervised learning framework which exploits the symmetrical and repetitive structural patterns in buildings to identify subsets of relevant clusters formed by these keypoints. Once an image is captured by a smart phone, uSee preprocesses it using variations in gradient angle- and entropy-based measures before extracting the building signature and comparing its representative SIFT keypoints against a repository of building images. Experimental results on 2 different databases confirm the effectiveness of uSee in delivering, at a greatly reduced computational cost, the high matching scores for building recognition that local descriptors can achieve. With only 14.3% of image SIFT keypoints, uSee exceeded prior literature results by achieving an accuracy of 99.1% on the Zurich Building Database with no manual rotation; thus saving significantly on the computational requirements of the task at hand.

1. Introduction

Given the impressive proliferation of digital images and videos nowadays which is partly caused by the ease of acquiring them using smartphones, opportunities for novel visual mining and search applications based on image processing are emerging as hot topics for further exploration. However, because of the energy aware computing trends and the somewhat still limited processing capabilities of these handheld devices, the traditional image and video processing techniques available would require adaptation to better fit an environment where “green”, “mobility”, and “on the go” are becoming prevailing.

We present, in this paper, uSee, a supervised learning framework designed for exploitation of the physical world and delivery of context-based services after an efficient usage of scale invariant feature transform (SIFT) keypoints that have been widely and successfully applied for processing

and mining images. Different from the means available for obtaining the absolute/relative location which provide a lot of information in a geographic sense, but modestly with regards to context, uSee is implemented as an on-demand pull service based on energy aware processing of building images. Any authorized user with a handheld device having both a camera and either an internet connection or MMS settings could take a picture of a residential, commercial building, or a landmark and receive from uSee, in near real-time, touristic, advertisement, or entertainment information specific to the building of interest. From an application perspective, uSee is intended to be customized and maintained for specific cities. It would help improve touristic experience, promote visually triggered marketing methods, and even assist in ubiquitous learning as people would relate better to their environment because of the information it provides. From a workflow perspective, and to the best of our knowledge, uSee's novelty relies in its computational phases. Different

from the prior work, it employs an entropy-based metric and exploits repetitive patterns and symmetry in man-made structures to identify a signature of selected SIFT keypoints representative of the building in question using correlation and weighted clustering measures. Experimental results showed noticeable decrease in computational requirements while maintaining minimal performance degradation when compared to building identification based on a full utilization of SIFT keypoints. The remainder of the paper is organized as follows: Section 2 surveys prior work related to building and landmark recognition followed by a detailed description in Section 3 of uSee entropy-based segmentation, signature creation, and building identification phases. Experimental results are discussed in Section 4 before concluding remarks and planned research extensions are presented in Section 5.

2. Related Work

Lately, similar applications to building recognition approached the problem of effective image search mainly on landmarks. Kennedy and Naaman in [1] developed a system that generates representative images for landmarks based on community-contributed collections of images. Using tags and locations to cluster landmarks and some visual features (color and texture as global features and SIFT as local ones) of the clustered images, they selected the images with the highest connections to other images in their group as representative images. Quack et al. [2], instead, exhaustively gathered geo-tagged data by first dividing the earth's surface into tiles. They then downloaded the geo-tagged photos related to each tile along with their tags and timestamps and used text and visual features as well as speeded up robust features (SURF) to group photos within a given tile. After the clustering stage, they proceeded to automatically annotate the clusters and finally, they linked them to related articles on the internet. In [3] Zheng et al. used travel guide articles in addition to geo-tagged images, hence, combining set of images from 2 distinct sources to mine for landmarks. They, next, clustered the images gathered based on direct keypoint matching. Although [1–3] used SIFT keypoints, none of them suggested to improve on the computational requirements in these image processing problems by reducing operational complexity.

