Graffiti-ID: Matching and Retrieval of Graffiti Images

Anil K. Jain
Michigan State University
East Lansing, MI 48824, USA
1-517-355-9282
jain@cse.msu.edu

Jung-Eun Lee
Michigan State University
East Lansing, MI 48824, USA
1-517-353-9319
leejun11@cse.msu.edu

Rong Jin
Michigan State University
East Lansing, MI 48824, USA
1-517-353-7284
rongjin@cse.msu.edu

ABSTRACT
Graffiti are abundant in most urban neighborhoods and are considered a nuisance and an eyesore. Yet, law enforcement agencies have found them to be useful for understanding gang activities, and uncovering the extent of a gang's territory in large metropolitan areas. The current method for matching and retrieving graffiti is based on a manual database search that is not only inaccurate but also time consuming. We present a content-based image retrieval (CBIR) system for automatic matching and retrieval of graffiti images. Our system represents each graffiti image by a bag of SIFT (Scale Invariant Feature Transform) features. The similarity between a query image and a graffiti image in the database is computed based on the number of matched SIFT features between the two images under certain geometric constraints. Experimental results on two graffiti databases with thousands of graffiti images show encouraging results.

Categories and Subject Descriptors
H.3.3 [**Information Search and Retrieval**]: Retrieval models, Search Process

General Terms
Performance, Design, Experimentation

Keywords
Graffiti, Gangs, Tagging, Content-based Image Retrieval

1. INTRODUCTION
Gangs and gang-related crimes in the United States have been rapidly growing, leading to serious social problems in large cities, prisons, and even in urban schools. It is estimated that there are at least 30,000 gangs with 800,000 members impacting 2,500 communities across the United States [1]. These gangs are responsible for up to 80 percent of crimes in communities all over the nation [2]. Furthermore, gang-related crimes are becoming extremely heinous, and the impacts of gang activities are no longer limited to only certain segments of communities. "In America's urban ganglands, and in L.A. particular, the ferocity of the thuggery has surged; gang members, their victims and police long on the gang beat tell the fighting has become more codeless, more arbitrary and more brutal than ever" and "unlike other categories of crime, gangs and gang-related crimes are spreading to formerly safe middle-class communities, or to a neighborhood near you." [2].

Figure 1. Graffiti of the Six Duce East Coast Crips. The Crips, a primarily, but not exclusively, African American gang founded in Los Angeles in 1971, is one of the largest and most violent association of street gangs in the United States, with an estimated 30,000 members. Notice the use of the basic lettering style. The spelling of six is done with a "c" to reinforce the Crip identity. The arrow is used among African-American gangs to express territoriality [5].

A gang is an organized group of individuals who collaborate for anti-social reasons. Like other organizations, a gang has a social structure that categorizes all of the members and uses recruitment techniques to bring new members into the group. Additionally, gangs provide members and their families with protection from rival gangs as well as any other perceived threats. This collective brotherhood is the main reason why people join a gang, and, as a group, they often rob, sell illicit drugs, steal cars and brutalize individuals. Representing its membership and setting up an effective means of communication among the members are essential for the success and growth of a gang. Gangs use specific clothing, brands, symbols, tattoos [3], hand signals, and graffiti to identify their group and interchange messages. Among these symbolisms, graffiti convey rich information about a gang (see Figure 1). It is the most visible form of gang criminal activity as well as a form of communication and demarcation of gang territory. Indeed, graffiti are regarded as newspapers or bulletin boards for gangs to communicate messages. Hence, recognition and interpretation of graffiti could aid in understanding gang characteristics and behavior.

Graffiti are any type of public markings that may appear in forms that range from simple written words to elaborate wall paintings. It has existed since ancient times, with examples dating back to Ancient Greece and the Roman Empire [4]. Graffiti are common sites in all the metropolitan regions in the United States and, increasingly, they have been viewed as growing problems for many cities in industrialized nations. Graffiti have been said to

MiFor'09, October 23, 2009, Beijing, China.

provide a unique insight into society, because messages conveyed through graffiti are often made without the social constraint that might otherwise limit free expression of political or controversial thoughts. In that sense, graffiti have been examined and interpreted to understand many social and cultural issues, such as adolescent personality, ancient cultures, and gang activities [5].

