Personal Photo Management and Preservation

Andrea Ceroni
ceroni.andre@gmail.com

Research performed at L3S Research Center in the context of the EU-funded project ForgetIT.
http://l3s.de/
https://www.forgetit-project.eu/en/home/

The ForgetIT project

A Computer that forgets?
Intentionally??
And in context of preservation???

However, nowadays we are facing:

- dramatic increase in content creation (e.g. digital photography)
- increasing use of mobile devices with restricted capacity
- inadvertent forgetting (loss of data) due to lack of systematic preservation

And: forgetting plays a crucial role for human remembering and life in general (focus, stress on important information, forgetting of details)

So: Shouldn't there be something like forgetting in digital memories as well? →

www.forgetit-project.eu

Scenario

Personal Photo Explosion

- Photo taking is fun, effortless, and tolerated nearly everywhere
- Hundreds of pictures taken during vacations, trips, ceremonies…

What to best do with all of these photos?

How to select important photos for future revisiting and preservation?

Problems

High User Investment

- Great effort in revisiting, annotating, organizing, making summaries
- Such effort increases with the size of the collections

Personal Collections become “Dark Archives”

- Photos are moved to some storage device
- Photos are rarely accessed and enjoyed again

Meeting user expectations

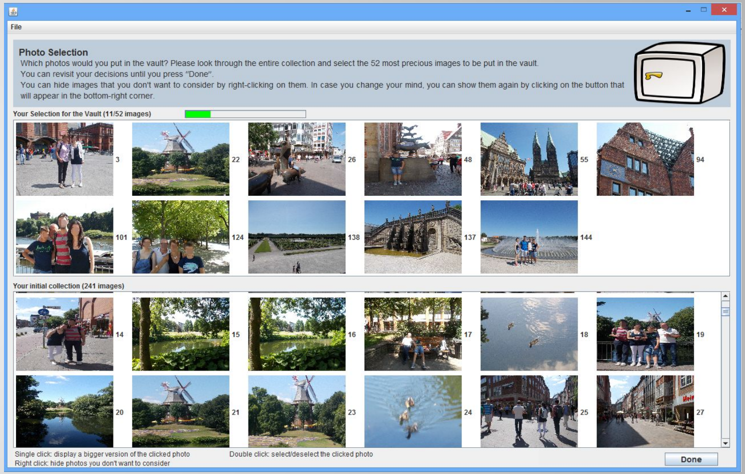
- What are the photos important to the user?
- What makes a photo important?
- Presence of personal (and hidden) attachment due to memories

Goals

- Select most important photos to keep them enjoyable and accessible
- Keep user investment low (avoiding user input like textual annotations)
- Meet user expectations and selection patterns