From a general classification perspective with respect to building recognition algorithms, the related work available in the literature is also different from the proposed uSee. While earlier efforts relied on combinations of global features and attempts to define effective local features, the most recent research focused mostly on invariant local features. Pass and Zabih in [4] used a joint histogram approach that combines global features such as color, edge density, texture, gradient magnitude, and rank features. In [5], Shao et al. achieved a recognition rate of 94.8% on the Zurich Building Database (ZuBuD) [6] by using scale invariant descriptors followed by a nearest neighbor search in the database for the best match based on “hyper polyhedron with adaptive threshold” indexing. Zhang and Kosecka in [7, 8] opted for a 2-stage-approach in their recognition phase. Selecting first the top candidates

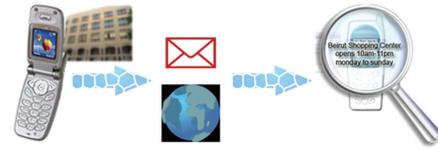


FIGURE 1: uSee pictorial workflow.

based on a localized color histogram that classifies gradient angles to the dominant vanishing points in the image, they applied SIFT features to the top candidates to detect the best match. Although they achieved 96.5% overall recognition rate, they had to manually prerotate some of the query images. Furthermore, the query images of the ZuBuD database do not pose major changes in views or illumination. Authors in [9] applied a global set of features invariant to illumination and clutter changes to extract the best candidates, thus, reducing the search space and then matched in a second stage exhaustively all keypoints. However, their global features performed poorly when pose, scale, and rotation changes were involved for selecting the best candidates of a match. The works in [10, 11] proposed using visual salient regions and attention model-based filtration so that only keypoints within the region of interest are used in the matching process while dropping those in the background. In building recognition, the building itself will often cover most of the image, and thus, a salient region will not help reduce much the amount of keypoints. Our previous work on HISI [12] proposed a soft recognition approach to the processing and identification of ZuBuB buildings. Using a coarse joint histogram technique, an image captured by a mobile user with a cell phone is preprocessed to reduce the search space to an adaptive list of potential buildings, after which a weighted fusion of different SIFT maps identifies the building in question.

3. uSee Preprocessing and Identification

uSee is a supervised framework for recognizing building in an energy aware fashion that will minimize the resource requirements associated with this task. Its preprocessing and identification phases could be performed either in the cloud or locally by the mobile application. In essence, the handheld device would either:

- (a) Perform all the pre-processing tasks and update the database if needed with new buildings and/or context aware services provided by different users. This would require the device to have enough computing power and memory to store the buildings' database and an Internet connection;
- (b) send the image to the cloud for pre-processing as shown in Figure 1 which represents the current implementation of uSee. The remote server is expected to identify the building and notify the user by SMS or email with the available location-based information.

3.1. Entropy-Based Segmentation. uSee image preprocessing phase starts with generating a map that highlights the areas



FIGURE 2: (a) Original image. (b) Entropy based preprocessed image.

of the image with high variation in gradient angle. This is done by checking the variation of the gradient angles in a given window's rows and columns. After classifying the gradient angles into 9 values equally spaced on a semicircle, we convolve the gradient angle image with a kernel that will compute a product of the entropy of the gradient angles in the kernel rows and columns based on the following.

- (i) The histogram along the rows of the window is formed of size $9 \times L$, where 9 is the number of classes of gradient angles on the semicircle and L is the size of the window around the given pixel.
- (ii) Based on the number of votes in the histogram defined above, a discrete probability distribution is formed for each of the L rows of the window.
- (iii) Steps 1 and 2 are repeated similarly for the columns of the window.
- (iv) Once the column and row discrete probability distributions are formed, the row and column entropies are computed in a straightforward manner as shown in (1) below:

$$\text{entropy}_R = \begin{bmatrix} -(p_{\text{row},R-(L-1)/2}) \cdot \ln(p_{\text{row},R-(L-1)/2})^T \\ \vdots \\ -(p_{\text{row},R+(L-1)/2}) \cdot \ln(p_{\text{row},R+(L-1)/2})^T \end{bmatrix}, \quad (1)$$

$$\text{entropy}_C = \begin{bmatrix} -(p_{\text{col},C-(L-1)/2}) \cdot \ln(p_{\text{col},C-(L-1)/2})^T \cdots \\ -(p_{\text{col},C+(L-1)/2}) \cdot \ln(p_{\text{col},C+(L-1)/2})^T \end{bmatrix},$$

where R and C are the row and column of the pixel whose noise level is to be calculated, L is the size of the kernel, and $p_{\text{row},i}$ and $p_{\text{col},j}$ are the discrete probability distributions of gradient angles along row i and column j , respectively, of the kernel.