Figure 2. Graffiti of the African-American gang, Six Duce East Coast Crips of Los Angeles. The bottom left of the image depicts their animosity towards the police in the inscriptions "Police K" and "LAPD 187." The "K" means Killer and "LAPD" stands for the Los Angeles Police Department and "187" means murder, from the California penal code. Law enforcement agencies view these writings as direct threats [5].

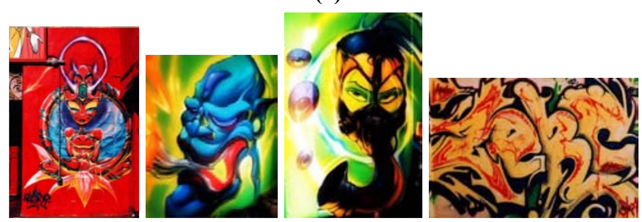
Graffiti play an important role in gang culture. Indeed, gang graffiti are also referred to as "tagging", because they are primarily composed of lines and symbols and primarily used for marking a gang's territory (see Figure 1); they warn intruders or trespassers from rival gangs and even police officers that they are not welcome (see Figure 2). Gang graffiti also transmit certain messages, symbolize a gang's power and advertise the sale of drugs. Graffiti are indeed the first indication that gang activity is present in a community. Consequently, this helps law enforcement agents to uncover the extent of a gang's territory by reading its graffiti. An accurate interpretation of gang graffiti can also assist in understanding its criminal intention in advance (see Figure 2). For these reasons, gang graffiti are not tolerated in any community. According to the *Broken Window Theory* [6], "If a broken window is left unfixed, it can quickly encourage more crime and vandalism to the neighborhood because it sends a message of indifference to observers. Graffiti is one element of the broken window theory. Once graffiti show up somewhere, if left untreated, generally more graffiti follow" [7]. Many communities have responded by creating special task force to combat and remove graffiti. As an example, the city of Riverside, California spends more that $1 million each year for graffiti abatement [7]. Indeed, graffiti are against the law in many places and punishable as a felony.

Many law enforcement agencies photograph and catalog gang graffiti patterns for the purpose of identifying gangs. When they find suspicious graffiti, a manual search is performed by checking against all the graffiti images in a database. This manual matching process is time-consuming with limited performance. To automate this process, we have designed and implemented a content-based image retrieval (CBIR) system for graffiti images, called Graffiti-ID. Since the objective of our system is gang identification, we expect the CBIR system to retrieve all the images in the database that are *visually* similar to the query graffiti image. In other words, our aim is to identify near-duplicates in the database. In contrast, most CBIR systems in the literature emphasize bridging the "semantic gap" by identifying images that are *semantically* similar to a given query [8]. To the best of our knowledge, the proposed CBIR system is the first for this application domain. We must, however, point out that Ke and Sukthankar [9] used a very small number of graffiti to illustrate pair-wise image matching using SIFT PCA-SIFT features.

2. GRAFFITI-ID SYSTEM

Given a query image, the goal of the Graffiti-ID system is to retrieve all visually similar graffiti images in the database. For a given query, database images are ranked in the descending order of their similarity to the query; the top-N (say, N=20) most similar images are first presented to the user. Additional database images are retrieved and presented to the user if so desired.

(a)



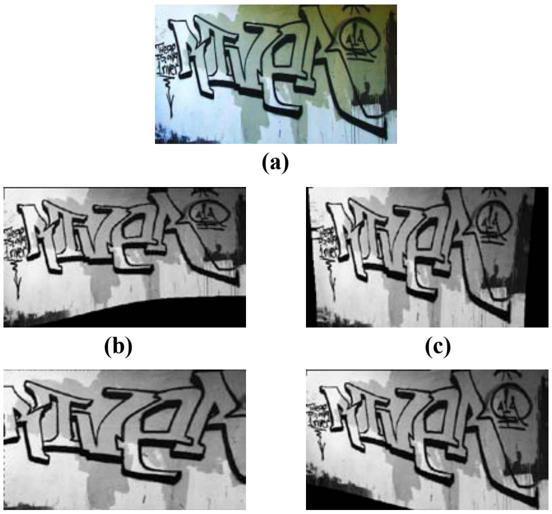
(b)

Figure 3. Examples of graffiti images in Web-DB: (a) gang-related and (b) general graffiti images [10, 11].