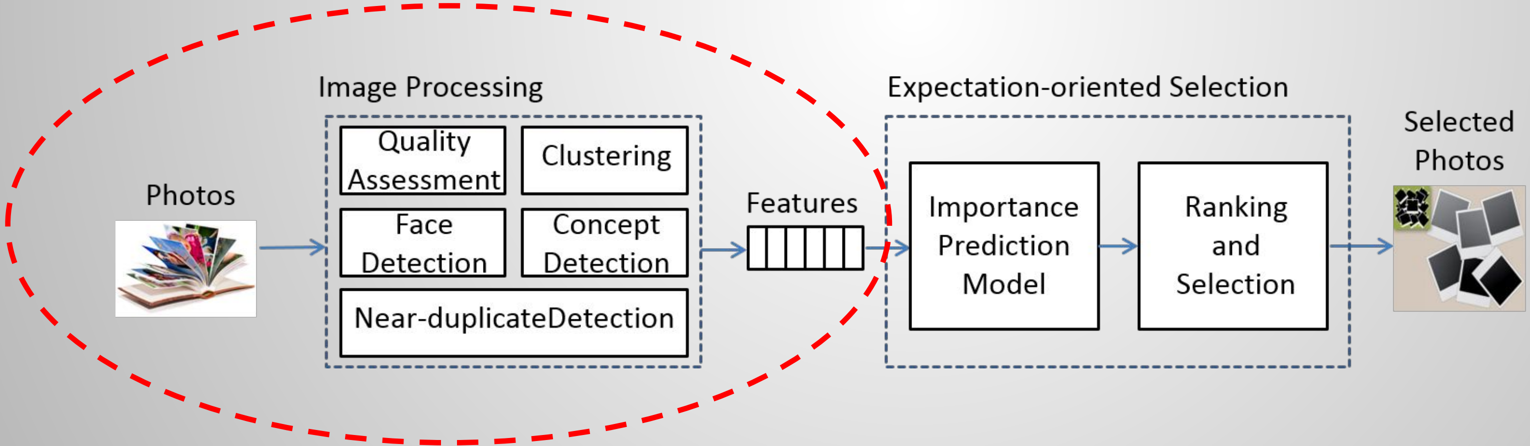
User Study

- Participants
  - 42 people
  - 91 collections
- Task definition
  - Each user provides one or more photo collections of personal events
  - Selecting 20% of photos from each collection for preservation and revisiting
- Insights
  - Image quality as least important selection criterion
  - Personal and hidden aspects rated as highly important
  - Event coverage also highly important



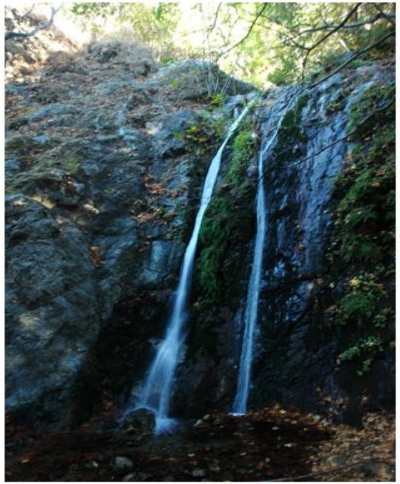
Expectation-oriented Photo Selection

- User selections from personal collections used to train the model
- Relaxed notion of coverage (features from collections, clusters, near-duplicates)
- No manual annotations or external knowledge is required



Quality-based Features

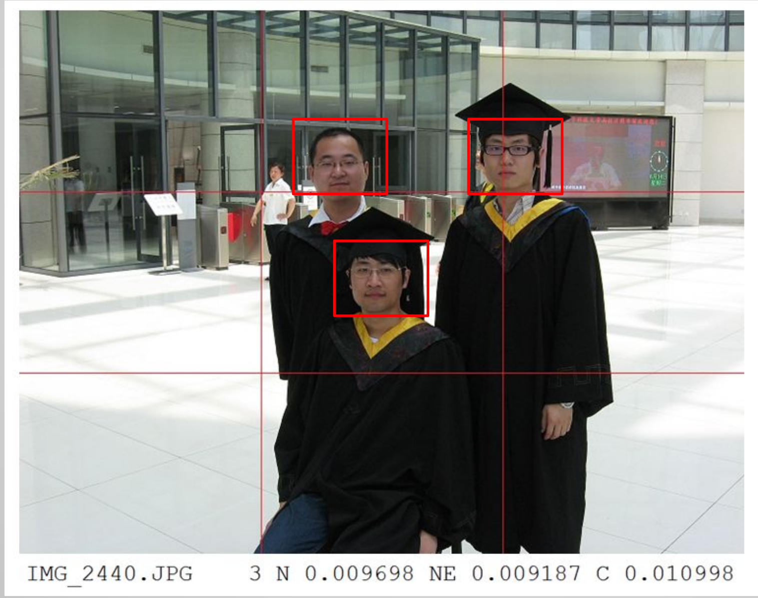
Blur, contrast, darkness, noise



	Left photo	Right photo
Blur	0.533219	0.241118
Contrast	0.157777	0.107511
Darkness	0.870238	0.433792
Noise	0.179392	0.167515

Face-based Features

Presence, position, relative size of faces in each of 9 quadrants



Concept-based Features

346 concept detectors represented by SVMs (concept set defined in TRECVID 2013 benchmark activity, 800 hours of video for training)

Top 10 concepts

- Outdoor: 0.9138
- Vegetation: 0.9
- Three_or_more_people: 0.89013
- Trees: 0.85785
- Building:0.83941
- Street: 0.81051
- Person: 0.79659
- Windows: 0.79222
- Sky: 0.76782
- Female: 0.75522

