Image-based Lifelogging: User Emotion Perspective

Prepared by J. Bum and H. Choo
**Definition of Lifelog**

- **Lifelog**: A digital record of everything that happens to a person
- **Lifelogging**: The action or practice of making a continuous record of one's daily activities by means of a digital device or computer application
[REF] Lifelogging Device
Motivation

- People record their lifelog in various forms.
- Especially, people record their life as photos thanks to smartphones equipped with cameras and high capacity memory.
- As the amount of photos preserved by individuals is growing rapidly, people often waste their time exploring them to remember moments.
- Thus, there is a growing need for a summary tool to help organize and browse for personal photos automatically.
  - Google Photos or Apple’s iPhoto automatically generate a summary based on people, places and recent timeline.
  - We provide a summary focusing on emotions in facial expressions such as surprise, joy and sadness.
Lifelogging system

- Basic architecture and modules

![Diagram showing the basic architecture and modules of a lifelogging system involving a mobile application, a local database, a data processing module, and a cloud service. The diagram highlights modules such as Data Collection Module, Diary Generation Module, SystemSens Plus Server, and Project Oxford (Cloud Service).]
SystemSens Plus (1/2)


- Due to resource constraints, we develop “SystemSens Plus” which is simpler operation and less power consumption.
  - Design the architecture of SystemSens Plus tool which includes: communication protocols, message format, polling time of sensor, uploading process etc
  - Develop a web server which can collect all data from clients
  - Develop new features for clients along with reducing unessential functionalities

SystemSens Plus(2/2)

- Architecture and Design

- SQLite Database / JSON message format

- Android Smartphone (client)
  - SensPlus Uploader
  - Statistic Service
  - Create SQLite Database
  - Android Service
    - Call sensor
    - Battery Sensor
    - Message Sensor
    - Application Usage Sensor

- HTTPS Post

- Server

- Battery
- Used apps
- SMS

Mobile Computing
Representative Emotional Data Extraction Scheme (REDES)

- Overall process of REDES

(a) Photostreams of a special day

(b) Finding user’s faces

(c) Clustering by similar emotions and ranking photos

(d) Displaying key photos
REDES*

1.1 Emotion Detection (1/3)

- MS Cognitive Service: Face Recognition API

* REDES: Representative Emotional Data Extraction Scheme
1.1 Emotion Detection (2/3)

- **Step 1: Face Recognition**
  1. **Create Person Group**: A person group is the container of the uploaded person data, including face images and face recognition features.
  2. **Create Person and add face**: A new person (personId & persistedFaceId) is created.
  3. **Train Person Group**: Training is a crucial step that only a trained person group can be used by 'Face – Identify'.
  4. **Identify Person**: 1-to-many identification to find the closest matches of the user’s face.
1.1 Emotion Detection (3/3)

- Step 2: Emotion detection
  - If the confidence of ‘Face Recognition’ > 0.7, we extract the faceId’s emotion fields.

```json
{
  "faceId": "d543ccdbd-9088-4794-9c9d-608d0f8e45f4",
  "faceRectangle": {
    "width": 227,
    "height": 227,
    "left": 459,
    "top": 124
  },
  "faceLandmarks": {
    "pupilLeft": {
      "x": 504.4,
      "y": 202.8
    },
    "pupilRight": {
      "x": 697.7,
      "y": 175.9
    },
    "noseTip": {
      "x": 598.5,
      "y": 250.0
    },
    "mouthLeft": {
      "x": 527.7,
      "y": 298.0
    },
    "emotions": {
      "anger": 0.103,
      "contempt": 0.003,
      "disgust": 0.038,
      "fear": 0.003,
      "happiness": 0.826,
      "neutral": 0.005,
      "sadness": 0.001,
      "surprise": 0.02
    }
  }
}
```
A photo is represented by time and emotion features.

- **Time (1D):** the data/timestamp from Exchangeable Image File Format (EXIf) header of the photos normalized to a real number between 0 and 1

- **Emotion (8D):** eight emotion scores for anger, contempt, disgust, fear, happiness, neutrality, sadness and surprise

### Example

<table>
<thead>
<tr>
<th>time</th>
<th>e₁</th>
<th>e₂</th>
<th>e₃</th>
<th>e₄</th>
<th>e₅</th>
<th>e₆</th>
<th>e₇</th>
<th>e₈</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1223</td>
<td>0</td>
<td>0.001</td>
<td>0</td>
<td>0.005</td>
<td>0.0663</td>
<td>0.0055</td>
<td>0</td>
<td>0.9273</td>
</tr>
</tbody>
</table>

< 9D photo representation >
REDES

2. K-means Clustering

- Using K-means algorithm, REDES groups the photo with similar emotion and time together.
### Key photo selection (1/2)

- **Emotion** ($f_1$): The maximum value among eight emotion scores for a photo is selected as emotion measure $f_1$.

- **Blur** ($f_2$): We adopt the technique proposed by Tong et al., which employs edge type and sharpness analysis using a Haar wavelet transform.

- **Hue count** ($f_3$): The number of hue is a measure of simplicity. Draw a 64-bin histogram to measure the hue values by converting the photos into the HSV representations.

- **Brightness** ($f_4$): The closer the average brightness is to the median of Y, the higher the score.

- **Contrast** ($f_5$): We split an image into 8-by-8 blocks based on the Y component of an image when the image is represented in the YUV color space, and then pick the central 6-by-6 blocks. The sub-contrast of each block is calculated by subtracting the minimum value from the maximum.

---

**Score** ($k$) = \[ \sum_{i=1}^{5} w_i \frac{f_i(k)}{\sigma_i} \]
3. Key photo selection (2/2)

- Blur and Brightness distributions of a photo collection

Graphs showing the distribution of blur and brightness in a photo collection.

- Blur (a) Blurry photos ($B > 0.9$)
- Brightness (a) Dark photos ($Y < 10$)
- Brightness (b) Photos close to average brightness ($120 < Y < 130$)
4. Diary Generation

- REDES displays a lifelog summary and user can write down texts on the recommended photos.

(a) Summary  
(b) Lifelog diary