2.1 Graffiti Databases

Two graffiti databases, i.e., Web-DB and CAL-DB, were collected to assist the evaluation of the system.

Web-DB This database contains 1,265 graffiti images that are downloaded from the Web [10, 11]. These images are classified into two groups, gang-related graffiti, and general graffiti. Gang related graffiti, also called "tags", consist of mostly lines and letters with plain colors. On the other hand, general graffiti are elaborate wall paintings with rich colors and textures. Figure 3 shows both the gang-related graffiti and general graffiti images. There are 198 gang-related graffiti and 1,067 general graffiti images in this database. Image sizes vary from 118x90 to 1280x960 pixels. Since there are no duplicate graffiti images in Web-DB, to facilitate the evaluation, we synthesize query images by transforming each image in Web-DB. In particular, given graffiti images (from building walls or other structures) are often captured under different viewpoints, we simulate the variation in viewpoints by applying affine transformations to the downloaded graffiti images. The Thin-Plate Spline (TPS) model [12] was used for generating the affine transformations. TPS has been widely used as a non-rigid transformation model for image alignment and shape matching. To compute the transformation (deformation matrix) for an image, we map a set of four randomly selected reference landmarks, one per quadrant, to another set of four randomly selected target landmarks from the image. This deformation matrix is then applied to the entire image to create a synthetically transformed image. In order to evaluate the performance of the Graffiti-ID system, we generated four TPS deformation matrices for each of the 198 gang-related graffiti images, and applied them to the original image. As a result, a total of 792 query images were created (Figure 4).



(a)

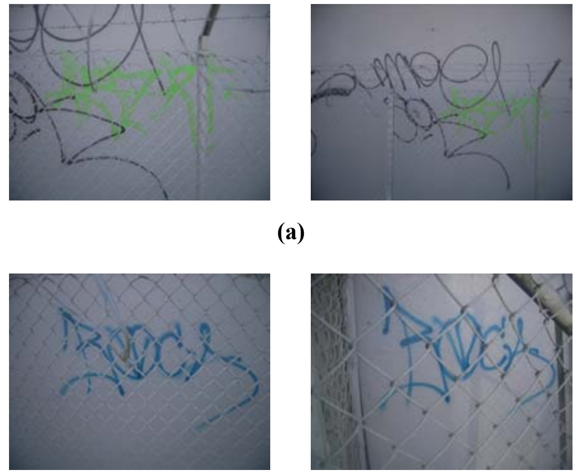
(b) (c)

(d) (e)

Figure 4. Examples of affine transformations applied to a graffiti image: (a) original image and (b)-(e) affine transformations of (a).

Figure 5. Examples of *good, bad and ugly* graffiti in CAL-DB. The red circles indicate the graffiti part (foreground).

CAL-DB This database consists of 5,000 graffiti images (1280x960 color images) provided by the city of Riverside, California. Figures 5-7 show example images in this database. This is a more challenging dataset than web-DB since images here were captured from various buildings and structures under different illumination, standoff and perspective distortion. A quality assessment of these images was made visually by the authors. Similar to the latent fingerprints in the NIST SD27 database [13], we assigned a subjective quality label to each image as "good", "bad" and "ugly"; only 60% of the images in this database are of good quality. The poor image quality is primarily due to their small size or poor imaging conditions, such as blurring and non-uniform ambient illumination (Figure 5). About 15% of the images in CAL-DB have multiple duplicates in this database, i.e., two or more instances of the same image (see Figure 6. To evaluate our system, we submitted one of the duplicates as a query and then determined if the system is able to find the duplicates of the query. Since the raw graffiti images in this database often contain multiple objects in the background, such as buildings and cars, a (manual) cropping of foreground (graffiti) from background is necessary (see Figure 7).



