Near-Duplicate Photo Detection in the Context of Events Shared on Social Networks

Thomas Steiner[1], Ruben Verborgh[2], Joaquim Gabarro[1], and Rik Van de Walle[2]
[1] Universitat Politècnica de Catalunya – Department LSI, Spain
{tsteiner,gabarro}@lsi.upc.edu
[2] iMinds – Multimedia Lab – Ghent University, Belgium
{ruben.verborgh,rik.vandewalle}@ugent.be

Abstract

In this paper, we report on the task of near-duplicate photo detection in the context of events that get shared on multiple social networks. When people attend events, they more and more share event-related photos publicly on social networks to let their social network contacts relive and witness the attended events. In the past, we have worked on methods to accumulate such public user-generated multimedia content so that it can be used to summarize events visually in the form of media galleries or slideshows. Therefore, methods for the deduplication of near-duplicate photos of the same event are required in order to ensure the diversity of the generated media galleries or slideshows. First, we introduce the social-network-specific reasons and challenges that cause near-duplicate photos. Second, we introduce an algorithm for the task of deduplicating near-duplicate photos stemming from social networks. Finally, we evaluate the algorithm's results and shortcomings.

Keywords: Near-duplicate photo detection, social networks, event summarization, media galleries, slideshows

1 Introduction

Mobile devices like smartphones, tablets, or digital cameras together with social networks enable people to create, share, and consume enormous amounts of media items like videos or photos. Mobile devices are omnipresent at all sorts of events, where—given a stable network connection—part of the event-related media items are published on social networks both as the event happens or afterwards, once a stable network connection has been re-established.

We have developed and evaluated an application and related methods [4] for media item enrichment to provide a scalable and near real-time solution to realize event summarization and media item compilation in form of media galleries. For any event with given event title(s), (potentially vague) event location(s), and (arbitrarily fine-grained) event date(s), with our approach we first extract binary media item data from social networks or media item hosting platforms and second, deduplicate exact and near-duplicate media items to then cluster similar media items for the ultimate goal of generating media galleries or slideshows. In this paper, we focus on the photo deduplication task in the context of social networks.

2 Definitions and Problem Statement

2.1 Exact and Near-Duplicate Duplicate for Photos

We define two media items of type photo as *exact duplicate* if their pixel contents are exactly the same. We define two media items of type photo as *near-duplicate* if their pixel contents differ no more than a given threshold after resampling. Examples of near-duplicate photos are scaled versions of the same photo, photos taken from slightly different viewing angles, photos with applied photo filters, slightly rotated photos, *etc.*

2.2 Social Network Challenges

Near-duplicate content in the context of social networks arises in a number of situations that we will illustrate in the following. The shown photos are examples of media items shared on social networks that were clustered correctly as near-duplicates by our clustering algorithm, detailed in Section 3.

Different Viewing Angle When two people attend the same event and take photos at roughly the same time covering the same scenes, their photos will be similar and only differ in the viewing angle. Figure 1 shows a concrete example.

Logo, Watermark, Lower Third, or Caption Insertion Oftentimes organizations or individuals insert logos, watermarks, or captions into media items to highlight their origin or convey related information, or to claim ownership of a media item.

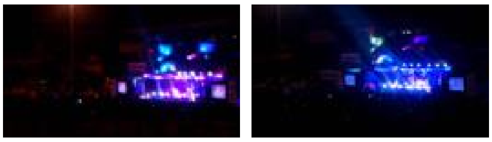
Cropping Cropping refers to the removal of the outer parts of a photo to improve framing, accentuate subject matter, or to (lossily) change aspect ratio. Cropping either happens manually via an image editing application, or, more often, the social networks themselves crop photos to have a square aspect ratio that better fits the timeline view of users.

Aspect Ratio Changes with Bulging or Stretching Aspect ratio changes can either happen combined with cropping (and thus losing parts of the photo), and/or combined with bulging or stretching (and thus deforming the photo).

Photo Filters Especially with the popularity of Instagram (`http://instagram.com/`), photo filters are a considerable reason for near-duplicate media content (see Figure 2).

3 Near-Duplicate Photo Clustering Algorithm

In this section, we detail our near-duplicate photo clustering algorithm, starting with its face detection component.