The product of the exponential of the 2 vectors entropy_C and entropy_R gives the pixel noise level as in (2) where the noise at pixel of row R and column C is incremented by the product of the exponential of the entropies. The exponential

is used to counter the logarithmic effect in the entropy as follows:

$$\text{noise}_{R,C} = \text{noise}_{R,C} + (e^{\text{entropy}_C} \times e^{\text{entropy}_R}). \quad (2)$$

The image is then divided using k-means into 2 clusters based on the pixel noise level, one with a low gradient angle variation and the other with a high gradient angle variation. The high dispersion in a window is interpreted as indicative of the presence of omnidirectional edges which are not characteristic of building edges, instead they could be the most nonstructural patches in the image such as trees, branches and bushes. Figure 2 illustrates the output of the gradient angle variation classifier using the entropy measure defined above. The SIFT keypoints of the image that are extracted and that lie in the high gradient angle variation region are eventually filtered out.

3.2. SIFT Signature Extraction. uSee exploits the embedded symmetry and repetitive patterns in man-made structures to ensure an energy aware framework. To select the most relevant keypoints, the autocorrelation matrix of the image's keypoints is generated by directly computing the inner product of the list of keypoints since they are of unit norm as shown:

$$C = [k_1 \cdots k_n]^T \cdot [k_1 \cdots k_n], \quad (3)$$

where C is the correlation matrix and k_i is the 128-dimensional feature vector of SIFT keypoint i for an image with n SIFT keypoints. C is then thresholded at Th , which is determined experimentally, and correlation clustering is then performed. uSee workflow, as shown in Figure 4, proceeds as follows.

- (i) Step 1: The SIFT keypoint b with the highest number of correlated keypoints in matrix C is selected as the head of the current cluster cc , and all m keypoints correlated to it above the predefined threshold Th will constitute the cluster samples of cluster i denoted as s_i .
- (ii) Step 2: All elements of cluster s_i are removed from n , the original set of SIFT keypoints, and the next cluster is formed similar to Step 1, with the remaining $(n-m-1)$ SIFT keypoints.

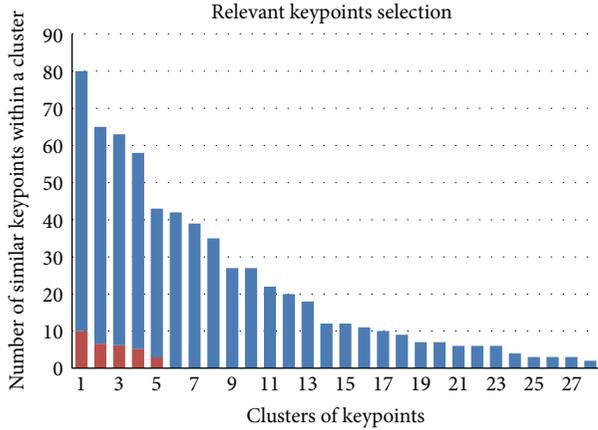


FIGURE 3: Graph of relevant keypoints selection.

- (iii) Step 3: Steps 1 and 2 are repeated until all n keypoints are clustered into c clusters.
- (iv) Step 4: The r most representative SIFT keypoints are selected based on a weighted score so as to avoid any bias towards one specific cluster of keypoints. r is predefined for any given image. In our experiments, as we shall see in Section 4, we tried out different values for r that ranged from 2.7% to 15.5% of the total n SIFT keypoints. Figure 3 shows the relevant keypoints selected by cluster in comparison to the total keypoints. Each bar represents a cluster of similar keypoints; the blue bars represent the total number of keypoints in a given cluster, and the red bars are the relevant keypoints selected within the given cluster. The ratio of the red bar to the blue one is the same as that of the blue one to the total keypoints of the image.

Note that selection of keypoints (red bars) stops when the specified limit number r is reached which, in the graph above, happens at the 6th cluster.

Each cluster s_i contributes in points to the overall signature proportionally to its cluster size such that $\sum_{i=1}^c s_i = r$ is reached, where s_i is the set of selected keypoints within a cluster (whose size corresponds to each red bar in the graph of Figure 3) and r is the limit chosen as the maximum number of selected keypoints. This guarantees a more diverse and balanced selection of the most highly correlated keypoints and will form the signature of the building.