(a)

(b)

Figure 6. Two examples of duplicates; a duplicate indicates two different images of the same graffiti.

Figure 7. Examples of graffiti and the corresponding cropped images from CAL-DB.

(a) 266 (b) 678 (c) 247

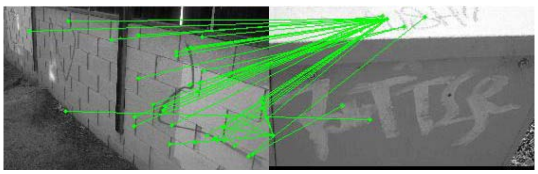
Figure 8. Examples of SIFT keypoint extraction along with the number of keypoints for three graffiti images.

2.2 Image Features

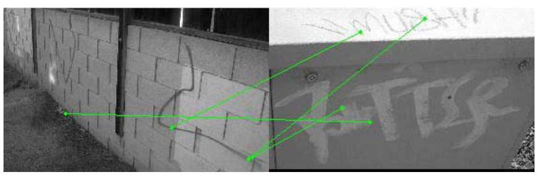
Scale Invariant Feature Transform (SIFT) [14] is a well-known local feature descriptor used for object recognition. It has been found to be highly distinctive in a number of studies [15]. SIFT extracts repeatable characteristic feature points at multiple image scales and resolutions, called keypoints. A 128-dimensional descriptor representing the texture around the keypoints is computed. The keypoints are generally invariant to image scaling and rotation, and therefore provide a robust mean for image matching across a reasonable range of affine distortion, viewpoints, additive noise, and illumination. We have empirically determined that 3 different octaves and 4 different image scales are adequate to extract SIFT keypoints from graffiti images. A threshold value of 0.015 was used to eliminate "weak" keypoints. The average (median) number of keypoints extracted per graffiti image is 860 (670) for Web-DB and 828 (652) for CAL-DB. Three examples of keypoint extraction are shown in Figure 8.

(a) 182 (b) 10

Figure 9. Matching (a) two similar and (b) two different graffiti images along with the match score.



(a) 45



(b) 4

Figure 10. Matching (a) without and (b) with geometric constraints along with the match score.

2.3 Keypoint Matching

Let $K_i = \{k_{i1}, k_{i2}, \ldots, k_{in}\}$ denote the set of keypoints detected in an image I_i. To compute the similarity between images I_i and I_j, we compare all the keypoints in K_i to the points in K_j. To determine if a key point k_{im} of image I_i matches with any key point of image I_j, we compute the Euclidean distance from k_{im} to all the points in K_j, and find the closest distance d_1 and the second closest distance d_2. Key point k_{im} is considered to have a matching key point in K_j if the ratio d_1/d_2 is small (i.e., less than 0.49 in our system).

In addition to this distance comparison, we also apply local geometric constraints to reduce the number of false matchings. Let M_{ij} represent the set of matching key points between image I_i and I_j. Then, M_{ij} can be expressed in terms of two subsets as $M_{ij} = M_{ij,T} \cup M_{ij,F}$, where $M_{ij,T}$ includes the true matching points and $M_{ij,F}$ includes all the false matching points. It is expected that removing the false matching points will increase the retrieval accuracy. Since, the SIFT descriptor is constructed as a fixed length (=128) histogram based on edge orientation, false matchings are possible in the presence of viewpoint variations or image blurring. When a key point belongs to $M_{ij,F}$, it is likely to match to many other key points. On the other hand, a key points in $M_{ij,T}$ is likely to match to one or a very small number of other key points. Given a query image Q, it is matched with all the images in the gallery database D and the number of matching points is obtained for each gallery image. Let L_m, $m=1, 2, 3, \ldots$, represent the set of keypoints in the query image in Q, that are matched to the same keypoint in a database image in D. We calculate the size of the area covered by L_m, and regard L_m as belonging to $M_{ij,F}$ if the size of the area covered by L_m (computed as the area of the polygon defined by L_m) is larger than a threshold t (0.2 in our system). All the matching key points not in $M_{ij,F}$ are regarded as true matching points. The number of key points that belong to $M_{ij,T}$ is used to define the image similarity. Figure 9 shows two matching examples, one between a pair of duplicate images and the other between two different images. Figure 10 illustrates the efficacy of the geometric constraints by comparing two different images. Note that the match score is significantly reduced from 45 to 4 after applying the constraints.