Collection-based Features

Temporal Clustering: groups of images belonging to the same sub event

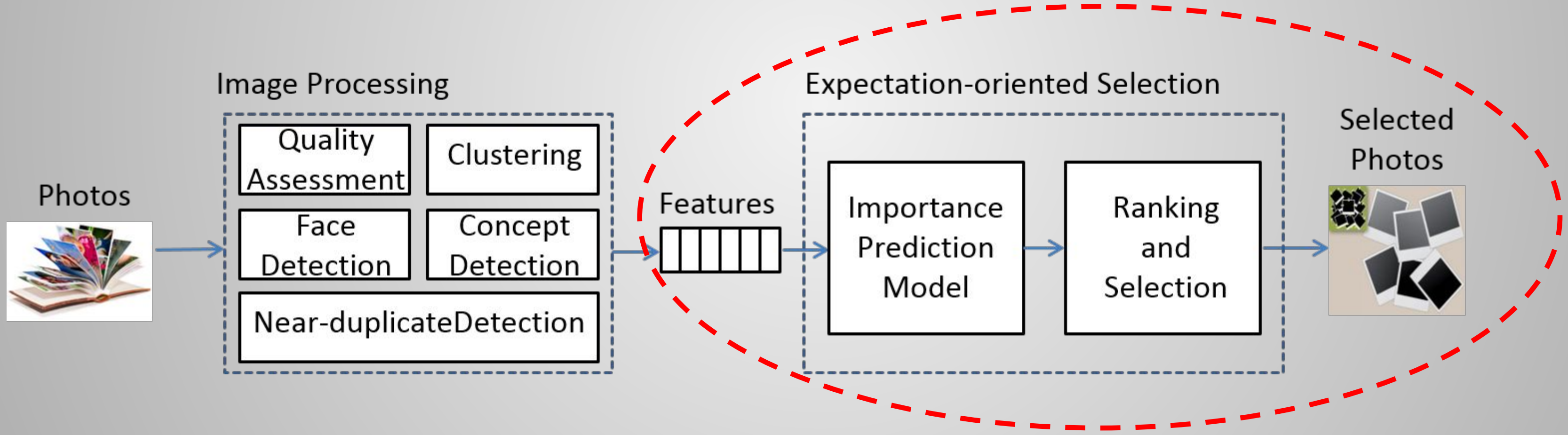
Near-duplicate Detection: identify similar shots of the same scene

Information about the clusters (sub events) and near-duplicate sets each image belongs to