While Figure 5 demonstrates how the SIFT keypoints of an image are reduced by uSee selection and clustering process, Figure 6 illustrates the weighting score selection mechanism of Step 4. In the left image, Figure 6 shows that r was reached within only one cluster whereas in the image on the right r is formed of SIFT keypoints from five different clusters. The map of Figure 7 clearly highlights the repetitive pattern of the corners in windows and shows where brighter dots reflect higher number of similar keypoints.

3.3. uSee Identification. When a new image is acquired, its signature is extracted and the L2 norm is computed between

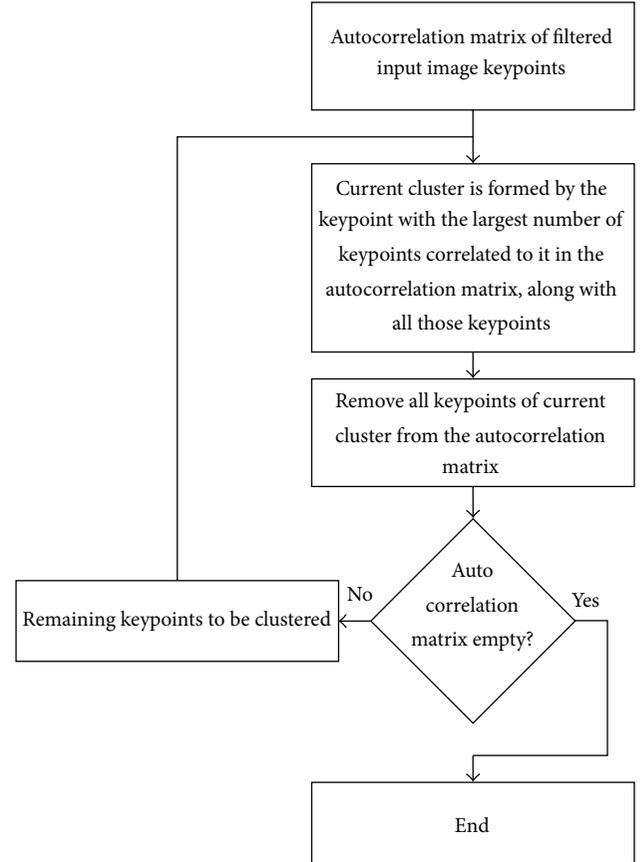


FIGURE 4: uSee SIFT keypoints clustering workflow.

its signature and the database's keypoints, or equivalently, the inner product can be computed then instead of selecting the lowest distance; the highest cosine angle is used. Building identification is based on a decision workflow that relies on a maximum voting scheme as shown below.

Assume that.

k_q is the keypoints matrix of size $[128 \times r]$ of the image to be classified,

k_d is the keypoints matrix of the database which is the concatenation of all keypoints matrices of all the database images; thus, it is of size $[128 \times (r \cdot d)]$ where d is the number of images in the database.

Then the correlation matrix formed by the inner product of k_q and k_d is given thus by $C_{q,d} = k_q^T \times k_d$, which is a matrix of size $[r \times (r \cdot d)]$.

For a given row i of $C_{q,d}$, the values in the columns represent the correlation between the database keypoints and the query's i th keypoint. The highest value in the i th row is the closest keypoint match. The corresponding image of the given column with highest value in row i gets a vote. Equation (4) maps the given column to its corresponding image.