3. EXPERIMENTAL RESULTS

To evaluate the retrieval performance of the Graffiti-ID system, one of the images from the duplicate graffiti pair is used as a query to retrieve its mate(s) in the database. For Web-DB, we use one of the affine transformed images as a query with the goal of retrieving the corresponding untransformed image (see Figure 4). As we mentioned earlier, this experiment consists of 792 query images that are searched against 1,265 gallery images. Since the CAL-DB already contains many duplicates of the same graffiti (see Figure 6), we use one of the duplicates as the query. The retrieval experiments were done in a leave-one-out fashion in which 200 queries were searched against a gallery of 4,999 images in CAL-DB. The cumulative matching characteristic (CMC) curve [16] is used as the evaluation metric. Figure 11 shows that for Web-DB, the system achieves an accuracy of 85.9% at rank-1 and 94.4% at rank-20. The false matchings are mainly due to the severe distortion introduced by the affine transformation (see Figure 12). The retrieval accuracy is substantially lower on CAL-DB; 65.2% at rank-1 and 70% at rank-20. This reveals both the limitations of the current approach as well as the difficulty in matching images in CAL-DB compared to that of Web-DB. The current matching procedure tends to produce false matches when there is a significant difference in the size between a query image and a database image (see Figure 13). This is because the number of extracted keypoints is proportional to the image size. Further, SIFT features are not invariant to drastic changes in viewpoints (see Figure 14) as observed in CAL-DB. Some retrieval examples are shown in Figure 15.

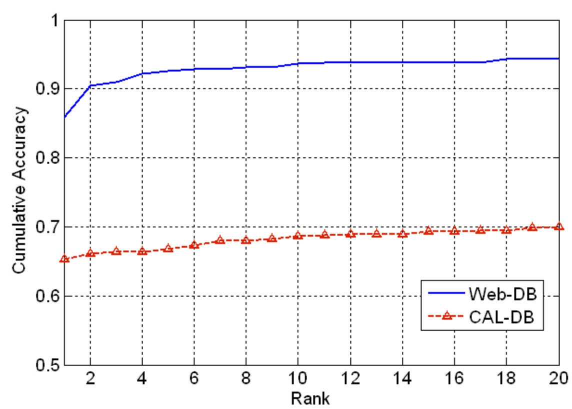
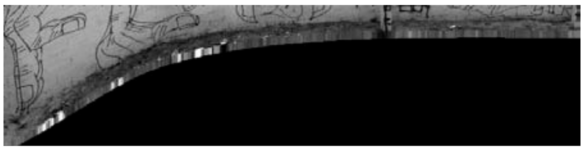


Figure 11. Cumulative Matching Characteristic (CMC) curve

4. CONCLUSIONS AND FUTURE WORK

We have presented a CBIR system for graffiti image search. Such a system is extremely useful to law enforcement agencies, particularly in large metropolitan areas that are affected by gang activities. Graffiti-ID system presented here computes the similarity between two graffiti images based on matching their corresponding SIFT key points. Experiments on two graffiti databases (one downloaded from the Web and the other provided to us by the city of Riverside) show promising results. We also identify several limitations of the current system, mostly arising from the nature of graffiti images captured under difficult imaging conditions (severe perspective distortion and large standoff). Our ongoing work is focused on (i) developing more robust features and matching algorithms, (ii) expanding the current database by including diverse graffiti images, (iii) investigating new ranking methods utilizing the keypoint matching scores, and (iv) adopting an indexing scheme to improve the retrieval efficiency of our system for large databases [17].