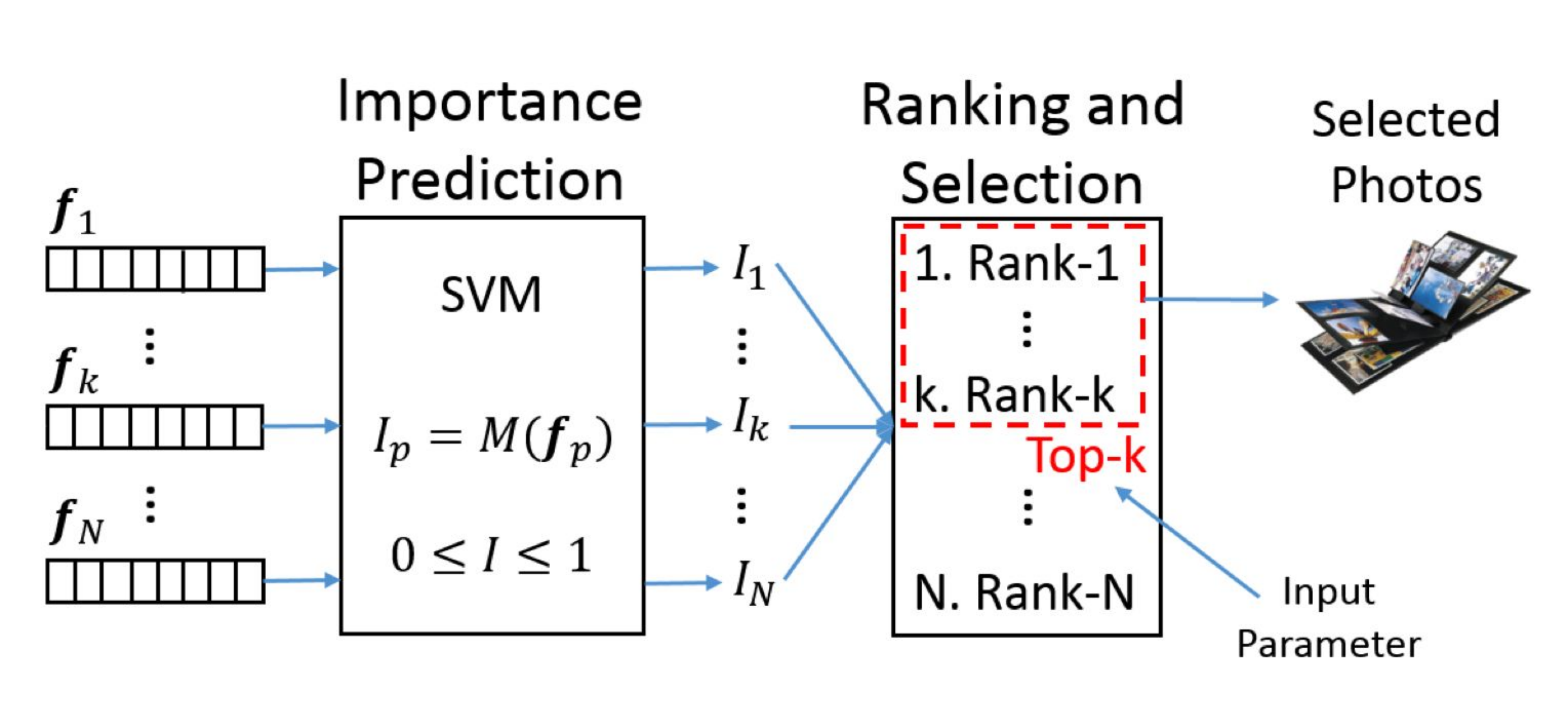
For each image:

- Size of its cluster
- Quality of its cluster (avg, std, min, max)
- Faces in its cluster (avg, std, min, max)
- Has near-duplicates?
- Size of its near-duplicates set

Expectation-oriented Photo Selection



Importance Prediction



Experiments

Dataset

- Photo collections representing events (e.g. vacations, business trips, ceremonies)
- 91 collections, 42 users, 18,147 photos
- 20% selected as most important for future enjoying/revisiting
- Each photo judged by its owner

Baselines

- Cluster → Iterate → Select (Rabbath et al., TOMM’11)
- Summary Optimization (Sinha et al., ICMR’11)

Baselines

Temporal Clustering

- Cluster photos based on time [Cooper et al., 2005]
- Iterate the clusters (round robin)
- At each round, select the most important photo according to:

$$I(p) = \underbrace{\alpha \cdot \|\mathbf{q}_{\mathrm{p}}\|}_{\text{Quality}} + \underbrace{(1 - \alpha) \cdot \dim(\boldsymbol{F}_p)}_{\text{Faces}}, \quad \alpha = 0.3$$

Summary Optimization [Sinha et al., ICMR'11]

- Compute the optimal summary of size k according to:

$$S^* = \arg\max_{S \subset P_C} F(Qual(S), Div(S), Cov(S, P_C))$$

- Qual = sum of quality and *portrait, group, panorama* concepts values of each photo
- Div = diversity within the summary
- Cov = number of photos in the collection that are represented in the summary

Results

Precision for different values of *k* and different subsets of features

Statistically significant improvement over baselines

Concepts are more discriminative than quality and faces

Modeling collection-level information as a set of features is more effective than explicitly imposing coverage

	P@5%	P@10%	P@15%	P@20%
Baselines				
Clustering	0.3741	0.3600	0.3436	0.3358
SummOpt	0.3858	0.3843	0.3687	0.3478
Expectation-oriented Selection				
quality	0.3431	0.3261	0.3204	0.3168
faces	0.4506▲	0.3968▲	0.3836△	0.3747△
concepts	0.5464▲	0.4599▲	0.4257▲	0.4117▲
photo-level	0.5482▲	0.4760▲	0.4434▲	0.4266▲
all (Expo)	0.7124▲	0.5500▲	0.4895▲	0.4652▲

Statistically significant improvements marked as ▲ ($p < 0.01$) or Δ ($p < 0.05$).