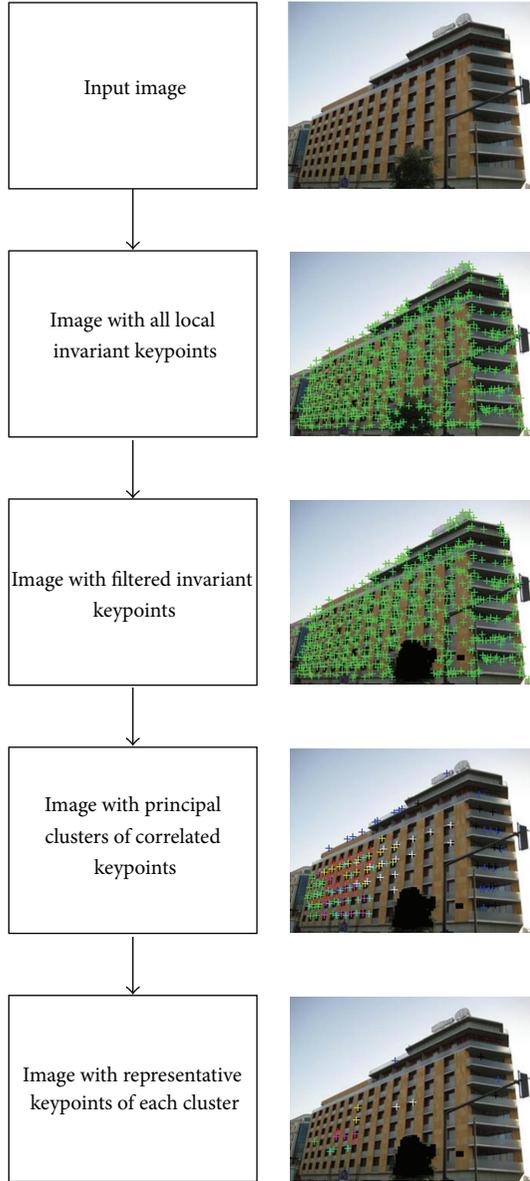


FIGURE 5: uSee selection workflow of relevant SIFT keypoints.

All in all, there will be r votes for a given query image, and the best match will be the image with highest votes as in (4) and (5) as follows:

$$k = \text{floor} \left(\frac{\max_j (C_{q,d}(i, j)) - 1}{r} \right) + 1, \quad (4)$$

$$M = \max_k V_m, \quad (5)$$

where V_m is the matching score vector of size $1 \times d$ (d is the total number of images in the database) in which the k th element of V_m is the number of votes that image k of the database received, and the best match M in (5) is the index of V_m corresponding to its maximum value. With the current

uSee implementation, once a building is identified, context-related information stored in the database will be retrieved and sent to the user.

4. Experimental Results

4.1. ZuBuD Database. The first experiments were run on ZuBuD, a popular benchmark used in computer vision and which has a total of 1005 database images captured at a resolution of 480×640 and 115 query images downsized to 240×320 . To be able to compare with published results, we tested uSee on ZuBuD based on the query set and reported recognition accuracy as the average for the 201 buildings. Several different values for r were tested for both the reference and the query images. The average number of all SIFT keypoints based on the implementation done in [13] in a given image was observed to be 740. Table 1 shows a summary of the results for $\text{Th} = 0.9$. It can be seen from the 3rd row of Table 1 that with a selection as low as 20 relevant keypoints (about 2.7% of the total SIFT keypoints), the recognition rate is above 90%. With $r = 50$ (6.8%), the recognition rate already matches the best result reported in the literature at 96.5% [7, 8]. The remarkable efficiency of uSee can be best demonstrated by this result where the computational complexity is reduced; thus, energy minimization is maximized compared to the case when all SIFT keypoints are used in building identification.

Increasing r to only 14.3% of the total SIFT keypoints exceeded all prior results achieved, to the best of our knowledge, on ZuBuD and reached 99.1% accuracy in building recognition. Also, uSee did not require any manual rotation for any image as was done in prior research work to be able to correctly identify rotated building images. The only image that failed in this case, as shown in Figure 8, had clutter that resulted after segmentation, with a small area of the image for building recognition.

The reduction in operational complexity that uSee allows for is substantial. When a query image is presented to uSee, the correlation matrix is computed then used to compare the query keypoints against all database keypoints before tallying the votes. With all n keypoints used, the complexity of creating the matrix and finding the winners is $O(n^2)$. Using only a subset r as suggested in uSee, the ratio of computational costs is $(r/n)^2$. For instance, with 50 keypoints in the query image ($r = 6.8\%$), we save 99.54% on computing energy without affecting accuracy results. The last column in Table 1 represents the percentage of the computational energy computed as the ratio of the proposed SIFT keypoints matching computation to the total keypoints comparison, that is, when no SIFT keypoint reduction is performed. As can be seen, the savings are significant and promising.