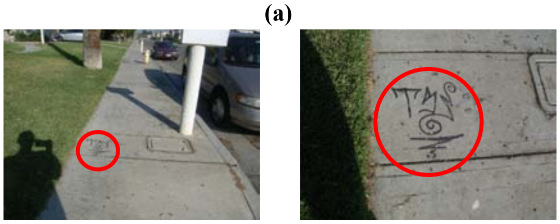


(a)

(b)

Figure 12. An example of difficult matching for Web-DB (a) query and (b) the correct duplicate was retrieved at rank 42

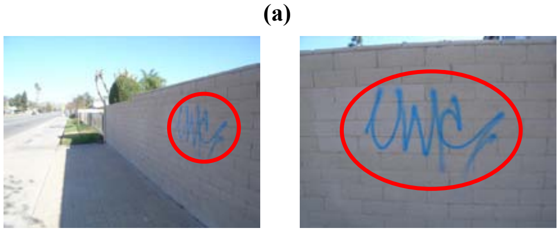
(a)



(b)

Figure 13. Examples of large size differences between two duplicate image pairs. The red circles indicate graffiti.

(a)



(b)

Figure 14. Examples of large viewpoint differences between two duplicate image pairs. The red circles indicate graffiti.

ACKNOWLEDGMENT

We thank Mr. Steve Reneker, CIO of city of Riverside, California for providing the gang graffiti images, and Nick Gregg for collecting graffiti images from the Web.

5. REFERENCES

[1] Johnson, K., Gang membership on the rise across U.S., USA Today, Jan. 30, 2009 http://www.azcentral.com/arizonarepublic/news/articles/2009/01/30/20090130gns-gangs0130.html

[2] L.A. Gangs: Nine miles and spreading, LA Weekly, http://www.laweekly.com/2007-12-13/news/l-a-gangs-ninemiles-and-spreading/2

[3] Lee, J-E., Jain, A. K., and Jin, R., Scars, Marks and Tattoos (SMT): Soft biometric for suspect and victim identification, Proc. Biometric Symposium, *BSYM*, 2008.

[4] Graffito, Oxford English Dictionary, Oxford University Press. 2006.

[5] Alonso, A., Urban graffiti on the city landscape, www.streetgangs.com

[6] Wilson, J. Q. and Kelling, G. L., The police and neighborhood safety, broken windows, http://www.manhattan-institute.org/pdf/_atlantic_monthly_broken_windows.pdf

[7] Graffiti vandalism in Riverside, http://www.riversideca.gov/graffiti/default.asp

[8] Pavlidis, T., "Limitations of content-based image retrieval", http://www.theopavlidis.com/technology/CBIR/PaperB/vers3.htm

[9] Ke, T. and Sukthankar, R., PCA-SIFT: A more distinctive representation for local image descriptors, CVPR, Vol. 2, pp. 506-513, 2004.

[10] Art crimes, the writing on the wall, http://www.graffiti.org/

[11] The golden age of gang graffiti, Flickr, http://www.flickr.com/photos/80643375@N00/sets/721575 94333288415/

[12] Bookstein, F. L., Principal Warps: Thin plate splines and the decomposition of deformation, IEEE Trans. PAMI, Vol. 11, pp. 567-585.

[13] NIST Special Database 27, Fingerprint minutiae from latent and matching tenprint images, http://www.nist.gov/srd/nistsd27.htm

[14] Lowe, D., Distinctive image features from scale-invariant keypoints, IJCV, Vol. 60, pp. 91-110, 2004.

[15] Mikolajczyk, K. and Schmid, C., A performance evaluation of local descriptors, IEEE Trans. PAMI, Vol. 27, pp. 1615-1630, 2005.

[16] Moon, H., and Phillips, P. J., Computational and performance aspects of PCA-based face recognition algorithms, Perception, Vol. 30, pp. 303-321, 2001.

[17] Li, F., Tong, W., Jin, R. Jain, A. K. and Lee, J.-E., An efficient key point quantization algorithm for large scale image retrieval, 1st Workshop on Large-Scale Multimedia Retrieval and Mining (LS-MMRM), 2009.