(a) Viewing angle 1 (b) Viewing angle 2

Fig. 1: Slightly different viewing angles of a concert stage

(a) Original (b) Photo filter

Fig. 2: Original, and version with an applied photo filter

Face Detection Face detection is a computer vision technology that determines the regions of faces in photos. Rotation-invariant face detection aims to detect faces with arbitrary rotation angles and is crucial as the first step in automatic face detection for general applications, as face photos are seldom upright and frontal. We use a face detection algorithm made available by Liu[3]. This algorithm is fast enough to be applied to hundreds of photos in well less than a second overall processing time on a standard consumer laptop.

Algorithm Description Our near-duplicate photo clustering algorithm belongs to the family of tile-wise histogram-based photo clustering algorithms. As an additional semantic feature, the algorithm detects faces as described above. The simplified algorithm pseudocode can be seen in Listing 1. For two photos to be clustered, the following conditions have to be fulfilled.

1. Out of m tiles of a photo with n tiles, ($m \leq n$), at most $tiles\_threshold$ tiles may differ not more than $similarity\_threshold$ from their counterpart tiles.
2. The numbers f_1 and f_2 of detected faces in both photos have to be the same, ($f_1, f_2 \geq 0$). We note that we do *not* recognize faces.

4 Evaluation

In our experiments, the algorithm performed very well. Typical social network challenges that lead to near-duplicate content can be dealt with efficiently and at scale. The algorithm had no problems with cropped, scaled, and slightly modified content by, *e.g.*, logo insertion. It handled photo filters very well and up to a certain value could cope with viewing angle variations. Some challenges with dark photos taken under low-light conditions remain. A short screencast[4] of the algorithm in action is available online.

[3] `http://liuliu.me/eyes/javascript-face-detection-explained/`, accessed 12/21/2012

[4] `http://youtu.be/jj7lTAuPiLI`, accessed 12/21/2012

```
ROWS = COLS = 10
TILES_THRESHOLD = floor(ROWS * COLS * 0.8)
SIMILARITY_THRESHOLD = 10
photos = getEventRelatedPhotos(anEvent)
distances = getDistances(photos)
clusters = {}

for outer in photos
  clusters[outer] = []
  for inner in photos
    if outer == inner continue
    similarTiles = 0
    distance = distances[outer][inner]
    for tile in distance
      if distance[tile] <= SIMILARITY_THRESHOLD
        similarTiles++
      end if
    end for
    if similarTiles >= TILES_THRESHOLD
      if faces[outer] == faces[inner]
        clusters[outer].push(inner)
      end if
    end if
  end for
end for
```

Listing 1: Near-duplicate photo clustering algorithm

5 Related Work

Work on ordinal measures that serve as a general tool for image matching was performed by Bhat *et al.* in [1]. Chum *et al.* have proposed a near-duplicate image detection method using min-Hash and term frequency–inverse document frequency (tf–idf) weighting [2]. The proposed method uses a visual vocabulary of vector quantized local feature descriptors based on Scale Invariant Feature Transform (SIFT) [3]. A method for photos and videos [5] has been proposed by Yang *et al.*

6 Conclusion and Future Work

In this paper, we have presented an algorithm targeted at the specific task of deduplicating near-duplicate event-related photos stemming from social networks. The algorithm addresses the use case very well. Future work will focus on the better handling of photos that were taken under low-light conditions.

References

1. Bhat, D.N., Nayar, S.K.: Ordinal Measures for Image Correspondence. IEEE Transactions on Pattern Analysis and Machine Intelligence 20(4), 415–423 (1998)
2. Chum, O., Philbin, J., Zisserman, A.: Near Duplicate Image Detection: min-Hash and tf–idf Weighting. In: Proceedings of the British Machine Vision Conference 2008. BMVA (2008)
3. Lowe, D.G.: Object Recognition from Local Scale-Invariant Features. In: Proceedings of the International Conference on Computer Vision. ICCV '99, vol. 2, pp. 1150 –1157. IEEE Computer Society (1999)
4. Rizzo, G., Steiner, T., Troncy, R., Verborgh, R., Redondo García, J.L., Van de Walle, R.: What Fresh Media Are You Looking For?: Retrieving Media Items from Multiple Social Networks. In: Proceedings of the 2012 International Workshop on Socially-Aware Multimedia. pp. 15–20. SAM '12, ACM (2012)
5. Yang, X., Zhu, Q., Cheng, K.T.: Near-duplicate Detection for Images and Videos. In: 1st ACM Workshop on Large-Scale Multimedia Retrieval and Mining. pp. 73–80. LS–MMRM '09 (2009)