4.2. Beirut Database. Because uSee is meant to be used as a real-life application, we decided to evaluate its performance on a more challenging set of images with more severe illumination, pose, and scale changes. Therefore, further tests were conducted on another database of buildings from the city of Beirut. The Beirut database was formed with



FIGURE 6: Selection of relevant keypoints based on one cluster (a) and 5 different clusters for better diversity in selection (b).

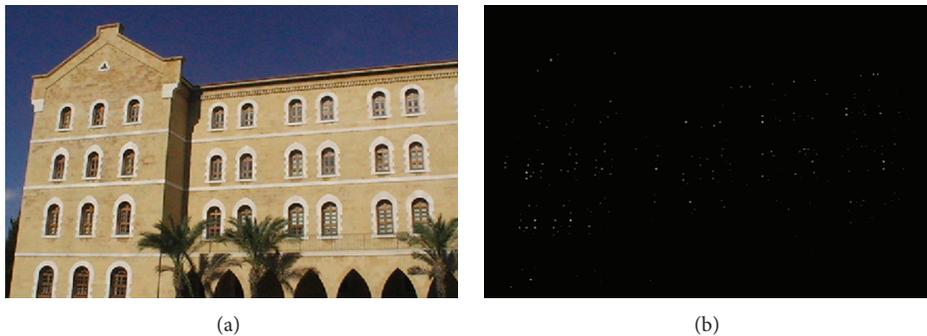


FIGURE 7: (a) Original image. (b) Map of recurrent keypoints; the brighter the dots the more similar keypoints there are to the dot.



FIGURE 8: Failed image in ZuBuD database with $r = 100$ keypoints.

5 reference images taken at the same time of day and a query image at different times and weather conditions. This means that the reference images set is more homogeneous in terms of illumination, but the test images are different from their corresponding reference in illumination, camera angle, and scale which are major image processing challenges not present in the ZuBuD database. All Beirut images were taken at a resolution of 480×640 and downsized to 240×320 . The Beirut database consists of a total of 38 buildings from downtown Beirut and from the American University of Beirut to which some context aware information was added. In order to compare uSee accuracy, we ran some live test using

a BlackBerry phone using the same approach adopted with ZuBuD, which means that the recognition rate reported is the average value computed after testing the query images for all 38 buildings.

To rule out the chance that the choice of Th is adversely affecting performance, we investigated the impact of different correlation threshold on the Beirut database. Table 2 shows the recognition rates achieved at varying values of Th . At lower thresholds, the algorithm will cluster together keypoints that are poorly correlated. At higher thresholds, it will filter out too many potential candidates of a common cluster. Hence, two similar keypoints will end up being considered different, thus, leading to poor diversity in keypoints selection for the image signature. The tests show that a value between 0.88 and 0.9 yields good results, and more specifically, 0.88 as a correlation threshold yields the best results and is optimal for the Beirut database. This threshold value results in an 84.93% recognition rate and is used for the remainder of this section.

We tested uSee performance against a random selection of SIFT keypoints. The left plot in Figure 9 shows the recognition curve plotted against the number of keypoints selected for a $Th = 0.88$. A log-scale is used for the horizontal axis to enhance the view at the lower end of the keypoints selection. The red curve shows the performance of the randomly selected keypoints, and the blue curve depicts the r SIFT keypoints selected by uSee. It can be seen as well

TABLE I: Summary of results for the ZuBuD database.

Method	Keypoints in reference images	Keypoints in query image	Ratio of selected keypoints to total keypoints	Recognition rate	Ratio of computational energy costs
Work in [2]	All in filtered region	All in filtered region	—	94.80%	n/a
Work in [1]	All in filtered region	All in filtered region	—	96.50%	0.52%
Relevant keypoints	20	20	2.7%	91.30%	0.07%
	30	30	4.1%	94.80%	0.16%
	40	40	5.4%	95.70%	0.29%
	50	50	6.8%	96.50%	0.46%
	100	75	10.1%	98.30%	1.37%
	100	115	15.5%	99.10%	2.10%