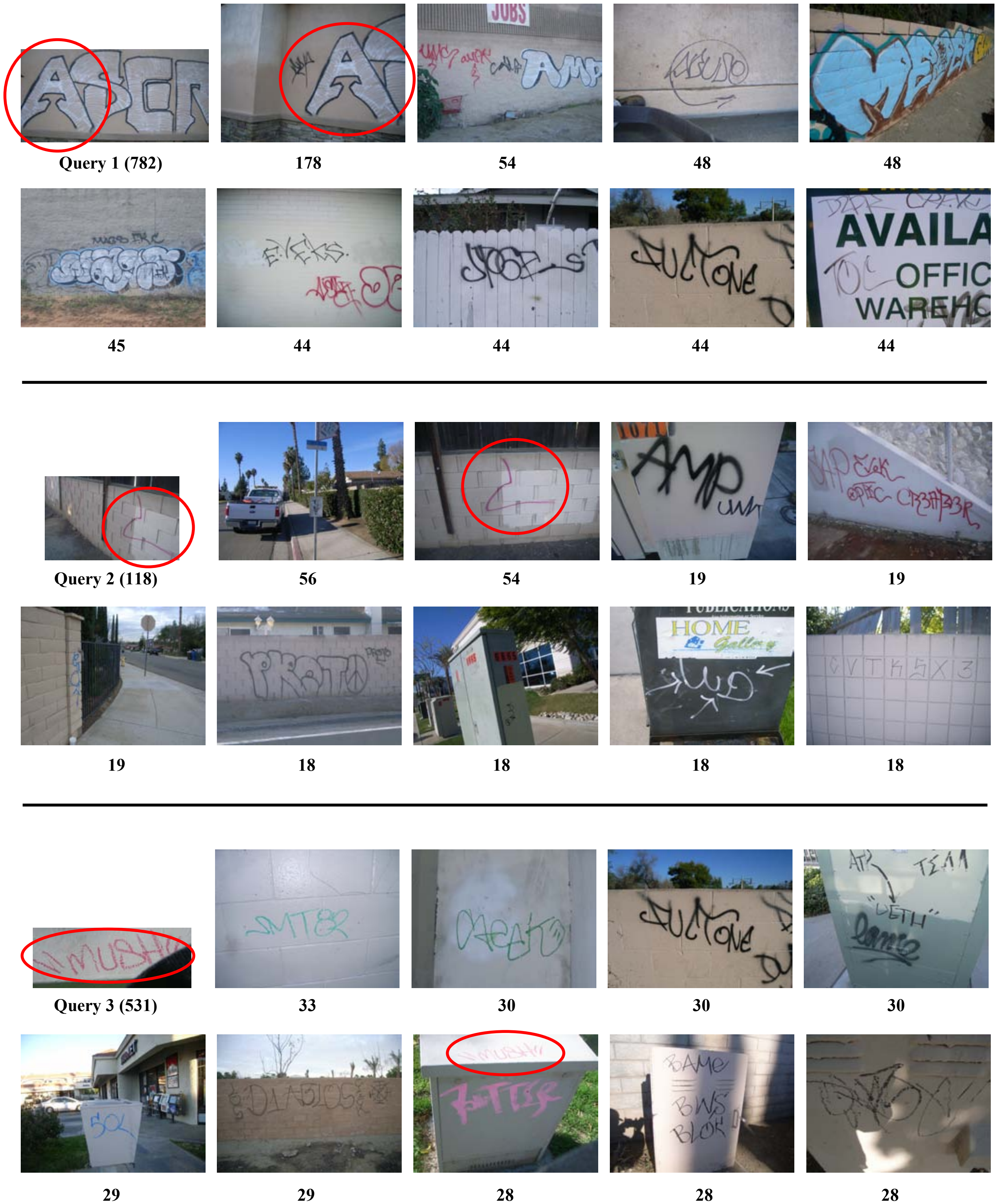


Figure 15. Three retrieval examples. Each example shows a query with the number of keypoints and top-9 retrieved images with the match score. The red circles indicate the duplicate graffiti present in the image.