Hybrid Selection

What is the role of coverage in personal photo selection?
Can we improve the selection by incorporating coverage within the model?

- ➢ Coverage-driven Selection
  - o Cluster $\rightarrow$ Iterate $\rightarrow$ Select
  - o Still a strict model of coverage

Importance
Prediction

- ➢ Summary Optimization
  - o Compute the optimal summary:
  - o More flexible

$$S^* = \arg\max_{S \subset P_C} F\left(Qual\left(S\right), Div\left(S\right), Cov\left(S, P_C\right)\right)$$

Results

Including importance prediction as quality measure in coverage-based methods improves their performances

A strict model of coverage via clustering gets smaller benefits

Expo is still better or comparable with the Hybrid Selection models

	P@5%	P@10%	P@15%	P@20%
Baselines				
Clustering	0.3741	0.3600	0.3436	0.3358
SummOpt	0.3858	0.3843	0.3687	0.3478
Coverage-driven Selection				
basic	0.4732▲	0.4113▲	0.3902△	0.3809△
filtered	0.5351▲	0.4617▲	0.4325▲	0.4170▲
filtered+greedy	0.6271▲	0.4835▲	0.4391▲	0.4262▲
SummOpt++	0.7115▲	0.5533▲	0.4937▲	0.4708▲
Expo	0.7124▲	0.5500▲	0.4895▲	0.4652▲

Statistically significant improvements marked as ▲ ($p < 0.01$) or Δ ($p < 0.05$).

Other Directions

- Inclusion of additional features in the model
- User personalization

Additional Features

Low-level visual info

Basic visual signals that might capture the attention and interest of the observer: HSV statistics, colors, textures, lines.

DCNN Features

Image representation given by a DCNN (GoogLeNet) pre-trained to predict the 1,000 categories of the ILSVRC.

Face Popularity

Face clustering applied to compute how frequently a face appears in a collection (cluster size).

Aesthetics

How an image is well posed, attractive and pleasant to an observer: rule of thirds, simplicity, contrast, balance.

Emotional Concepts

Concept detectors of SentiBank: nouns (concepts) and adjectives carrying sentiments are combined together to associate emotions to concepts.

Additional Features

Moderate yet statistically significant improvement

Concepts (**DCNN**) and concepts (**SentiBank**) improve **concepts** features

Face popularity only slightly improves **faces** features alone

Both **low level** and **aesthetics** features are better than **quality** features

	P@5%	P@10%	P@15%	P@20%
Expo				
quality	0.3431	0.3261	0.3204	0.3168
faces	0.4506	0.3968	0.3836	0.3747
concepts	0.5464	0.4599	0.4257	0.4117
all	**0.7124**	**0.5500**	**0.4895**	**0.4652**
Expo++				
low level	0.4399	0.3913	0.3729	0.3697
aesthetics	0.4406	0.3923	0.3732	0.3639
face popularity	0.4692	0.4101	0.3977	0.3945
concepts (DCNN)	0.5694	0.4945	0.4553	0.4436
concepts (SentiBank)	0.6124	0.5172	0.4674	0.4502
all	**0.7426**$^{\triangle}$	**0.6155**$^{\blacktriangle}$	**0.5330**$^{\blacktriangle}$	**0.5121**$^{\blacktriangle}$

User Personalization

Personalized photo selection model

- Adapts to user preferences by exploiting user feedback
- Based on retraining the model every time a new annotated collection is available

Promising adaptation capabilities

- Including new annotated collections of the same user can benefit future selections
- Exploiting annotated collections from other users can alleviate the cold-start problem

Evaluation on a large number of users and collections is required to make the results more evident and significant

Applications for PhotoPrism

- Semi-automatic photo selection/summarization (fine-tuning DCNNs)
- Event-based clustering and near-duplicate detection
- Face clustering and recognition
- User personalization (selection model)
- Emotion detection as additional feature (SentiBank library)
- Low-level information (e.g. textures, colors, etc.) as additional features
- Rules of aesthetics as additional features (code in OpenImaJ library available)

For more information, visit photoprism.org
or github.com/photoprism/photoprism