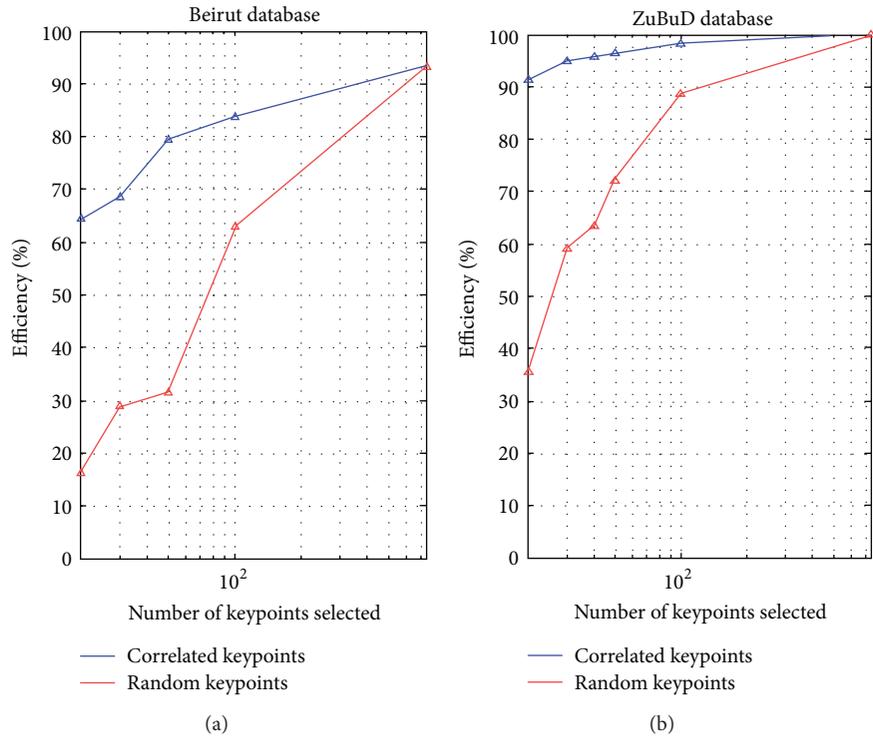


FIGURE 9: Recognition curve for Beirut (a) and ZuBuD (b) databases.

as intuitively understood that increasing r will have both curves converge towards each other and towards the maximum performance, that is, that of selecting all n keypoints. However, the more interesting part of the curves occurs at the lower end and specifically between 20 and 50 keypoints since 50 keypoints achieved the same best result reported by [7, 8] in the literature for the ZuBuD database. The bigger discrepancies between the 2 curves in the Beirut database is an indication of the higher complexity of that database, as is the fact that when all n keypoints are used, a maximum of only 93.2% accuracy was achieved.

Samples of buildings that failed to be identified by uSee in Beirut database are shown in Figure 10, while sample of the query and reference images for the ZuBuD database are displayed in Figure 11. To the left of the green line in Figure 11 is the query image that has been successfully recognized, and to its right are the reference images in the database.

The drastic changes between query and reference images of the Beirut database are shown in Figure 12. The first column in Figure 12 refers to the query images that were successfully recognized, and the 5 columns to the right of the green line represent the corresponding reference images used during the supervised learning. Note the visually apparent difficulty of the Beirut database with respect to difference in illumination, pose, and scale over ZuBuD. Despite all, the results achieved on Beirut database are quite acceptable.

5. Conclusion

We presented, in this paper, uSee, a supervised learning framework that selects in an energy aware fashion a reduced subset of relevant SIFT keypoints for image matching. With only 14.3% of an image SIFT keypoints and without any manual rotation for selected images as was done in prior



FIGURE 10: uSee failed images from the Beirut database.



FIGURE 11: Some ZuBuD query and reference images.

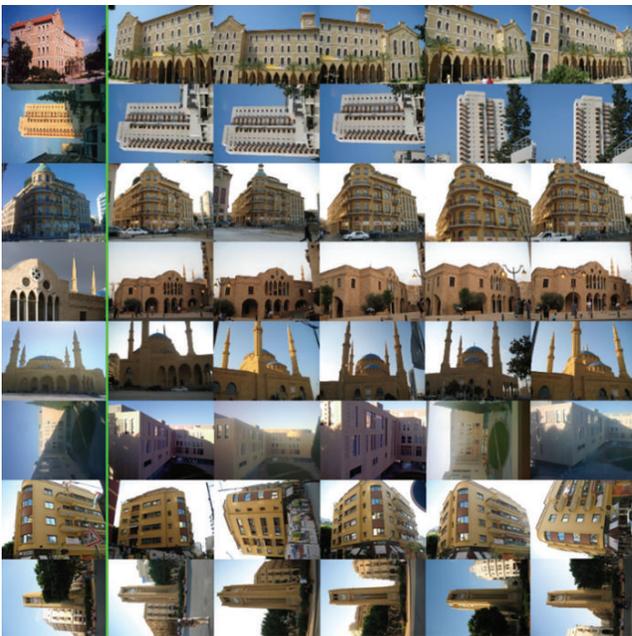


FIGURE 12: Some Beirut query and reference images.

research work, uSee exceeded published literature results and identified with an accuracy of 99.1% buildings in the Zurich Building database. Experimental results on Beirut Database showed that even with only 7% of SIFT keypoints, uSee can still deliver very good results in an energy aware paradigm,

TABLE 2: Effect of varying T_h on the recognition rate of the Beirut database.

Correlation threshold (T_h)	Recognition rate (%)
0.8	73.97%
0.85	79.45%
0.88	84.93%
0.89	78.08%
0.9	82.19%
0.91	78.08%
0.92	78.08%
0.95	73.97%

compared to the case where all SIFT keypoints are used for building recognition. With a more elaborate building segmentation in the preprocessing stage, uSee performance could be enhanced further more. Another extension of uSee that has been partially implemented is to redesign it and test it so that it completely runs on the mobile device, yet, the major future work planned for uSee is to architect it as an unsupervised learning framework and to compare it to feature saliency work.

Acknowledgment

This work was supported by the Lebanese National Center for Scientific Research (LNCSR) grant to conduct and promote research at the American University of Beirut.

References

- [1] L. Kennedy and M. Naaman, "Generating diverse and representative image search results for landmarks," in *Proceedings of the 17th International Conference on World Wide Web (WWW '08)*, pp. 297–306, April 2008.
- [2] T. Quack, B. Leibe, and L. V. Gool, "World scale mining of objects and events from community photo collections," in *Proceedings of the International Conference on Content-Based Image and Video Retrieval*, pp. 47–56, 2008.
- [3] Y.-T. Zheng, M. Zhao, Y. Song et al., "Tour the world: building a web-scale landmark recognition engine," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 1085–1092, June 2009.
- [4] G. Pass and R. Zabih, "Comparing images using joint histograms," *ACM Journal of Multimedia Systems*, vol. 7, no. 3, pp. 234–240, 1999.

- [5] H. Shao, T. Svoboda, T. Tuytelaars, and L. van Gool, "HPAT indexing for fast object/scene recognition based on local appearance," in *Proceedings of the 2nd International Conference on Image and Video Retrieval (CIVR '03)*, pp. 71–80, 2003.
- [6] H. Shao, T. Svoboda, and L. van Gool, "ZuBuD—Zurich buildings database for image based recognition," Tech. Rep. 260, Computer Vision Laboratory, Swiss Federal Institute of Technology, 2003, <http://www.vision.ee.ethz.ch/showroom/>.
- [7] W. Zhang and J. Kosecka, "Localization based on building recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p. 21, June 2005.
- [8] W. Zhang and J. Kosecka, "Experiments in building recognition," Tech. Rep. GMU-CS-TR-2004-3, 2004.
- [9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [10] J. J. Bonaiuto and L. Itti, "Combining attention and recognition for rapid scene analysis," in *Proceedings of the 3rd International Workshop Attention and Performance*, 2005.
- [11] K. Gao, S. Lin, Y. Zhang, S. Tang, and H. Ren, "Attention model based SIFT keypoints filtration for image retrieval," in *Proceedings of the 7th IEEE/ACIS International Conference on Computer and Information Science*, pp. 191–196, May 2008.
- [12] M. Awad and A. El Mobacher, "HISI: a soft framework for building recognition," in *Proceedings of the 5th International Conference on Soft Computing, Computing with Words and Perceptions in System Analysis, Decision and Control*, September 2009.
- [13] SIFT implementation in MATLAB by Andrea Vedaldi from the university of UCLA, <http://www.robots.ox.ac.uk/~vedaldi/code/sift.html>.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

